AAUSC 2012 Volume – Issues in Language Program Direction

Hybrid Language Teaching and Learning: Exploring Theoretical, Pedagogical and Curricular Issues

Fernando Rubio (Editor)
Joshua J. Thoms (Editor)
Stacey Katz Bourns (Series Editor)

HEINLE
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

AAUSC 2012 Volume – Issues in Language Program Direction: Hybrid Language Teaching and Learning: Exploring Theoretical, Pedagogical and Curricular Issues
Fernando Rubio, Joshua J. Thoms and Stacey Katz Bourns

Editorial Director: P. J. Boardman

Publisher: Beth Kramer

Editorial Assistant: Gregory Madan

Managing Media Editor: Morgen Gallo

Executive Brand Manager: Ben Rivera

Market Development Manager: Courtney Wolstoncroft

Executive Marketing Communications Manager: Jason LaChapelle

Rights Acquisitions Specialist: Jessica Elias

Manufacturing Planner: Betsy Donaghey

Art and Design Direction, Production Management, and Composition: PreMediaGlobal

Library of Congress Control Number: 2012950573

ISBN-13: 978-1-285-17467-9

ISBN-10: 1-285-17467-4

Heinle
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil and Japan. Locate your local office at: **international.cengage.com/region**

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

For your course and learning solutions, visit **www.cengage.com.**

Purchase any of our products at your local college store or at our preferred online store **www.cengagebrain.com.**

Instructors: Please visit **login.cengage.com** and log in to access instructor-specific resources.

Printed in the United States of America
1 2 3 4 5 6 7 16 15 14 13 12

Chapter 4

Beyond Hybrid Learning: A Synthesis of Research on E-tutors Under the Lens of Second Language Acquisition Theory

Luis Cerezo

For several reasons—logistical, theoretical, pedagogical, and evidence based (Goertler, 2011)—higher education institutions are increasingly advocating for technology-enhanced curricula for second language acquisition (SLA). Research in this field of inquiry has grown exponentially, and recent literature reviews concluded that technology can be effectively used to develop different areas of a second language (L2) (Abraham, 2008; Egbert, Huff, McNeil, Preuss, & Sellen, 2009; Felix, 2005a; Hubbard, 2005; Peterson, 2010; Stockwell, 2007). Even more interestingly, recent meta-analyses of developmental studies found greater effect sizes for conditions using technology than conditions based on other forms of instruction, including face-to-face (F2F) instruction (Grgurovic, Chapelle, & Shelley, in preparation, as cited in Chapelle, 2010; Zhao, 2003) and individual study with printed texts (Taylor, 2009). These results presage a bright future for technology-enhanced language learning (TELL), although several limitations should be considered. For example, more replications and longitudinal studies are needed; also, studies must exert tighter control over external and internal factors (e.g., time on task, type of pedagogical practice, amount of exposure to input, and learners' individual differences); and existing meta-analyses are based on a small pool of studies, so their results must be interpreted with caution (see also Felix, 2005b; and Norris & Ortega, 2010, for a discussion of issues in meta-analytic literature).

From a different perspective, it is important to avoid overgeneralizations when making claims about technology because TELL can take many forms, and each of them may yield very different learning outcomes. For example, whereas in *hybrid or blended* curricula, 30 to 80 percent of F2F meetings are replaced by online discussions, in *online or open* curricula, this percentage is surpassed (Allen & Seaman, 2011). Additionally, curricula can incorporate *technology-mediated* or *technology-based* components, depending on whether learners respectively use technology to interact with other humans or to work independently. Overall, much ink has been spilled to discuss different types of TELL environments in descriptive studies, action research studies, and position papers. However, among the studies that used pretest–posttest designs to measure L2 development, the majority have focused on hybrid and technology-mediated learning (e.g., Tang, Yin, & Lou, 2009). The field has caught full speed and new studies in computer-mediated communication (CMC) appear in journals issue after issue. In contrast, relatively few developmental studies have investigated online learning, in which F2F instruction is dramatically reduced or eliminated, and even fewer studies

have investigated technology-based learning, in which learners work independently with technology without interacting with a teacher or a peer (Blake, 2009; Cerezo, 2010; Hulstijn, 2000). In this latter scenario, so-called *e-tutors* occupy a central role. An e-tutor is software that, in addition to displaying language input, has the capability of engaging learners in some sort of pedagogical practice and evaluating their responses, parsing them as correct or incorrect, and reacting accordingly by, for example, providing modified input or corrective feedback (see, e.g., Levy, 1997; Taylor, 1980). There are different types of e-tutors, depending on how they process learners' responses (with or without natural language processing tools) and how they react to them (with or without using learning modeling techniques) (see Heift & Schulze, 2007; Schulze, 2008, for a review).

By nature, the development of an e-tutor involves expertise in two different areas, linguistics and computer science. For this reason, scholarly research on e-tutors is spread out over a wide range of journals—of applied linguistics, computer science, computational linguistics, and educational technology—and is not easily accessible. As a result, TELL researchers often have the perception that the literature on e-tutors is very limited (Schulze, 2008). Although this is not the case, research on e-tutors is largely descriptive, and only a fraction of studies have investigated whether and how e-tutors may facilitate L2 development by using pretest–posttest designs, especially in the realm of grammar (e.g., AbuSeileek, 2009; Cerezo, 2010; Nutta, 1998). This relative lack of developmental studies is particularly striking in light of the growing presence and potential of e-tutors inside and outside academia. Outside academia, the sales of commercial e-tutors (e.g., *Rosetta Stone*, *Tell Me More*, or *Rocket Languages*) are booming; inside academia, more and more university administrators are pushing for the adoption of online learning curricula—65 percent of 2,500 surveyed institutions in the United States reported that online learning is a critical component of their long-term strategic plans (Allen & Seaman, 2011). Several factors may have caused this imbalance of research. For example, product developers may be reluctant to finance investigations on e-tutors because they do not see an immediate return on the investment, and research-granting agencies may deride the "applied" nature of these investigations (Hulstijn, 2000). Also, from a pedagogical perspective, some researchers are more interested in using technology as a medium rather than a tutor because it best befits their perception of language learning as socially constructed knowledge. Or, from a motivational perspective, some teachers and scholars may perceive e-tutors as a potential threat to their jobs, so they choose to ignore them (see, e.g., Blake, 2001, 2007, for a review of teachers' negative attitudes toward technology).

