Gomes de Freitas Menezes Martins, Cristiana; Levis, John Michael; Cabral Borges,
Vládia Maria

The design of an instrument to evaluate sotware for EFL/ESL pronunciation teaching

The design of an instrument to evaluate software for EFL/ESL pronunciation teaching

Cristiana Gomes de Freitas Menezes Martins[*]
Universidade Federal do Ceará
Fortaleza, CE, BR

John Michael Levis[**]
Iowa State University
Ames, IA, USA

Vládia Maria Cabral Borges[***]
Universidade Federal do Ceará
Fortaleza, CE, BR

Abstract

This study aimed at designing an instrument composed of 72 questions that evaluate the extent to which software programs teach pronunciation of English as a Foreign Language and/or Second Language (EFL/ESL) following the principles of the Communicative Approach (Celce-Murcia et al, 2010). The designed instrument was tested for reliability and validity. For this purpose, 46 EFL/ESL teachers used the instrument to analyze an online version of the software program *Pronunciation Power 2*. The participants' answers to the questions were subjected to reliability and validity tests. The results of these statistical tests suggest the instrument is potentially valid for evaluating EFL/ESL pronunciation teaching software.

Keywords: Pronunciation Teaching Software Evaluation; Communicative Approach; Computer-Assisted Pronunciation Teaching

Introduction

With the advent of globalization, there is a growing need to communicate with people of different nationalities. This communication includes the use of oral language, and helps explain the increasing diffusion of English teaching and learning throughout the world. According to the British Council (2012), a non-ministerial department of the UK government responsible for education, one out of four of the world's population speaks English with some level of competence. English is the primary language used for international contacts and business transactions. Thus, English has become an international language

---

[*] Cristiana Martins teaches English as Foreign Language in undergraduate programs at private colleges. She has a doctoral degree in Linguistics from the Federal University of Ceará, Brazil with a "sandwich" period at Iowa State University, USA. Her research interests are pronunciation instruction, and computer assisted pronunciation teaching. Email: cristiana.martins@uol.com.br

[**] John Levis is Professor of Applied Linguistics and TESL at Iowa State University, USA. His articles on pronunciation and intonation have been published in a variety of professional journals. He was co-editor for *Social Dynamics in Second Language Accent* (De Gruyter Mouton) and the *Handbook of English Pronunciation* (Wiley). He initiated the annual *Pronunciation in Second Language Learning and Teaching* conference and is founding editor of the new *Journal of Second Language Pronunciation* (John Benjamins). Email: jlevis@iastate.edu

[***] Vládia Borges is Associate Professor of English as Foreign Language in the Department of Foreign Languages and of Applied Linguistics in the Linguistics Graduate Program at the Federal University of Ceará, Brazil. She has a Ph.D. in Education with a major in Teaching English as a Second Language, from the University of Rhode Island, USA (2006). She has been the dean of the School of Humanities since 2011. Her research interests are second language acquisition and processing and computer assisted language learning and teaching. Email: vladiaborges@gmail.com

as described by Jenkins (2000) and Walker (2010), among others.

Indeed, studying English has become a basic need for acceptance and qualification in almost all professions. Aware of the growing number of people interested in learning this language worldwide, the publishing industry has released thousands of books, CDs, DVDs and software programs claiming to help learners master English. In order to make their business more profitable, the publishing industry is increasingly seeking to tailor their materials to particular audiences (Assumpção Filho, 2011).

With the advent of computer technology and growing technological appeal, the use of software programs has been one more option among the set of materials used when teaching and prioritizing pronunciation. Levis wrote that

> the use of computers is almost ideally suited to learning pronunciation skills. Computers can provide individualized instruction, frequent practice through listening discrimination and focused repetition exercises, and automatic visual support that demonstrates to learners how closely their own pronunciation approximates model utterances. (2007, p.184)

Given these facts, software programs for developing English pronunciation have increasingly been launched in the market. Publishers have promoted these programs stating they are effective and of good quality. However, when it comes to the advertising discourse of English language teaching publishers, promotion is not always factual, as investigated by Carvalho (2011). Teachers and learners should not be seduced by the strong appeal of the marketing done by publishers. Instead, it is necessary to analyze English as a Foreign Language and/or Second Language (EFL/ESL) pronunciation teaching software programs as to their potential for developing English pronunciation. There is an unquestionable need to analyze these programs from a critical perspective using pedagogically coherent and technically elaborated criteria (Navarro, 1999).

