Original Article

Foreign Languages and Computer-Assisted Learning: New Principles for Language Assessment in Teletandem?

Douglas Altamiro Consolo

1. Professor, Departamento de Letras Modernas, State University of Sao Paulo (UNESP), Campus of Sao Jose do Rio Preto, Brazil

Abstract

In this article I review principles and practical aspects of language assessment in foreign language (FL) learning, more specifically with a focus on Computer-Assisted Language Learning (CALL) and other contexts in which computers are used as a means for distance learning and for language assessment (Computer-Aided Assessment - CAA). Assessment constitutes an essential dimension of learning experiences and it is an aspect of most formal processes of language teaching and learning. Language assessment may involve language testing, as well as other procedures and instruments such as observations, performance tasks, portfolios and self-evaluation, and by combining information from various sources of assessment one is able to obtain more valid and reliable results. I draw on the literature on CALL and on language assessment, and on data collected within the scope of the Teletandem Brazil Project: foreign languages for all (henceforth TBP), to support my position on principles that may or may not characterize language assessment in the context of teletandem interactions. CAA is defined as any type of activity in which computers are used to support a process of assessment apart from and beyond their simple function to store and transmit information. CAA helps faster assessment, increases the quality and quantity of information detected and maximizes the provision of feedback about language assessment processes. In the TBP project, undergraduate students from a Brazilian university interacted with students from universities abroad, by means of computer programmes for synchronous communication, microphones and webcams. Besides the claims about CALL from the literature, I make reference to teletandem interactions in both EFL and Portuguese as a foreign language, considering occasions in which teletandem agents evaluate each other linguistic performances. I also analyse a questionnaire for evaluation in teletandem, which does not focus on language assessment but rather on the experience of interacting in the teletandem context and on the tools used for communication. No clear distinctions were found to exist between CAA and more traditional procedures for language assessment. Principles for CAA seem to combine traditional bases for language assessment and testing with a number of pedagogical principles that underpin distance learning. Nevertheless, CALL and CAA can contribute in various aspects of language education, especially when large numbers of learners are involved in teaching and learning processes. However, these principles do not characterize a new paradigm in language assessment, since the linguistic criteria on which teletandem agents base their evaluation are very similar to criteria that underpin language assessment and testing by means of paper-and-pencil tests, for example. I conclude the article indicating the need for further investigation and the establishment of principles for language assessment in electronic contexts.

Keywords

Assessment, Computer-Assisted Language Learning (CALL), Distance learning, Foreign languages, Teletandem.

*Corresponding Author: douglas.consolo@unesp.br

Introduction

Assessment constitutes an essential dimension of learning experiences, and it is an aspect of most formal processes – as well as some informal processes – of language teaching and learning. Formative assessment, for example, allows learners to improve in terms of language development and can help to facilitate successful learning experiences. Achievement testing, on the other hand, indicates learning goals reached by means of teaching and learning experiences. Assessment as a process reveals the results of teaching and learning experiences in relation to expected learning aims or proficiency standards.

Given the scope of different types of learning environments available at present – from more standard language classrooms to distant learning, and contexts in which language learning occurs by means of or with the help of computers and considering that the areas of language assessment and language testing are grounded on various consolidated theoretical principles formulated mainly with regards to standard classrooms, it seems relevant to reflect and discuss which principles characterize and underpin computer-aided language assessment and testing. Therefore, motivated by a desire to investigate and understand electronic language assessment, in this paper I deal with principles and aspects of language assessment in foreign language (FL) learning, more specifically with a focus on CALL and other contexts in which computers may be used as a means for distant education and for language assessment. Assessment involves language testing, as well other means and procedures to verify whether language learning has occurred, and considers possible backwash effects of language assessment on language learning and teaching as well. Although backwash effects from language assessment and testing have been the subject of attention in the current literature on Applied Linguistics, I do not make explicit reference to those effects here. Some positive effects of assessment are mentioned insofar as they are seen as contributions for language learning.

The advancement of technology has contributed to facilitate language learning and teaching [1]. However, there is a lack of knowledge about online language assessment, especially concerning valid measures of proficiency outcomes and how to assess online language learning more effectively [2, 1]. On the one hand, electronic assessment tools have advantages when compared to paper-and-pen(cil) tests, for example, the use of multimodalities, easier access to data banks, and faster correction and provision of feedback to candidates [3, 4].