Additionally, e-tutors are not only underresearched but also underreported in research syntheses and meta-analyses. Although e-tutors have been discussed at different lengths in specialized publications on TELL (e.g., Blake, 2009; Heift & Schulze, 2007; Schulze, 2008), they remain largely ignored by wider scope publications on SLA. For example, none of the four major meta-analyses of empirical literature on corrective feedback (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russell & Spada, 2006) included an e-tutor study. One wonders why this is so. Some may argue that studies on computerized feedback are not informative

to oral interaction literature because the sources of feedback are fundamentally different. However, computerized feedback is ultimately a human product, and e-tutors are likely to provide more consistent feedback in terms of type and frequency, which ultimately results in more tightly controlled studies (see, e.g., Hulstijn, 2000; Sanz, 2000; Tsutsui, 2004). From a different perspective, others may argue that modality should not be overlooked and that oral and written feedback should be analyzed separately. However, on the one hand, some meta-analyses did include studies of written feedback in CMC (Li, 2010; Mackey & Goo, 2007), and on the other, computerized feedback does not necessarily have to be written, as illustrated by recent research with prerecorded oral feedback (Cerezo, 2010).

Addressing these gaps, the present chapter provides a synthesis of empirical research on e-tutors for L2 grammar development, revolving around three central questions. First, the *whether* question addresses whether e-tutors can effectively promote L2 grammar development per se and compared with other types of instructional technologies based on workbooks, videos, and F2F. Second, the *why* question addresses why e-tutors can promote L2 development in light of three psycholinguistic or psychocognitive theoretical approaches: the theories of dual coding and multimedia learning, the noticing hypothesis, and the interaction framework. And third, the *which* question addresses which e-tutor features promote the highest L2 grammar development outcomes, focusing on the type of computerized practice and the type of computerized feedback. The chapter closes with a summary of findings and a set of recommendations for further research on e-tutors.

Literature Search and Selection Criteria

The following sections provide a thematically organized synthesis of published and unpublished primary sources on the use of e-tutors for L2 development with a focus on grammar. To qualify for this synthesis, a study had to meet all of the following four criteria: (a) The study involved the use of an e-tutor, understood as software that has the capability of parsing learners' responses as correct or incorrect, responding accordingly; thus, if the study involved the use of software to merely display L2 input, it was excluded (e.g., Doughty, 1991; Robinson, 1996). (b) The e-tutor was the only source of instruction; studies in which e-tutors were used as a supplement to other type of instruction, such as F2F, were excluded (e.g., Sagarra & Zapata, 2008). (c) The study measured L2 grammar development in the form of learning outcomes (e.g., accuracy, latency or fluency, complexity). (d) The study involved a pretest–posttest design; if assumptions were made based only on posttreatment measures, the study was excluded (e.g., Chen, 2006).

To identify primary sources, four techniques were used. First, the reference lists of published literature reviews and meta-analyses were consulted. Second, two electronic databases were searched, Linguistics and Language Behavior Abstracts (LLBA) and ProQuest Dissertations & Theses (PQDT). Search queries were based on combinations of the query phrase *second/foreign language* with the following terms: *computer-based instruction*, *computerized instruction*,

computerized tutors, *e-tutors*, *electronic tutors*, *distance learning*, *online learning*, *open learning*, *CALI*, *CALL*, *CASLA*, *iCALL*, *MALL*, *TELL,* and *WELL*. No time bracket was selected—retrieved sources included years up to 2011. Third, the reference lists of recent doctoral dissertations were consulted. And fourth, feedback from both colleagues and anonymous reviewers was incorporated.

Retrieved sources were classified in three sections, depending on whether they were informative to the *whether*, *why*, or *which* questions discussed below. This synthesis is by no means exhaustive; rather, its purpose is to provide tentative answers to important questions about e-tutors under the lens of SLA theory.

The "Whether" Question: Whether E-tutors Can Promote Second Language Grammar Development

Since Nagata's (1993) seminal publication, at least 20 empirical studies have used pretest–posttest designs to investigate whether e-tutors can successfully facilitate L2 grammar development. Interestingly, all of the retrieved studies answered this question affirmatively, showing that independent work with e-tutors can be pedagogically effective for specific L2 grammatical structures, learners, and contexts (AbuSeileek, 2009; AbuSeileek & Rababah, 2007; Bowles, 2008; Camblor, 2006; Cerezo, 2010; Hsieh, 2007; Lado, 2008; Lin, 2009; Medina, 2008; Moreno, 2007; Morgan-Short & Wood Bowden, 2006; Nagata, 1998a, 1998b; Nutta, 1998; Petersen, 2010; Rosa & Leow, 2004; Sachs, 2011; Sanz & Morgan-Short, 2004; Torlaković & Deugo, 2004; Watts, 1989). Additionally, a fraction of these studies compared the effectiveness of e-tutors against other instructional technologies. In Watts's (1989) study, e-tutors were compared with non-interactive videos; in Nagata's (1996) study, they were compared with instruction based on workbooks; and in Nutta (1998), Torlaković and Deugo (2004), AbuSeileek and Rababah (2007), Abuseileek (2009), and Petersen (2010), the comparison group was F2F instruction.