While several studies sought to establish criteria and evaluation instruments for analyzing educational software programs in general, others investigated criteria and evaluation instruments for analyzing language learning software programs. Table 1 provides an overview of some of the studies done to establish criteria and evaluation instruments for analyzing educational software programs.

Table 1

Studies on Criteria and Instruments for Analyzing Educational Software Programs

General educational software programs		Language learning software programs	
Figueiredo (2000)	Seven evaluation categories: 1) Traditional Techniques; 2) Strategies; 3) Metaphors; 4) Goals; 5) Picture; 6) Sound; 7) Animated screensavers.	Burston (2003)	Four evaluation categories: 1) Technical features; 2) Activities (Procedure); 3) Teacher fit (Approach); 4) Learner fit (Design).
Jackson (2000)	Six evaluation categories: 1) Platform requirements; 2) Goals and objectives; 3) Content; 4) Pedagogy; 5) Ease of use; 6) Costs.	Jamieson, Chapelle and Preiss (2005)	Six evaluation categories: 1) Language learning potential; 2) Learner fit; 3) Meaning focus; 4) Authenticity; 5) Impact; 6) Practicality.

Oliveira, Costa and Moreira (2001)	Four evaluation categories: 1) Interaction student-software program-teacher; 2) Pedagogical basis; 3) Content; 4) Programming.	Borges (2006)	Three evaluation categories: 1) Technological features; 2) Pedagogical features; 3) Individualized-learning features.
Fino (2003)	Criteria for the use of educational software programs were more important than evaluation criteria and the teacher.	Hubbard (2006)	Six evaluation categories: 1) Technical preview; 2) Operational description; 3) Teacher fit; 4) Learner fit; 5) Implementation schemes; 6) Appropriateness judgments.
Lacerda (2007)	The necessity of individualized evaluation criteria due to the use of software programs with different kinds of content in different contexts and by different users.	Chapelle and Jamieson (2008)	Five evaluation categories: 1) Learner fit; 2) Explicit teaching; 3) Interaction with the computer; 4) Evaluation and visual feedback; 5) Strategy development.
Abreu (2010)	Evaluation categories derived from software engineering and educational theories in order to promote technical and pedagogical usability of software programs.	Zardini (2009)	Utilized three of the four evaluation categories proposed by Oliveira, Costa and Moreira (2001) to evaluate EFL/ESL software programs: 1) Interaction student-software program-teacher; 2) Pedagogical basis; 3) Content.

The analysis of these studies indicates that there are no specific criteria for evaluating pronunciation software programs, although some studies, such as Chapelle and Jamieson (2008), establish criteria for evaluating the skill of speaking in general. Therefore, there is a need for studies that attempt to establish specific criteria to evaluate programs for teaching pronunciation. There might be pronunciation software programs whose interface may look attractive but fail at reflecting solid grounded principles for teaching pronunciation.

Thus, this study aimed at designing an instrument to evaluate the extent to which software programs teach EFL/ESL pronunciation following the principles of the Communicative Approach (Celce-Murcia, Brinton, Goodwin, & Griner, 2010). More specifically, this study focused on answering the following questions:

1. How reliable is the EFL/ESL pronunciation teaching software program evaluation instrument? That is to say, does it generate similar results when used again in similar circumstances?

2. How valid is the instrument in evaluating the extent to which an EFL/ESL pronunciation teaching software program follows the principles of the Communicative Approach (Celce-Murcia et al, 2010)?