I draw on the literature on CALL and on language assessment, and on data collected within the scope of the Teletandem Brazil Project: foreign languages for all (henceforth TBP), to support my position on principles that may or may not characterise language assessment in CALL and in teletandem interactions. Besides the claims about CALL from the literature and the support of data from the TBP, represented by teletandem interactions in both EFL and Portuguese as a FL (henceforth PFL), I make brief reference to a type of blended learning environment for EFL – an English language course in the curriculum of a Letters course in Brazil.

In order to explain the meaning of 'teletandem', I quote the explanation from the web page of the TBP:\

Language learning in tandem involves pairs of native or non-native speakers of different languages working collaboratively to learn each other's language. Teletandem Brazil matches up Brazilian university students who wish to learn a foreign language, with students in other countries who are learning Portuguese. With tandem language learning, each partner is a student for one hour, learning and practicing a language from the other partner. Then they switch roles and switch languages.

Teletandem can thus be defined as a process of cooperative language learning by means of electronic communication. Students at UNESP, a public university in the state of Sao Paulo, in Brazil, have been encouraged to register on the TBP webpage to obtain student partners from universities in other countries where agreements for the project have been established – for example, in Argentina, France, Germany, Italy, Mexico, Sweden and the USA. The students

abroad are learners of PFL and the students in Brazil are learners of English, French, German, Italian and Spanish. As stated above, in a partnership, students are expected to help each other learn the languages in which they are proficient users.

Interactions in the TBP are grounded on the principle of learner autonomy, that is, language learning is no longer the responsibility of a class teacher alone. Learners are responsible for their own process of language learning, and this responsibility requires that learners decide about their learning goals, the content of learning and the resources to be used. In this sense, learners benefit from the possibility of negotiating the aforementioned aspects with their partners – that is, decisions which can contribute - or not - to the success of a collaborative language learning experience, or which can possibly reduce the benefits of the teletandem experience, are the agents' prerogative.

Reflection is another principle of teletandem and, according to Schön [5] and Mezirow [6], reflection may bridge the traditional didactic asymmetry usually found in standard classrooms, in the sense that the student also becomes a 'teacher'. Moreover, reflection offers the learners the possibility of negotiating the course of the interactions and, as a result, the route of their learning experience.

Reciprocity is a third principle that supports interactions in the TBP, that is, both agents are expected to act as language 'teachers' and 'learners' so that they can not only experience language development as learners but also learn how to behave as the partner who is more proficient in one of the languages involved. Based upon language proficiency, on previous experience of foreign language learning and on teaching experience, if that is the case, and on reflection, the most proficient agent is expected to decide on appropriate courses of action so as to help his or her partner learn a foreign language. These actions involve how the most proficiency agent deals with situations in which the learner lacks linguistic competence or any other type of knowledge to express his or her ideas, or when the learner makes language mistakes – given the fact that language mistakes may or may not impair communication. Because lack of linguistic competence and language mistakes are two phenomena that are commonly related to when and how teachers assess language learners in teaching processes, I shall return to this assertion later, in the 'Discussion' section of this paper. Conversely, the most proficient agent's decision on whether to provide corrective or non-corrective feedback in the course of a teletandem interaction, akin to what happens in face-to-face classroom interaction, brings the nature of such decisions closer to those taken by teachers in standard language lessons.

Teletandem interactions occur by means of online chat, audio or video communication, with the help of communication devices and software such as MSN, Skype and Zoom, and generate a corpus of written and spoken data. Focusing on spoken language and for research purposes, oral data has been recorded by means of a software called Easy Recorder, which is available on the internet, free of charge. Written data produced in interactions by MSN were also been recorded by means of the command to record MSN files.