In Watts's (1989) study, 30 adult English students of beginning French interacted with two different versions of the *Dès le Debut* video series for four hours total. One group watched a traditional video and completed exercises in paper-and-pen format, and the other group used an e-tutor consisting of an interactive version of the video with online exercises. Results showed that the interactive video group experienced a combined gain of 30 to 35 percent on word translation, sentence translation, and listening comprehension and a decrease in pronunciation errors. Conversely, the traditional video group experienced only negligible gains and "minimal improvement in pronunciation." These results must be interpreted with caution because of several caveats. For example, the use of a combined gain score makes it difficult to quantify L2 grammar gains in the sentence translation task; also, no specific grammatical forms were investigated.

Focusing specifically on a grammatical structure (Japanese particles), Nagata (1996) compared the effectiveness of an e-tutor versus workbook-mediated instruction. In her study, 26 English students of Japanese completed a preliminary grammar explanation session and four practice sessions including workbook

or online activities. Results showed that the e-tutor group significantly outperformed the workbook group in production (fill-in-the-blank and sentence production), both immediately after the treatment and three weeks later, but not in comprehension (sentence translation into L1 English). Several limitations urge caution in interpreting these results. First, the feedback received by the two conditions was qualitatively different. The e-tutor group received concurrent feedback that pinpointed and explained the nature of errors, and the workbook group had access to an answer sheet with positive evidence only. Therefore, it is impossible to determine whether the observed differences were because of the use of different instructional technologies or the quality of the feedback. Also, the study did not control for timing and amount of pushed output. The students in the e-tutor group were given three opportunities to modify their answers, but the answer sheet approach of the workbook group precluded any opportunity for modified output.

Switching the scope from poorly interactive workbooks to human instructors, Nutta (1998) compared the effects of an e-tutor against teacher-directed instruction for the acquisition of English past and conditional tenses. In her study, 53 participants were exposed to seven one-hour instructional sessions. Students in the teacher-directed group used the *Focus on Grammar* textbook series and engaged in a variety of communicative activities with their teacher. In contrast, students in the e-tutor group followed a checklist to navigate through the *ELLIS Mastery* multimedia system, completing online activities with feedback. For the past tense, the students in the e-tutor group outperformed the teacher-directed group in production posttests (fill-in-the-blank and open-ended sentence completion), but no differences were observed for recognition (multiple choice). These results were partially maintained after two weeks, when again the e-tutor group performed best for open-ended sentence production. The results for the English conditional tense, on the other hand, were less differential. Both groups experienced similar gains on all tests except on the delayed open-ended sentence production test, in which the e-tutor group once again performed significantly better. This study, however, did not specify whether both conditions received comparable practice in terms of type and quantity; consequently, caution must be exerted when interpreting its findings.

In a similar study, Torlaković and Deugo (2004) compared the effects of instruction with an e-tutor versus teacher-fronted instruction for English adverb placement. Twenty-one college students were exposed to six hours of instruction over the course of two weeks and were divided into three groups. One group completed a number of pedagogical activities with an e-tutor, *Adverbial Analyzer*; another group completed similar activities in the context of teacher-directed instruction; and a control group was involved only in testing. Results showed that the e-tutor group was the only one that experienced significant learning in the assessment tasks, a grammaticality judgment task (both immediately after the treatment and two weeks later), and written controlled production (in the immediate posttest only). These results, however, must be interpreted with caution because the feedback received by the two experimental conditions during practice was qualitatively different. The e-tutor group had access to rules on adverb placement

with every practice item, but in the teacher-fronted group, students received positive feedback only when they made correct generalizations.

Building on previous studies, AbuSeileek and Rababah (2007) compared four different types of instructional settings, determined by the source (e-tutor vs. teacher) and the pedagogical approach (presentation of rules before or after practice activities). The targeted grammatical structures were five English verb tenses: simple present, simple past, present continuous, present perfect, and simple future. Participants were 128 male college students, and instruction took place over the course of four weeks. Results showed that although all conditions experienced significant learning gains, the combined e-tutor group outperformed the teacher-driven group in all verb tenses except for the present perfect, for which both groups experienced comparable learning. These results must be interpreted in light of the fact that during some parts of the experiment, students worked in groups rather than individually, which introduced a hybrid component in the e-tutor groups.

In a partial replication study, AbuSeileek (2009) again investigated the effects of learning grammar inductively or deductively with an e-tutor or a human teacher. The targeted structures this time were four types of English sentences: simple sentence, compound sentence, complex sentence, and compound complex sentence. Participants were 72 undergraduate male students, instruction time was four weeks, and assessment tests comprised a combination of multiple-choice recognition and written sentence production. Results showed that for the two easiest sentence structures (simple and compound sentences), the e-tutor and teacher-centered groups learned comparably; in contrast, for the more difficult sentence structures (complex sentence and compound complex sentence), the e-tutor groups performed significantly better. As in AbuSeileek and Rababah (2007), these results should be interpreted with caution because the e-tutor groups did include a hybrid component based on small group discussions.

Finally, Petersen (2010) compared the effects of instruction using oral recasts delivered by a teacher versus written recasts delivered by an intelligent e-tutor, *Sasha*. Participants were 56 high school students, and the targeted structure was English question formation. Results indicated that participation in recast-intensive interaction was a significant predictor of learning, in terms of both question development and morphosyntactic accuracy. However, the mode of the interaction did not have an effect on development—participants who received oral recasts from a teacher learned comparably to those who received written recasts from the e-tutor.