Theoretical framework

Because no criteria for the evaluation of pronunciation software programs currently exist, the purpose of this study was to design an instrument to evaluate the extent to which software programs designed

for teaching English pronunciation to FL/SL learners follow the principles of the Communicative Approach (Celce-Murcia et al, 2010). The designed instrument was tested for reliability and validity. Underlying this process was the belief that software programs should be based on an understanding of Second Language (SL)[1] acquisition, more specifically SL phonological acquisition, and the principles of the Communicative Approach (Celce-Murcia et al, 2010). In this section, the principles which were used to ground the instrument are briefly described.

SL Phonological Acquisition

Phonology brings together a set of segments (phonemes) that can distinguish meaning. It is by combining these segments that spoken language is perceived and produced by an individual. As the acquisition of lexical, syntactic, semantic and pragmatic aspects, the acquisition of phonological aspects is also essential in the perception and production of spoken language (Pennington, 1999).

Research studies have been done to explain how SL phonological acquisition occurs. In this field, one of the most influential researchers is James Emil Flege. Flege (1995, 2002, 2003) developed the Speech Learning Model (SLM), which posited that an individual's perception level of phonological aspects was positively related to the accurate production of these aspects.

Another model that deals with SL phonological acquisition is the Perceptual Assimilation Model (PAM) proposed by Best (1995) and extended by Best and Tyler (2007). In the second model (PAM-L2), Best and Tyler (2007) stated that adults perceived phonological aspects of a SL by comparing the similarities and differences of the SL with the first language (L1), especially when they had little contact with the SL. It was through this contact that phonological aspects were perceptually learned.

The model proposed by Best and Tyler (2007) shows the relevance of perceiving distinctive language sounds to SL phonological acquisition, while the model proposed by Flege (1995, 2002, 2003) shows the relevance of perceiving and producing these sounds to SL phonological acquisition. Therefore, software programs designed for teaching English pronunciation to Foreign

Language/Second Language (FL/SL) learners should have activities that work on the perception and production of different phonological aspects of the English language. Figure 1 shows examples of the questions in the designed instrument that analyze the activities that develop sound perception and production (for more examples of the questions that analyze this characteristic, please refer to the Evaluation Instrument in the Appendix).

Questions in the Instrument	The ACTIVITIES work on:	
2.22	the perception of sounds.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.23	the production of sounds.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 1[2]. Questions for analyzing activities for sound perception and production.

The Place of Pronunciation in a Communicative Approach

Most researchers agree that communicative competence does not necessarily mean to sound like a native speaker of a language. Instead, the goal of teaching pronunciation is to have the learners acquire intelligible pronunciation, in other words, being able to make themselves understood by others without communication breakdown (Jenkins, 2000). Pronunciation teaching in this approach is also associated with the learners performing interactive pronunciation activities in the classroom. The teacher can make use of technical explanations, but it is important that learners have the opportunity to interact with others while making use of both segmental and suprasegmental aspects of language (see Figure 2). Suprasegmental aspects are as important as segmental ones and both should be taught in a communicative approach (Celce-Murcia et al, 2010; Levis & Levelle, 2009).

Questions in the Instrument	The program addresses:	
1.3	the vowel sounds.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
1.10	sentence stress.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 2. Questions in the instrument for analyzing the presentation of segmental and suprasegmental aspects.

Celce-Murcia et al (2010) proposed a communicative framework for teaching pronunciation and suggested that a pronunciation lesson should be divided into five phases:

> **1. DESCRIPTION AND ANALYSIS** – *oral and written illustrations of how the feature is produced and when it occurs within spoken discourse;* **2. LISTENING DISCRIMINATION** – *focused listening practice with feedback on learners' ability to correctly discriminate the feature;* **3. CONTROLLED PRACTICE** – *oral reading of minimal-pair sentences, short dialogues, etc., with special attention paid to the highlighted feature in order to raise learner consciousness;* **4. GUIDED PRACTICE** – *structured communication exercises, such as information-gap activities or cued dialogues, that enable the learner to monitor for the specified feature;* **5. COMMUNICATIVE PRACTICE** – *less structured, fluency-building activities (e.g., role-play, problem solving) that require the learner to attend to both form and content of utterances.* (p.45)

Celce-Murcia et al (2010) also stated that it is essential for the learners to get systematic feedback in all these phases (see Figure 3).