A full teletandem session usually lasts two hours. One hour is dedicated to each of the two languages used by the agents. In principle each one-hour session comprises three parts: (a) conversation, (b) feedback on language and (c) evaluation of the session. In the first part of the session the agents engage in a conversation in the target language, about one or more topics, for around thirty minutes. In the second part, which takes approximately twenty minutes, the agents discuss the language used in their previous conversation and the most proficiency agent has the opportunity to provide linguistic feedback to his or her partner, with the help of notes written during the conversation or, in the case of written communication (chat), by referring to the previous lines of their interaction. The third part of the session lasts around ten minutes and is dedicated to evaluating the whole session, comprising a discussion about the difficulties faced by the participant while interacting in teletandem and suggestions for future action. Once the agents have completed an interaction period at least twelve weeks, they may decide to continue or to end

their partnership. A final evaluation of the experience of engaging in teletandem interactions is provided to the TBP research team by means of an evaluation questionnaire answered by the agents, available on the web page (see Appendix). The 'Teletandem evaluation questionnaire' does not focus on language assessment but rather on the experience of interacting in the teletandem context and on the tools used for communication. Issues concerning linguistic aspects and language use in teletandem interactions have been dealt with in scientific initiation studies and MA dissertations [7, 8, 9, 10].

In the next section I review a theoretical background to foreign language learning and assessment, and in section 3 aspects of computer-aided assessment are reported. I then proceed to a discussion about the issues concerning language assessment in distance learning of languages and teaching and present my position about principles for electronic language assessment.

Background to foreign language learning and assessment

Because this discussion involves two related concepts sometimes used as synonyms, but which imply different aspects and dimensions of language teaching and learning – assessment and evaluation, definitions must be provided. According to Garrison and Anderson [11], evaluation refers to a comparison between course units or programs and some determined criteria for course results. These results may include students' or customers' satisfaction with the course results attained. Assessment refers to the process of critically evaluating students' performance and development towards educational goals, which include language knowledge, language skills and linguistic performance. Language assessment should follow and be aligned with the same concepts and principles chosen as support for a given language learning process, that is, views on what language, language use and language learning mean should be reflected in the criteria to assess language development and effective language use.

Assessment can be more effective when the principles of 'multiplism' [12] are followed. According to the authors, one has to be aware of several facets involved in assessing language learning, and of the instruments and procedures available – for example, observations and grids, questionnaires, tasks and tests - to better map out or verify if learning has occurred.

Language assessment, in many circumstances and especially in formal contexts, aims at verifying at which level of language proficiency a given learner can be classified. For Stern [13], language proficiency means the actual performance of a learner in a given language, and it involves the mastery of (a) the forms, (b) the linguistic, cognitive, affective and sociocultural meanings of those forms, (c) the capacity to use the language with a focus principally on communication rather than attention on form, and (d) creativity in language use. Based on this definition one may interpret communicative language ability (or communication by means of language use) as being constituted of two components: linguistic proficiency and communicative proficiency.

Scaramucci [14] adopts two senses of proficiency with regards to terminology: a technical and a non-technical sense. The non-technical sense generally encompasses impressionistic judgments based on a holistic view and values a concept of proficiency that can be regarded as monolithic, stable and unique. This concept is usually pre-defined and represents a boundary that distinguishes, in overall terms, between proficient and non-proficient learners. However, the author emphasizes that such a concept of proficiency is to be understood as dependent on other variables like the teaching context, its characteristics and objectives, which also makes it relative and variable as well. In its technical sense, the concept of language proficiency encompasses levels within which the descriptions of language ability and use fall in order to indicate what a language user is able to do and under which circumstances. In this sense, proficiency takes into account the real aims of using language in social contexts, for example, in distance-learning interactions by means of computers.

Computer Aided Assessment (CAA)

Computer-Aided Assessment (henceforth CAA) is defined as any type of activity in which computers are used to support a process of assessment apart from and beyond their simple function to store and transmit information [15]. CAA helps faster assessment, increases the quality and quantity of information detected, and maximizes the provision of feedback about language assessment processes.

An essential principle underpinning CAA is that it requires pedagogical bases that are consistent and coherent with the pedagogical principles that support distance learning of languages by means of computers [15]. This is why the principles that illuminate teletandem interactions (autonomy, co-operation, reflection and reciprocity), as reported in the first section of this paper, also have a role in the scope of principles to support CAA. As stated in section 2, principles that support language learning under any given set of conditions and with regards to clear language learning goals, should support the actions and techniques used in assessing learning, for consistency and coherence between learning and assessment.