To sum up, then, accruing evidence suggests a positive answer to the *whether* question, that is, whether or not e-tutors can facilitate L2 grammar development. On the one hand, more than 20 empirical studies unanimously showed that e-tutors did facilitate L2 grammar development. On the other hand, incipient empirical research comparing the effectiveness of e-tutors versus other instructional technologies showed that under certain circumstances and for certain targeted grammatical structures, learners using e-tutors can at least parallel, and even supersede, the accuracy gains experienced by learners in instructional conditions with video, workbooks, and human teachers. However, this latter claim must be

interpreted with caution given important limitations. For one, although there is a growing body of empirical research comparing the effectiveness of hybrid versus F2F instruction, there is a blatant need for more studies comparing the effectiveness of e-tutors against other instructional conditions. And second, some of the existing studies present serious methodological caveats, involving lack of control for targeted linguistic structure, pedagogical practice, or corrective feedback, in terms of quantity, quality, or both.

The "Why" Question: Why Can E-tutors Promote Second Language Grammar Development?

Arguably, at least three theoretical frameworks can be proposed to account for the success of e-tutors in promoting L2 grammar development: the dual coding theory (or the related theory of multimedia learning), the noticing hypothesis, and the interaction framework. Paivio's (1986) dual coding theory and Mayer's (2001) theory of multimedia learning contend that visual and verbal information are processed differently and separately, along distinct channels, and they can enhance processing and learning of information if displayed concurrently through a multimedia environment—as long as they do not compete with each other, causing working memory overload. These two theories could then explicate why e-tutors led to superior learning than workbook instruction in Nagata (1996) because computers have greater multimedia capabilities than printed workbooks. However, this does not seem to be the case because Nagata's e-tutor displayed written text only, yet it promoted higher learning gains than her workbook approach. Similarly, these theories would fail to explicate why e-tutors yielded better results than videos (Watts, 1989) and teacher-centered instruction (AbuSeileek, 2009; AbuSeileek & Rababah, 2007; Nutta, 1998; Petersen, 2010; Torlaković & Deugo, 2004) because both movies and teachers can concurrently display written, aural, and visual information in effective ways.

From a different standpoint, e-tutors can be argued to increase students' attention to L2 input, which according to Schmidt's noticing hypothesis is "the necessary and sufficient condition for the conversion of input into intake" (Schmidt, 1993). According to Watts (1989, pp. 19–20), one of the advantages of the e-tutor over the traditional video approach in her study was that it eliminated students' passivity toward the screen. For example, Watts reported that the vast majority of students in the e-tutor group (29 of 30) actually spoke aloud when instructed to do so, but only a minority in the video group (7 of 30 students) followed the instructions. Interestingly, some students in the traditional video group declared that they felt stupid talking to a TV because it did not give them "the same confidence that a computer gives you." Even the students who reported enjoying both instructional approaches demonstrated lower rates of attention in the video treatment, completing only half of the worksheets. This suggests that computers have a greater ability than other media (e.g., books or non-interactive videos) to direct students' attention to L2 input, perhaps because they increase students' motivation to learn. On this token, the participants in the e-tutor group in Nagata's

(1996) study reported significantly greater satisfaction in their Likert-scale ratings of the format, interest, and content of the practice exercises, "in spite of the fact that both groups received the same exercises in the same format" (p. 66). However, beyond this evidence, more research addressing the role of attention and motivation as independent variables (and their possible correlation) is certainly warranted.

Finally, from a third perspective, the better results of e-tutors over other instructional technologies may be because of their inherent "interactive features" (Nagata, 1996, p. 66). Interestingly, a number of early comparative studies (1960–1980s) reviewed by Pederson (1987) showed no significant advantages of the computer versus other learning technologies based on audiovisual labs, teachers, or paper and pencil. Notably, the computers used in these earlier studies were not capable of evaluating the students' responses; therefore, they were not e-tutors in the strict sense. As argued by Nagata (1996), when e-tutors "simply transfer printed textbooks to the computer screen, the computer is nothing more than an electronic page-turner, and we cannot expect better results" (p. 54). It appears, then, that the advantage of e-tutors versus other technologies may lie in their capability to interact with learners or to present them with large amounts of interactive practice. Interestingly, though, very few studies have empirically addressed this.

Drawing on the interaction hypothesis (Long, 1996), it can be argued that participation in interactive practice can promote SLA because it allows learners to notice, comprehend, and process input; produce output; form and revise hypotheses in response to feedback; and produce modified output, processes that have been shown to be developmentally helpful (Mackey, 2007). However, the interaction hypothesis does not specify whether the effects of these processes are determined by the type of *agency* of the learner, that is, whether it is necessary for learners to engage proactively in the aforementioned processes or whether mere exposure to interactive practice by others (e.g., watching a video in which different interlocutors engage in these processes) may suffice. To date, only two e-tutor studies have investigated the impact of practice versus exposure to practice on L2 grammar development. In a seminal study, Hsieh (2007) obtained mixed results for Spanish *gustar* constructions, with both conditions experiencing similar gains in written recognition and production and the practice condition outperforming the observers in oral production. In turn, using an e-tutor that allowed students to interact with prefilmed human avatars, Cerezo (2010) found that learners who had the chance to interact with the e-tutor significantly outperformed their observers in written and oral production of L2 Spanish grammar but only under more demanding conditions (with less explicit feedback as opposed to feedback with grammatical explanations) and for easier rather than complex structures (prepositional relative clauses as opposed to present subjunctive).

To sum up, then, incipient literature suggests that the reason why e-tutors may facilitate grammar development (per se or compared to other instructional technologies) lies in their ability to engage learners in greater amounts of proactive interaction. However, this advantage may arise only in conditions that impose higher cognitive demands on the part of the student (e.g., inferring a rule,

producing rather than recognizing grammar) and for structures that are governed by simpler rather than complex rules (i.e., rules that learners can infer through repeated practice in line with their developmental level).

The "Which" Question: Which E-tutors Best Promote Second Language Grammar Development?

Arguably, e-tutors may differ in two main components: the type of pedagogical practice and the type of corrective feedback that they provide. Practice is typically thought of as a prerequisite for interaction (DeKeyser, 2007), and corrective feedback is the feature that distinguishes testing from learning tasks (Loschky & Bley-Vroman, 1993). The following two sections discuss the effects of these variables separately.