Questions in the Instrument	The program uses:	
2.20	listening comprehension activities.	$\bigcirc\ 0\ \bigcirc\ 1\ \bigcirc\ 2\ \bigcirc\ 3\ \bigcirc\ 4$ (barely) (completly)
Questions in the Instrument	Regarding the feedback of the ACTIVITIES:	
3.38	it gives explanations about the user's errors.	$\bigcirc\ 0\ \bigcirc\ 1\ \bigcirc\ 2\ \bigcirc\ 3\ \bigcirc\ 4$ (barely) (completly)

Figure 3. Questions in the instrument for analyzing feedback for activities.

The use of phonetic symbols is also considered valuable in pronunciation teaching (Kelly, 2000), (see Figure 4). Apart from helping learners visually perceive pronunciation, the use of phonetic symbols can help them become more autonomous in their learning (Celce-Murcia et al, 2010).

Questions in the Instrument	The program uses:	
2.18	phonetic symbols to present and practice vowel and consonant sounds.	$\bigcirc\ 0\ \bigcirc\ 1\ \bigcirc\ 2\ \bigcirc\ 3\ \bigcirc\ 4$ (barely) (completly)

Figure 4. Questions in the instrument for analyzing the use of phonetic symbols and transcriptions.

Another relevant pedagogical procedure is to contrast the learners' mother tongue sounds to the target language sounds (see Figure 5). This procedure may help the learners raise awareness of the differences between the two languages, identify which sounds are more difficult and set strategies to work on typical errors (Dale, 2001).

Questions in the Instrument	Regarding the program:	
4.57	it contrasts the user's mother tongue sounds to the English sounds.	$\bigcirc\ 0\ \bigcirc\ 1\ \bigcirc\ 2\ \bigcirc\ 3\ \bigcirc\ 4$ (barely) (completly)

Figure 5. Questions in the instrument for analyzing the contrast of the user's mother tongue sounds to the English sounds.

Exposing the learners to highly variable input leads to gains in pronunciation learning (Bybee & Torres Cacoullos, 2008; Thompson, 2011). The learners should be aware of different regional phonological varieties to understand the English spoken by different speakers of English (Kelly, 2000). They should also develop intelligible rather than native-like pronunciation (Celce-Murcia et al, 2010; Levis & Levelle, 2009). Acquiring a pronunciation that can be understood by any other speaker of English is a more realistic goal for pronunciation teaching (see Figure 6).

Questions in the Instrument	Regarding the program:	
4.58	it incorporates highly variable input (e.g., regional phonological varieties, male and female voices).	$\bigcirc\ 0\ \bigcirc\ 1\ \bigcirc\ 2\ \bigcirc\ 3\ \bigcirc\ 4$ (barely) (completly)

Figure 6. Questions in the instrument for analyzing the incorporation of highly variable input.

Therefore, software programs designed for teaching English pronunciation to FL/SL learners that follow the principles of the Communicative Approach (Celce-Murcia et al, 2010) should explore pedagogical characteristics which can be identified in the communicative framework as just described.

Ways to Best Utilize the Potential of Computer-Assisted Pronunciation Teaching (CAPT)

After decades of research on language teaching, we have come to a time when digital technologies

enrich learning and assist in the acquisition process. Levis (2007) stated that "for any teacher who thinks that pronunciation is essential, Computer-Assisted Pronunciation Teaching (CAPT) is immensely promising" (p. 196).