The literature on CAA brings some types of questions and procedures that are adequate for formal assessment in electronic contexts. These procedures include multiple-choice questions, textual answers, problem-solving tasks and peer evaluation.

According to Carter et al. (mimeo.), multiple-choice questions can be used to assess knowledge at many different cognitive levels. If the questions are well-designed – that is, if principles of content and language coherence between the question and the alternatives are respected, and the alternatives are properly formulated, multiple-choice questions can be really effective in assessing knowledge, as opposed to demonstrating merely logical competence. Multiple-choice questions can take the format of alternatives to be chosen, or to be combined, or gap-filling tasks. CAA allows for the generation of different permutations in alternatives, and the setting of different values for question variables and, in this way, tests can be individualized. Levels of complexity can also be increased with the aid of permutation. Drawbacks to assessment by means of multiple-choice questions, however, include memorization of answers and cheating.

Textual assessment requires students to write short answers or essays. Depending on the skill and experience of the question designer, textual assessment can be used to test both lower and higher order learning skills. There are several approaches to supporting automatic assessment of text answers. Some question systems base the assessment on either direct text comparison or collocations, for example, and these support short-answer questions. There are also more sophisticated approaches to the assessment of textual content [16, 17].

Problem-solving tasks require theoretical knowledge and ability to solve specific problems of a more practical nature. These tasks are usually assessed according to the probable efficiency of the proposed solution for a given problem, or the quality of certain materials indicated or produced to help solve the problem.

Peer assessment is especially indicated in situations in which large numbers of students are involved, situations in which computers are very helpful to organize pairs or groups of students to work together. Electronic submission of tasks allows for random distribution in peer assessment, and for answers to be presented to assessors anonymously, that is, respondents' names can easily be omitted. Marks can also be easily stored and compared for consistency. Thus, with the aid of computers, peer assessment can be greatly developed. Finally, peer assessment can also be used for assessing students' evaluation skills, by comparing their assessment to their teacher's assessment or to each other's.

The literature on CAA reviewed so far suggests that, in most cases, it is common to combine CAA use with manual marking than to rely on it totally for fully electronic assessment. For example, cases that are regarded as not clearly correct can be submitted to human inspection, or automatic assessment can be complemented with human assessment [18]. In a study about automated essay scoring, Dikli [19] compared feedback to ESL students' essays from automated

scoring with feedback provided by teachers in written form. It was noticed that the types of feedback differed. Teachers provided shorter and more focused feedback, while the automated essay scoring was longer, and considered generic and redundant. Such results indicate the need for improvement in computer programs used in language assessment.

The data from the TBP, however, reveals that characteristics of language assessment occur in the process of teletandem interactions, that is, while agents are involved in conversations, prior to the use of any formal instrument to assess or to test language performance or levels of language proficiency attained by the agents. Some characteristics of language assessment in teletandem interactions are dealt with in the next section.

Language assessment in teletandem interactions

Assessment in teletandem interactions occurs within the conversational process, that is, during online interactive sessions between the agents. In some cases, this assessment aims at clarifying the message and in other cases the most proficient agent helps his or her partner to learn language items that are relevant to convey meaning. Freschi [20], for example, studied types of linguistic feedback provided by most proficient agents to their partners during teletandem interactions. These language items include mainly grammar and vocabulary, as reported below and illustrated in Table 1.

In order to exemplify coherence between the views about the role of grammar in FL learning, as discussed by Brocco [21], data from teletandem interactions in the scope of the TBP has been analyzed and presented elsewhere (for example, [7]; [9]; [21]) in order to discuss difficulties faced by agents in the USA, who are most proficient agents in English, when producing the Portuguese language in particular occasions on which lack of grammar competence disturbed or impeded clear communication. When these occasions motivated explicit attention from the most proficient agents in Portuguese, or some type of request from the learners, it can be said that the quality of language use was judged. Moreover, this kind of judgment was usually corrective, or sometimes accompanied by non-corrective feedback.