Type of Computerized Practice

Existing studies on the effects of type of practice in e-tutors can be classified into two groups: (a) type of practiced skill, such as comprehension, interpretation, recognition, or production (Morgan-Short & Wood Bowden, 2006; Nagata, 1998a, 1998b), and (b) type of task features, such as task essentialness or task complexity (Medina, 2008; Moreno, 2007). Studies addressing the effects of type of practiced skill have compared input- versus output-focused practice for Japanese honorifics (Nagata, 1998a), Japanese nominal modifiers (Nagata, 1998b), and Spanish preverbal direct object pronouns (Morgan-Short & Wood Bowden, 2006). In Nagata (1998a, 1998b), five types of exercises were used to elicit interpretation and production at the word, sentence, or paragraph level. Input-focused practice involved selecting, from three choices, the correct English interpretation for the Japanese targeted form. In contrast, output-focused practice involved using Japanese in translation, editing, or retelling activities. Similarly, in Morgan-Short and Wood Bowden (2006), referential and affective activities were included to elicit interpretation or production in the aural and written modes, although the focus was on the word and sentence levels rather than the paragraph level.

Taken together, these three studies found that both input- and output-focused practice led to comparable gains in the interpretation of the target grammatical structures, but for production, output-focused practice yielded the best results, both in the short term (Morgan-Short & Wood Bowden, 2006; Nagata, 1998a, 1998b) and weeks later (Nagata, 1998a, 1998b). Furthermore, the superior effectiveness of output-focused instruction manifested itself for both complex and easy structures (Nagata, 1998a, 1998b). When comparing these findings with the general SLA literature, the following conclusions can be drawn: (a) These findings lend support to the growing number of empirical studies showing the facilitative role of output in L2 grammar development (e.g., Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Loewen, 2005; Mackey, 1999; McDonough, 2005); (b) they contradict some studies on processing instruction that found no significant differences in production and an advantage of input-focused practice for interpretation (e.g., Benati, 2005; Cadierno, 1995; Farley, 2001); and (c) they seem to contradict skill

acquisition theory (Anderson, 1993), according to which both input- and output-focused practice develop comprehension and production skills, respectively (e.g., DeKeyser & Sokalski, 2001).

Focusing on task features rather than practiced skills, Medina (2008) compared the pedagogical effects of task complexity in input-focused tasks. In her experiment, participants were asked to read a text containing exemplars of the targeted structure, Spanish imperfect subjunctive. Participants in the more complex condition were asked to spot the structures and select them in the text, and those in the less complex condition were instructed to re-highlight the forms that were already underlined for them. Results showed that task complexity had a significant effect on the written production of the target structure, with participants in the more complex task group achieving significantly higher learning outcomes. Extrapolation of these findings to the general SLA literature yields the following two conclusions: (a) there is increasing support for the cognition hypothesis, according to which more complex tasks promote higher levels of accuracy (e.g., Iwashita, McNamara, & Elder, 2001; Kuiken, Mos, & Vedder, 2005), and (b) the observed lack of effects of task complexity for recognition confirms Flynn's (1986) hypothesis that comprehension precedes production, so the effects of more complex tasks are more likely to arise in production measures.

Also addressing task features, Moreno (2007) investigated the effects of task essentialness in input-focused practice on the development of Spanish preverbal direct object pronouns. Participants in her study were presented with a picture cue and had to compose or select a sentence that described it. In the more task-essential condition, students composed the sentence by navigating a tree that successively branched out with a set of object pronouns. In contrast, students in the less task-essential condition completed an interpretation task by selecting one of two fully composed sentences. Crucially, in this latter condition, attention to the targeted form was not enforced because the two sentences differed only in terms of their subjects rather than their objects. Results showed that both conditions performed statistically similarly. Notably, Moreno's (2007) findings are hard to extrapolate given the lull of empirical research on task essentialness.

In sum, existing literature on the effects of type of computerized practice on L2 grammar development suggests that output-focused practice yields better results than input-focused practice because it can promote comparable learning gains in interpretation and superior learning gains in production. In terms of task features, increased task complexity has been shown to have a positive effect on L2 grammar development (Medina, 2008), but task essentialness has not (Moreno, 2007). Clearly, however, the marked paucity of studies in this field of inquiry warrants further investigation to elucidate the role of these and other task features.

Type of Computerized Feedback

In contrast to the scant literature on the type of computerized practice, there is a considerably larger body of research investigating the effects of the type of computerized feedback. Interestingly, although existing studies all agree that feedback, regardless of its type, has a positive effect on L2 grammar development, the jury is still out as to the relative effectiveness of different types of feedback. Most studies

to date have shown a significant advantage of more "explicit" types of feedback in which learners are given an explanation of the sources of their errors (Bowles, 2005; Cerezo, 2010; Lado, 2008; Lin, 2009; Nagata, 1993, also reported in Nagata & Swisher, 1995; Rosa & Leow, 2004; Sachs, 2011). However, counterevidence exists, with other studies finding no significant differences between more and less explicit types in which learners were merely informed that their responses were incorrect (Camblor, 2006; Hsieh, 2007; Moreno, 2007; Sanz & Morgan-Short, 2004). Additionally, one study showed significantly greater gains of less over more explicit feedback in oral delayed measures (Moreno, 2007). Notably, these mixed findings are in line with the longer established literature on feedback in classroom interaction, in which recent meta-analyses found that although corrective feedback has facilitative effects (Li, 2010; Lyster & Saito, 2010; Russell & Spada, 2006), it is still unclear which type of feedback yields the highest effect sizes. Russell and Spada (2006) and Mackey and Goo (2007) concluded that more studies are needed to decide whether implicit and explicit types of feedback are differently beneficial. In Li's (2010) meta-analysis, explicit feedback outperformed implicit types in the short to medium run, and in Lyster and Saito (2010), the developmental effects of recasts and explicit correction were comparable, and prompts yielded the best results.