In fact, there are many advantages to the use of CAPT. When comparing the teaching of pronunciation mediated by a phonetician or a teacher of the same area, Pennington (1996) argued that CAPT may prove to be superior in some aspects. With the computer, it is possible to analyze the users' pronunciation and give them feedback faster than a human does. Since people usually suffer from limitations such as patience, hearing and judgment, this analysis via computer can be performed more accurately and more reliably an infinite number of times (see Figure 7).

Questions in the Instrument	Regarding the ACTIVITIES:	
3.45	they can be redone by the user.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 7. Questions in the instrument for analyzing the redo of the activities by the user.

The computer may present various phonological aspects in more ways than a human can. Thus, the users can easily receive different types of input, become more conscious about their pronunciation, have a better understanding of many phonological aspects and increase their motivation to work on their pronunciation (Pennington, 1996).

In addition, the computer may be able to provide more prominent feedback than a human, and this prominence arises from the use of visual aids such as videos, animations, pictures, graphics, sounds, and other multimodal presentations that only a computer can provide (Pennington, 1996), (see Figure 8).

Questions in the Instrument	Regarding the program:	
4.52	it uses animation to demonstrate the production of sounds.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 8. Questions in the instrument for analyzing the use of visual aids.

The computer can individualize the teaching of pronunciation, allowing the users themselves to choose the phonological aspects they want to work on or suggesting which aspects the users need to improve according to the pronunciation analysis made by the computer. The users may also receive input from the computer whenever and wherever they need. The computer can also give the users the opportunity to take responsibility for their own learning, which, according to Benson (2007), is a key element in the development of learner autonomy. By developing autonomy, the learners can compensate classroom learning limitations and boost pronunciation learning without the presence of a teacher.

However, Chapelle and Jamieson (2008) advised that the users need guidance for choosing what to learn and how to learn and it is the teacher's job to provide them with the necessary guidance for selecting and evaluating instructional materials for CALL based on the teacher's expertise and experience in language teaching and learning (see Figure 9).

Questions in the Instrument	Regarding the ACTIVITIES:	
3.44	the user can choose their content.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 9. Questions in the instrument for analyzing the possibility of choosing the content of the activities.

Even with so many positive attributes that the computer has to help in the development of pronunciation, CAPT still has its limitations: many educational software programs are attractive, but do not meet educational requirements (Neri, Cucchiarini & Strik, 2002), while others work on pronunciation in a decontextualized way presenting only mechanical exercises with no focus on meaning (Pennington, 1996), (see Figure 10).

Questions in the Instrument	The ACTIVITIES work on:	
2.24	the perception/production of sounds in a contextualized way.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

Figure 10. Questions in the instrument for analyzing the use of contextualized activities.

There are software programs that present only a single pronunciation model, either a female or a male one and a single regional phonological variant,

limiting the input received by the users as well as their perception and production. As asserted by Cardoso, Smith and Garcia Fuentes (2015) and Soler-Urzua (2012), different kinds of quality input can enhance learners' perception and production. Thus, the learners need to be exposed to different models in order to be able to recognize different patterns of pronunciation (see Figure 11).

Questions in the Instrument	Regarding the ACTIVITIES:	
3.47	they allow the user to choose between a male or female voice.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completely)

Figure 11. Questions in the instrument for analyzing the possibility of choosing a pronunciation model.

The pronunciation feedback provided to the users should be easily understandable. The users should be able to understand through the feedback received by the computer which phonological aspects they need to work on. Although feedback can be provided in several ways by the computer, the most common ones are those that utilize visual displays such as spectrograms, waveforms and pitch tracings. Many of these visual displays require some sort of practice to help teachers build a kind of knowledge that they, in general, do not have (Levis, 2007), (see Figure 12).

Questions in the Instrument	Regarding the program:	
4.56	it uses visual acoustic displays (e.g., waveforms, spectrograms, pitch tracings) to provide immediate feedback on the user's speech recordings.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completely)

Figure 12. Questions in the instrument for analyzing the use of visual displays.