Goertler, Schenker, Lesosk and Brunsmeier [22] conducted a study on success in learning through telecollaboration with focus on language outcomes and intercultural competence. The authors state that "The assessment of student learning through telecollaboration is a challenge many instructors face" (p.23). Research participants were university students of an advanced German language course in the USA and pre-service teachers of English at a university of Education in Germany. Various instruments were used to asses learning outcomes, including in-class assignments and class feedback, a role-play similar to those in the Oral Proficiency Interview conducted by ACTFL (which was rated by means of ACTFL can-do-statement assessment), and blogs and recordings from the telecollaborative project, which involved the completion of tasks performed online. According to the results from their study, the authors report that although "students were linguistically able to complete the tasks without major cases of miscomunnication during their interactions with their partners" and "clearly enjoyed the telecollaboration and saw it as beneficial", most of them "fell short of language skill goals and language production goals" (p. 33).

In this discussion I refer to registers of online interactions conducted in Portuguese and in English, between Brazilian and English-speaking agents, which occurred by means of online chat and MSN Messenger, which generated audio and video data for language analysis. The participants were university students who studied Portuguese as a foreign language at a university in the USA, and Brazilian students of a Letters course who had English as a course subject. Each participant in Brazil interacted with a participant in the USA by means of MSN in a minimum of eight teletandem sessions of one hour each.

The analysis of interactions in Brazilian Portuguese shows the types and frequency of linguistic feedback provided to an agent in the USA by a Brazilian agent. Feedback refers to all

types of reflection about linguistic items, including grammar, vocabulary, spelling, discourse and phonology. Table 1 presents the frequency of the three types of feedback found in the data:

Table 1. Types of linguistic feedback.

Type of linguistic item	Number of occurrences	Percentage
Grammar	44	28.4%
Vocabulary	78	50.3%
Spelling	8	5.2%
Discourse	21	13.5%
Phonology	4	2.6%
	Total: 155	Total: 100%

The information in Table 1 reveals that most of the linguistic feedback (50.3%) focused on vocabulary, as might be expected. Foreign language learners usually need help in learning new words and when they face a lack of words in the course of the language learning process. The amount of feedback on grammar, the focus of this investigation, was not very high. However, the frequency of feedback on grammar observed (28.4%), together with an overview of the cases of grammar mistakes raised in the corpus, suggests that grammar needs attention in foreign language learning.

Feedback may be provided by the most proficient agent when s/he notices formal deviations, lack of vocabulary or pauses in his or her partner's speech, which indicate limited language proficiency. In this sense, feedback is associated with a process of language assessment, on the part of the MPI, and with corrective techniques s/he chooses to use. Some types of feedback encountered in data from teletandem interactions [10] are explicit corrections, reformulations (recast), requests for information and requests for clarification.

The three examples presented below are from teletandem interactions conducted in English.

Example 1 below illustrates a request for clarification:

Example 1 (from Rossi dos Santos [10])

```
001-NAAg   do you have a car?
002-BrAg   no… I/ I drives my father's car
003-NAAg   you do WHAT?
004-BrAg   I/ I drives my FATHER'S car
005-NAAg   oh ((laugh))
006-BrAg   ((laugh))
007-NAAg   what kind of car is it? (…)
```

The North-American agent (NAAg) does not understand the answer given by the Brazilian agent (BrAg) in turn 002, probably because of the incorrect use of the verb form drives. The NAAg indicates the existence of a problem in the BrAg's answer by asking for clarification, in turn 003 – you do WHAT? – but the BrAg simply repeats the answer and emphasizes the word father but makes no correction to the verb form, drives (turn 004). The NAAg seems to understand the information in the BrInt's statement though and the conversation continues (turn 007). According to Rossi dos Santos [10], requests for clarification usually refer to part of a preceding statement and are contingent with a form of pronoun – you, in the case of example 1, for instance, You did what?, You saw what?, You went where?, and so on. Request for clarification can also take the form of interrogatives such as Sorry? Or I beg your pardon?, or even statements such as I don't understand.