Arguably, several external factors may explicate the conflicting findings in the literature on e-tutors (see, e.g., Nassaji, 2007, for a review of potentially mediating factors in classroom interaction research). One such factor is the type of computerized task at hand. Some studies used input-focused tasks, such as assembling sentence constituents (Bowles, 2005; Hsieh, 2007), matching pictures with sentences (Lado, 2008; Lin, 2009; Moreno, 2007; Sanz & Morgan-Short, 2004), matching the content of a text with its title (Sanz & Morgan-Short, 2004), truth-value judgments (Sachs, 2011), or filling in blanks using a list of possible forms (Rosa & Leow, 2004). In contrast, other studies used output-focused tasks, such as writing sentences to describe pictures (Camblor, 2006) or responding to context cues, both at the phrase level (Cerezo, 2010) and at the sentence level (Nagata, 1993). Another possibly mediating factor may have been the type of targeted grammatical structure, which included Japanese particles and passivization (Nagata, 1993), Japanese reflexive constructions (Sachs, 2011), Latin assignment of semantic functions (Lado, 2008; Lin, 2009), Spanish dative experiencer constructions with *gustar* (Bowles, 2005; Hsieh, 2007), Spanish noun–adjective gender and number agreement (Camblor, 2006), Spanish direct object pronouns (Moreno, 2007; Sanz & Morgan-Short, 2004), Spanish past counterfactual conditional sentences (Rosa & Leow, 2004), and Spanish present subjunctive and preposition pied-piping in adjectival relative clauses (Cerezo, 2010). Additionally, one study (Nagata, 1993) failed to isolate the role of corrective feedback from prior grammatical explanation, and only three studies (Lado, 2008; Lin, 2009; Sachs, 2011) addressed the role of learners' individual differences such as language aptitude or working memory to explicate within- and between-group variance.

In addition to external factors, the conflicting findings in the literature may also be explicated by a poor definition of the intrinsic nature of feedback. Specifically, most studies on types of computerized feedback have exclusively focused on

the explicitness of negative evidence (i.e., explaining the rule or not), overlooking the fact that feedback moves can also encapsulate positive evidence (i.e., models of correct language) and prompt learners to repair their errors. For example, except for Camblor (2006) and Cerezo (2010), all studies included models of correct grammar either in the feedback message or implicitly in the task, so it is impossible to assess whether the observed learning gains resulted from processing positive or negative evidence. Similarly, all studies but Cerezo (2010) pushed learners to repair their errors at all times. On this token, the findings of seminal studies in the oral interaction literature by Leeman (2003) and McDonough (2005) suggest that under specific circumstances, the role of positive evidence and prompting for repair may outweigh the impact of negative evidence per se. Addressing this gap, Cerezo (2010) studied the separate contributions of prompts and negative evidence in the absence of positive evidence. In his study, participants interacted with an e-tutor to complete two different tasks, reporting a theft and finding an apartment. Some received grammatical explanations of their errors, but others were only informed about the place and nature of their errors and did not receive the rule. Some participants were asked to modify their output after an error, but others were not. Results showed that in the absence of positive evidence, more explicit negative evidence positively affected learning, but prompting did not. This provides additional support to the claim that prompting is only likely to have facilitative effects when learners have access to positive evidence, either concurrently in the ambient input or retrospectively by rescuing previously stored knowledge (see, e.g., VanPatten, 2004; but see Toth, 2006, for counter-arguments).

In sum, then, the growing body of literature on the effects of the type of computerized feedback on L2 grammar development has largely focused on comparing different degrees of explicitness of negative evidence (explaining the rule at play or not), with inconclusive findings. This may result from a lack of control over external and internal factors. At an external level, a variety of input- and output-focused tasks were used, different grammatical structures were targeted, and individual differences in language aptitude were controlled for only in a fraction of studies. At an internal level, studies have largely failed to isolate the relative and combined contributions of the three components of feedback (positive evidence, negative evidence, and prompting for repair). Specifically, most studies have neglected the role of positive evidence because they included models either in the feedback message or implicitly in the task. Similarly, they mostly did not control for the role of pushed repair, prompting learners to fix their errors at all times. On this token, incipient research showed that in the absence of positive evidence, the role of explicit negative evidence outweighed prompting, which was not beneficial.

Conclusion and Further Research

In the new millennium, the field of TELL is expanding its horizons to address ever more interesting research questions. Besides testing the pedagogical potential of technology per se and against other forms of instruction, researchers have also started to compare different technologies with each other to identify the most

efficient ones. Also, premised on current theories of SLA and using hybrid designs that triangulate quantitative and qualitative analysis, researchers have started to explain why technologies can promote L2 development.

So far, however, the spotlight has been cast primarily on the use of technology as a medium of communication among tutors and peers in hybrid learning curricula. Yet given the increasing interest of university administrators in online learning curricula, language program administrators must turn to the incipient but steadily growing body of research on e-tutors to make better informed curricular decisions. So far, research has shown that under certain circumstances, e-tutors can successfully facilitate L2 grammar development per se and even better than other instructional technologies. Although several psycholinguistic and psychocognitive frameworks may explain this, it seems likely that the pedagogical potential of e-tutors lies in their ability to engage learners in large amounts of proactive practice, but not always—only under more demanding learning conditions (e.g., with output-focused tasks and less explicit feedback). Additionally, incipient research comparing different types of e-tutors tentatively suggests that the most efficient systems are those that use output-focused and complex tasks. As for the type of corrective feedback, the jury is still out. Although a growing body of studies suggests that, overall, more explicit feedback yields better results, a number of studies have shown that in some cases, minimally explicit feedback can work equally well, and learning outcomes may vary depending on several factors, including the complexity of the targeted structures.