A case of feedback on vocabulary can be seen in Example 2:

Example 2 (from Rossi dos Santos [10])

```
008-NAAg   have you (tried?) waterski?
009-BrAg   No
010-NAAg   no?
011-BrAg   WHAT?
012-NAAg   Waterski
013-BrAg   WATERSKI?
014-NAAg   yeah do you know what it is?
015-BrAg   No
016-NAAg   like you know skiing… right?
017-BrAg   ski yeah
018-NAAg   you can ski on the snow… you can ski on the water as well
019-BrAg   ah ok
```

The BrAg answers the question on waterski (turns 008 and 009) with a no but she did not know the meaning of waterski. The NAAg realizes the BrAg's lack of knowledge of the lexical item, checks whether she knows the word (turn 014) and provides an explanation (turns 016 and 018).

Example 2 also includes a case of negotiation of meaning, that is, unknown language items are discussed in the context of language used in interaction. This is in line with the view of language as a means of communication and as a support for mental processes of comprehension and production, for cognitive activities [23]. When engaged in verbal interaction, interlocutors aim at making sense and understanding the conversation and, at the same time, at thinking about the language, and testing and confirming previous linguistic knowledge. Feedback can therefore contribute to this process of analyzing language within interactive processes, and foster language development.

Example 3 illustrates the last part of a teletandem session, in which the focus is explicitly on linguistic and interactive feedback:

Example 3 (from Rossi dos Santos [10])

```
020-BrAg   (8 s) ahm… let's start to talk about my mistakes?
021-NAAg   Ok
022-BrAg   ok?
023-NAAg   you really didn't make very many
024-BrAg   a:h
025-NAAg   you/ you did good… one thing… was when you were talking about ages… you
said… my mom has… forty
026-BrAg                  [ah yeah
027-NAAg   i/ it's… she is… and then however many years old
028-BrAg   ah yes
029-NAAg   so l/ you know?
030-BrAg   ok… ok
031-NAAg   so like I'm 21 years old
032-BrAg   yeah I'm/ I'm nervous
033-NAAg   ok… don't be nervous you speak very well
034-BrAg   (laugh)
035-NAAg   you could come to America and live… fine without a problem… (2 s) your
English is very good
036-BrAg   ahm thanks
037-NAAg   and then the only others… thing that you said wrong the entire time… was you
```

said my grandfathers
    038-BrAg   ah ok
    039-NAAg and you should have said my grandparents… I think (laugh)
    040-BrAg   (laugh)
    041-NAAg and (   ) (audio failure) (4 s) can you believe it?
    042-BrAg   No
    043-NAAg yeah only two/ only two errors
    044-BrAg   yeah?
    045-NAAg yeah… did you have problems understanding?
    046-BrAg   no:… no

The linguistic problems faced by the BrAg (turns 25, 27 and 37) are reviewed and the NAAg praises his partner by saying that she has a good level of English (turn 35). The BrAg answers the NAAg's question (turn 45) and informs him that she did not "have problems understanding". This answer may refer either to the fact that her mistakes did not impede communication, or it may indicate that she was able to understand which mistakes she had made and confirm that she has learnt the correct linguistic items. As a whole, the agents seem to evaluate the level of understanding and the quality of communication in English as satisfactory.

Discussion
The literature reviewed for this article points to principles in common in both language learning and language assessment, with which I entirely agree, and to the existence of a variety of procedures and techniques available to assess language learning and language performance in electronic environments – also referred to as CAA.

    Data from teletandem interactions reveals a degree of coherence between language-learning principles and actions taken by the MPI to assess his or her partner's performance and to provide help towards language development, in a type of formative assessment. Agents follow their beliefs about language learning when they act as 'the teacher' in teletandem interactions, usually based on previous experiences in learning one or more foreign languages. Agents also act according to procedures that have been discussed or negotiated in the interactions, and try to respect each other's learning needs and preferences. It may then be stated that interactive patterns and language learning in teletandem interactions are characterized according to aspects that somehow differ from face-to-face communication in standard language classrooms. Procedures for language assessment, however, seem to be grounded on the same principles that support language assessment in standard classrooms.

    As stated in the first part of this paper, lack of linguistic competence and language mistakes are two phenomena that are commonly related to when and how teachers assess language learners in standard lessons. So, despite the fact that language use and assessment occur online, the most proficient agent's decisions and actions are similar to those in non-electronic contexts, for example, standard language lessons, on which their previous experiences in language learning are probably based. By the same token, most of the tasks suggested for CAA – for example, multiple-choice questions, textual assessment, problem-solving tasks and peer assessment, seem to be grounded on principles that support paper-and-pen(cil) language tests as well. Thus, no significant differences have so far been found to be exclusive to CAA in such a way as to support a fully new paradigm for assessment in distance learning.

    Computers, in fact, offer a variety of resources for storing and distributing data, and are helpful regarding aspects of random and anonymous assessment. Despite the contributions and innovations introduced with the use of CALL and CAA, the principles underlying language assessment do not seem to differ from those supporting standard testing and assessment conducted in standard classrooms.

Implications for language learning

Before agents engage in teletandem interactions they are oriented about some principles that should be followed during the sessions, as mentioned in the introduction to this chapter, and one of these principles states that there must be mutual co-operation between agents so that they can both benefit from the interactions inasmuch as they can share experiences, knowledge about various subjects and knowledge about the languages involved.

As for language assessment, it is suggested that agents negotiate how they prefer to be corrected when they make mistakes, for example, "on the spot" or at the end of the session. During the main part of an interaction, agents may take notes of mistakes and linguistic aspects they would like to discuss with their partner later. However, when one of the agents lacks language to express an idea, it is common that his or her partner provides help with vocabulary or grammatical structures, as illustrated in examples 2 and 3 above. Agents sometimes comment on their partners' good linguistic abilities, as illustrated in example 3 as well. These types of linguistic feedback can be seen as peer assessment in teletandem interactions and they very much resemble the types of feedback teachers usually provide in standard language lessons. A difference between formal language instruction in classrooms might be the fact that when the most proficient agent does not know how to explain a linguistic aspect s/he may tell so to his or her partner, and compromise to look for the explanation and bring it for the next session. Similar situations may also happen in standard classrooms, when teachers tell their students they will bring a(n) (better/more detailed) explanation about a teaching point or a student's question in a future lesson.

It is expected that the possibilities concerning feedback on language use in teletandem interactions presented above can contribute for language development, as well as provide a type of "teaching" experience on language assessment for more proficient agents.

Conclusion

In this study I reported on and discussed principles and aspects of language assessment in FL learning, with a focus on CALL and CAA. Data from interactions in the Teletandem Brazil Project has been used to illustrate some of my claims, and at this point I present the following conclusions.

It seems that principles for CAA combine the more traditional bases for language assessment and testing with a number of pedagogical principles that underpin distance learning, but these principles do not characterize a new paradigm in language assessment. Conversely, CALL and CAA contribute in various aspects, especially when large numbers of learners are involved in teaching and learning processes. However, the principles for language assessment followed in CAA are not significantly different from those for assessment and testing followed in more traditional teaching and learning contexts.

Further investigation is needed in order to analyze larger amounts of data, from several agents in the scope of the TBP, as well as an expanded review of the literature on CALL and CAA. This should make it possible to verify the validity of the conclusions reached in this discussion and result in the definition of a paradigm for language assessment by means of computers in online interactions.