In light of several methodological caveats and the overall paucity of studies on e-tutors, further research is certainly warranted. There is a blatant need for original and replication studies comparing the pedagogical effectiveness of e-tutors against other instructional technologies, including but not limited to F2F instruction. Also, researchers should compare different types of computerized features to contribute to the development of more sophisticated and efficient e-tutors. Specifically, further research should investigate how different types of computerized practice affect learning, comparing subtypes of input- and output-focused practice with varying degrees of complexity and task essentialness, at different levels (word, sentence, paragraph). Additionally, studies on computerized feedback should investigate the separate and combined contributions of different components of feedback (e.g., prompts, differently explicit positive and negative evidence). There is also a need for more studies addressing the role of modality (written, aural, visual, or a combination) in computerized instruction, in terms of both the L2 input and feedback. Finally, more studies should empirically address what kinds of learner external and internal variables may moderate the effects of computerized instruction.

Clearly, e-tutors still have a long way to go until they can effectively train language learners in all levels of communicative competence—grammatical, sociolinguistic, and strategic. Actually, the question of whether or not e-tutors might be able to ever replace human language instructors is a futile one: most learners study languages to communicate with other humans, not machines. But in a time when using technology for language learning is becoming the norm rather than the exception, when university administrators have identified online instruction

as a critical goal, and given the many potential contributions of e-tutors to provide learners with additional access to language practice and corrective feedback, e-tutors should clearly occupy a more prominent position in our research agendas and language program administrators should remain informed.

References

Abraham, L. B. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary learning: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199–226.

AbuSeileek, A. F. (2009). The effect of using an online-based course on the learning of grammar inductively and deductively. *ReCALL, 21*(3), 319–336.

AbuSeileek, A. F., & Rababah, G. A. (2007). The effect of computer-based grammar instruction on the acquisition of verb tenses in an EFL context. *The JALT CALL Journal, 3*(1-2), 59–80.

Allen, I. E., & Seaman, J. (2011). Going the distance: Online education in the United States, 2011. Retrieved March 25, 2012, from http://www.onlinelearningsurvey.com/reports/goingthedistance.pdf.

Anderson, J. R. (1993). *Rules of the mind.* Hillsdale, NJ: Erlbaum.

Benati, A. (2005). The effects of processing instruction, traditional instruction, and meaningful-based output instruction on the acquisition of English past simple tense. *Language Teaching Research, 5*, 95–127.

Blake, R. (2001). What language professionals need to know about technology. *ADFL Bulletin, 32*(3), 93–99.

Blake, R. (2007). New trends in using technology in the language curriculum. *Annual Review of Applied Linguistics, 27*, 76–97.

Blake, R. (2009). The use of technology for second language distance learning. *Modern Language Journal, 93*(Focus Issue), 822–835.

Bowles, M. A. (2005). Effects of verbalization condition and type of feedback on L2 development in a CALL task. (Unpublished doctoral dissertation). Georgetown University, Washington D.C.

Bowles, M. A. (2008). Task type and reactivity of verbal reports in SLA: A first look at a L2 task other than reading. *Studies in Second Language Acquisition, 30*(3), 359–387.

Cadierno, T. (1995). Formal instruction from a processing perspective: An investigation into the Spanish past tense. *Modern Language Journal, 79*(2), 179–193.

Camblor, M. (2006). *Type of written feedback, awareness, and L2 development: A computer-based study.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Cerezo, L. (2010). *Talking to avatars: The computer as a tutor and the incidence of learner's agency, feedback, and grammatical form in SLA*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Chapelle, C. A. (2010). The spread of computer-assisted language learning. *Language Teaching, 43*(1), 66–74.

Chen, L. (2006). The effect of the use of L1 in a multimedia tutorial on grammar learning: An error analysis of Taiwanese beginning EFL learners' English essays. *The Asian EFL Journal Quarterly, 8*(2), 76–110.

DeKeyser, R. M. (2007). Situating the concept of practice. In R. M. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1–18). New York: Cambridge University Press.

DeKeyser, R. M., & Sokalski, K. J. (2001). The differential role of comprehension and production practice. *Language Learning, 51*(supplement 1), 81–112.

Doughty, C. J. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*(4), 431–469.

Egbert, J., Huff, L., McNeil, L., Preuss, C., & Sellen, J. (2009). Pedagogy, process, and classroom context: Integrating teacher voice and experience into research on technology-enhanced language learning. *The Modern Language Journal, 93*, 754–768.

Farley, A. P. (2001). Processing instruction and meaning-based output instruction: A comparative study. *Studies in Applied Linguistics, 5*, 57–93.

Felix, U. (2005a). Analysing recent CALL effectiveness research? Towards a common agenda. *Computer Assisted Language Learning, 18*(1-2), 1–32.

Felix, U. (2005b). What do meta-analyses tell us about CALL effectiveness? *ReCALL, 17*(2), 269–288.

Flynn, S. (1986). Production vs. comprehension: Differences in underlying competences. *Studies in Second Language Acquisition, 8*, 135–164.

Goertler, S. E. (2011). Blended and open/online learning: Adapting to a changing world of language teaching. In N. Arnold & L. Ducate (Eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 471–502). San Marcos, TX: CALICO.

Grgurovic, M., Chapelle, C. A., & Shelley, M. (in preparation). Effectiveness of CALL: A meta-analysis and research synthesis.

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

Hsieh, H. C. (2007). *Input-based practice, feedback, awareness and L2 development through a computerized task.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Hubbard, P. (2005). A review of subject characteristics in CALL research. *Computer Assisted Language Learning, 18*(5), 351–368.