Acknowledgements

References
[1] S. Link, J. Li, (Eds.), Assessment across online language education, Sheffield & Bristol: Equinox & CALICO, (2018).
[2] L. F. Bachman, Ongoing challenges in language assessment, in: A. J. Kunnan (Ed.), The Companion to Language Assessment, New York: John Wiley & Sons, (2014).
[3] Chapelle, C. A. (1999). Validity in language assessment. Annual Review of Applied Linguistics, 19, pp. 254-272.
[4] D. A. Consolo, P. P. Anchieta, Avaliação de proficiência em língua estrangeira em meios virtuais: Alguns princípios e características de exames e testes eletrônicos, in: P. T. C. Szundy, J. C. Araújo, C. S. Nicolaides, K. A. Silva, (Eds.), Linguística Aplicada e Sociedade: Ensino e Aprendizagem de Línguas no Contexto Brasileiro. Campinas, Brazil: Pontes Editores, (2011), pp.197-215.
[5] D. A. Schön, How professionals think in action, New York: Basic Books, (1983).
[6] J. Mezirow, Transformative dimensions of adult learning, San Francisco: Jossey-Bass, (1991).
[7] Brocco, A. S. (2007). The systematization of grammar via a deductive approach in the teaching/learning of Portuguese in a teletandem context. Report of a Scientific Initiation study. Sao Jose do Rio Preto, Brazil: UNESP, Department of Modern Languages.
[8] Brocco, A. S., "A Gramática no Contexto Teletandem e em Livros Didáticos de Português como Língua Estrangeira", MA dissertation. Sao Jose do Rio Preto, Brazil: UNESP (2009).
[9] Custódio, C. M. (2007). Final report on a scientific initiation project. Sao Jose do Rio Preto, Brazil: UNESP, Department of Modern Languages.
[10] Rossi dos Santos, G. "Características da Interação no Contexto de Aprendizagem In-Tandem", MA dissertation, Sao Jose do Rio Preto, Brazil: UNESP (2008).
[11] D. R. Garrison, T. Anderson, E-Learning in the 21st Century: A Framework for Research and Practice, Oxon & New York: Routledge Falmer, (2003).
[12] E. Shohamy, O. Inbar, The Language Assessment Process: A "Multiplism" Perspective, CALPER, (2006).
[13] H. H. Stern, Fundamental Concepts of Language Teaching, Oxford: Oxford University Press, (1983).
[14] Scaramucci, M. V. R. (2000). Proficiência em LE: considerações terminológicas e conceituais, Trabalhos em Lingüística Aplicada, 36, pp.11-22. https://periodicos.sbu.unicamp.br/ojs/index.php/tla/article/view/8639310/6904
[15] Carter J, Ala-Mutka K, Fuller U, Dick M, English J, Fone W et al. "How shall we assess this?" In Proceedings of the ITiCSE-WGR '03 Working group reports from ITiCSE on Innovation and technology in computer science education, ACM Press, (2003). pp. 107-123.
[16] Burstein, J., Leacock, C & Swartz, R. (2001). Automated Evaluation of Essays and Short Answers. Proceedings of 5th Annual Conference on Computer-Aided Assessment, Loughborough, UK.
[17] Mason, O. & Grove-Stephenson, I. (2002). Automated free text marking with Paperless School. Proceedings of 6th Annual Conference on Computer-Aided Assessment, Loughborough, UK.
[18] Jackson, D. (2000). A Semi-Automated Approach to Online Assessment. Proceedings of ITiCSE'00, Helsinki.
[19] Dikli, S. (2010) The Nature of Automated Essay Scoring Feedback. CALICO Journal 28 (1), pp.99-134.
[20] Freschi, A. C. (2017) A Avaliação por Pares no Teletandem Institucional Integrado: um

Estudo de Caso sobre o Feedback Linguístico nas Sessões Orais em Português. MA dissertation. Sao Jose do Rio Preto, Brazil:UNESP.

[21] Consolo, D. A., Brocco, A. S. & Custodio, C. M. (2009). Grammar and communication in Portuguese as a foreign language: A study in the context of teletandem interactions. Proceedings of the 1st International CSEDU Conference, Lisbon, Portugal (CD-ROM), pp.62-66.

[22] Goertler, S.; Schenker, T.; Lesosk, C. & Brunsmeier, S. (2018) Assessing Language and Intercultural Learning during Telecollaboration. In Link, S. & Li, Jinrong (Eds.), Assessment Across Online Language Education, Sheffield & Bristol: Equinox and CALICO, pp.21-48.

[23] Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. Modern Language Journal 82 (3), pp.320-337.

[24] Yang, C. C. R. (2008). Second Language Classroom Interaction Patterns: An Investigation of Three Case Studies. IABR & TLC Conference Proceedings. San Juan, Puerto Rico, USA.

[25] Walsh, S. (2006). Investigating Classroom Discourse. New York: Routledge.

[26] Bloom, B. (1956). Taxonomy of Educational Objectives: The classification of Educational Goals: Handbook I. New York: Longman.

[27] Furtoso, V. A. B. (2009). Para além do Gostei muito da conversa: avaliação no contexto de aprendizagem in-tandem. In J. A. Telles (Ed.). Teletandem: um contexto virtual, autônomo e colaborativo para a aprendizagem de línguas estrangeiras no século XXI, pp.297-314, Campinas, Brazil: Pontes Editores.