Hulstijn, J. H. (2000). The use of computer technology in experimental studies of second language acquisition: A survey of some techniques and some ongoing studies. *Language Learning & Technology, 3*(2), 32–43.

Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning, 51*, 401–436.

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition, 21*, 421–452.

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. P. García & J. Cenoz (Eds.), *EUROSLA Yearbook* (Vol. 5, pp. 195–222). Amsterdam: John Benjamins.

Lado, B. (2008). *The role of bilingualism, type of feedback and cognitive capacity in the acquisition of non-primary languages: A computer-based study.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition, 25*, 37–63.

Levy, M. (1997). *CALL: Context and conceptualisation*. Oxford: Oxford University Press.

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

Lin, H. J. (2009). *Bilingualism, feedback, cognitive capacity, and learning strategies in L3 development.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27*(3), 361–386.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition: Vol. 2. Second language acquisition* (pp. 413–468). New York: Academic Press.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. M. Gass (Eds.), *Tasks and language learning* (pp. 123–167). Clevedon, UK: Multilingual Matters.

Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition, 32*, 265–302.

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

Mackey, A. (2007). Introduction: The role of conversational interaction in second language acquisition. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 1–26). Oxford: Oxford University Press.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.

Mayer, R. E. (2001). *Multimedia learning*. New York: Cambridge University Press.

McDonough, K. (2005). Identifying the contributions of negative feedback and learners' responses to L2 development. *Studies in Second Language Acquisition, 27*, 79–103.

Medina, A. D. (2008). *Concurrent verbalization, task complexity, and working memory: Effects on L2 learning in a computerized task.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Moreno, N. (2007). *The effects of type of task and type of feedback on L2 development in CALL.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Morgan-Short, K., & Wood Bowden, H. (2006). Processing instruction and meaningful output-based instruction: Effects on second language development. *Studies in Second Language Acquisition, 28*(1), 31–65.

Nagata, N. (1993). Intelligent computer feedback for second language instruction. *Modern Language Journal, 77*, 330–339.

Nagata, N. (1996). Computer vs. workbook instruction in second language acquisition. *CALICO Journal, 14*, 53–75.

Nagata, N. (1998a). Input vs. output practice in educational software for second language acquisition. *Language Learning & Technology, 1*, 23–40.

Nagata, N. (1998b). The relative effectiveness of production and comprehension practice in second language acquisition. *Computer Assisted Language Learning, 11*(2), 153–177.

Nagata, N., & Swisher, M. V. (1995). A study of consciousness-raising by computer: The effect of metalinguistic feedback on SLA. *Foreign Language Annals, 28*(3), 336–347.

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

Norris, J. M., & Ortega, L. (2010). Research timeline: Research synthesis. *Language Teaching, 43*(4), 461–479.

Nutta, J. W. (1998). Is computer-based grammar instruction as effective as teacher-directed grammar instruction for teaching L2 structures? *CALICO Journal, 16*, 49–62.

Paivio, A. (1986). *Mental representations: a dual coding approach.* Oxford: Oxford University Press.

Pederson, K. M. (1987). Research on CALL. In W. F. Smith (Ed.), *Modern media in foreign language education: Theory and implementation* (pp. 99–131). Lincolnwood, IL: National Textbook.

Petersen, K. A. (2010a). *Implicit corrective feedback in computer-guided interaction: Does mode matter?* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Peterson, M. (2010b). Computerized games and simulations in computer-assisted language learning: A meta-analysis of research. *Simulation & Gaming, 41*(1), 72–93.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18*(1), 27–67.

Rosa, E. M., & Leow, R. P. (2004). Computerized task-based instruction in the L2 classroom: The effects of explicitness and type of feedback on L2 development. *Modern Language Journal, 88*(2), 192–216.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar. In J. M. Norris (Ed.), *Synthesizing research on language learning and teaching* (pp. 133–162). Philadelphia: John Benjamins.

Sachs, R. (2011). *Individual differences and the effectiveness of visual feedback on reflexive binding in L2 Japanese.* Unpublished doctoral dissertation, Georgetown University, Washington, DC.

Sagarra, N., & Zapata, G. (2008). Computer-assisted instruction and L2 grammar accuracy. *Hispania, 91*(1), 93–109.

Sanz, C. (2000). Implementing LIBRA for the design of experimental research in SLA. *Language Learning and Technology, 13*(2), 27–31.

Sanz, C., & Morgan-Short, K. (2004). Positive evidence versus explicit rule presentation and explicit negative feedback: A computer-assisted study. *Language Learning, 54*, 35–78.

Schmidt, R. W. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206–226.

Schulze, M. (2008). AI in CALL: Artificially inflated or almost imminent? *CALICO Journal, 25*(3).

Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. *ReCALL, 19*(2), 105–120.

Tang, H. W., Yin, M. S., & Lou, P. J. (2009). A meta-analytic review of current CALL research on second language learning. *IT in Medicine & Education, 1*, 677–684.

Taylor, A. M. (2009). CALL-based versus paper-based glosses: Is there a difference in reading comprehension? *CALICO Journal, 27*(1), 147–160.

Taylor, R. P. (Ed.). (1980). *The computer in the school: Tutor, tool, tutee*. New York: Teachers College Press.

Torlaković, E., & Deugo, D. (2004). Application of a CALL system in the acquisition of adverbs in English. *Computer Assisted Language Learning, 17*(2), 203–235.

Toth, P. D. (2006). Processing instruction and a role for output in second language acquisition. *Language Learning, 56*(2), 319–385.

Tsutsui, M. (2004). Multimedia as a means to enhance feedback. *Computer Assisted Language Learning, 17*(3), 377–402.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–32). Mahwah, NJ: Lawrence Erlbaum Associates.

Watts, C. (1989). Interactive video: What the students say. *CALICO Journal, 7*(1), 17–20.

Zhao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal, 21*(1), 7–27.