However, in order for the computer to analyze the users' pronunciation and provide feedback, an Automatic Speech Recognition (ASR) Mechanism is necessary. Derwing, Munro and Carbonaro (2000) tested different versions of ASRs and found out that these mechanisms could analyze quite accurately (95%) the pronunciation of English native speakers, but when analyzing the pronunciation of non-native speakers, the level of accuracy dropped to approximately 70%.

According to these researchers, the problem with the inaccuracy of these mechanisms was that they were not designed for non-native speakers.

A recent study by Liakin, Cardoso and Liakina (2015) tested the use of ASR for SL pronunciation instruction in a mobile-assisted learning environment. Forty-two learners were divided into three groups: 1) a group using ASR to do pronunciation activities; 2) a group not using ASR; and 3) a control group. The results of this study showed that only the learners in the group using ASR improved their pronunciation significantly. Therefore, it is relevant to evaluate ASR mechanisms in software programs. The following question checks the ability of ASR devices to provide feedback (Figure 13).

Questions in the Instrument	The program uses:	
5.67	Automatic Speech Recognition to provide immediate feedback on the user's pronunciation.	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completely)

Figure 13. Questions in the instrument for analyzing the use of ASR.

CAPT can also take place in language labs, allowing the teacher to monitor multiple users on different computers at the same time. The teacher can guide users on how to use the computer, revise their pronunciation analysis made by the computer, compare the analysis of one user with the others as well as provide feedback.

Therefore, EFL/ESL pronunciation teaching software programs should explore technological characteristics effective for computer-mediated interactive pronunciation learning as just discussed.

Method

In order to design an instrument with questions that encompass characteristics to be observed during the evaluation of software programs for teaching pronunciation to EFL/ESL learners, criteria and the evaluation models reported in Table 1 were reviewed.

The following criteria were used in the design of the instrument (see Appendix): 1) the presence of pedagogical characteristics consistent with the Communicative Approach (Celce-Murcia et al, 2010) and 2) the presence of technological characteristics

effective for computer-mediated interactive pronunciation learning.

Based on these criteria, 72 questions were elaborated. It is noteworthy that the instrument was divided into two parts: 1) the Descriptive Analysis, which describes the technical and pedagogical characteristics of the EFL/ESL pronunciation teaching software program provided by the developer/distributor and 2) the Critical Analysis, which analyzes the extent to which the software program follows the principles of the Communicative Approach (Celce-Murcia et al, 2010).

To assess the extent to which the software program meets the characteristics described in the questions, a rating scale of 0-4 was used, where 0 indicates absence of the characteristic and 4 indicates a completely satisfying exploration of the characteristic (See Figures 1 to 13 for examples of questions in the instrument with this rating scale). The greater the overall rating of the software program, the more the potential to improve the pronunciation of the English language learner.

In order to test the reliability and the internal consistency of the instrument and make sure that it has some validity, 46 participants used the instrument to analyze an online version of the software program *Pronunciation Power 2*[3], because it didn't require any installation procedure. An online version of the instrument was also used, because it allowed the researchers to get the responses provided by the informants by email.

The participants were selected based on the following criteria: 1) teachers of English teaching in language schools, secondary schools or colleges, and 2) teachers graduated in English and/or with a specialization, Master's Degree and/or Ph.D. in English, linguistics and/or applied linguistics. 779 teachers were invited to be part of the study by email, but only 58 sent their evaluation responses. These 58 teachers were first divided into two groups according to their context of teaching English - foreign language and second language teaching. Then they were grouped according to their level of expertise in English phonetics and phonology – teachers and experts. These procedures resulted in three groups: Group 1

– Brazilian EFL teachers; Group 2 – Brazilian experts in English phonetics and phonology; and Group 3 – Foreign experts in English phonetics and phonology.

However, from the 58 teachers, only the responses sent by 46 were considered. Twelve teachers had to be discarded because 10 of them assigned 3 and/or 4 points to all questions of the instrument and two did not fit into any of the three groups. Thus, in this study, the responses sent by 46 teachers/participants were analyzed and the teachers were divided as follows: 27 participants in Group 1; 11 participants in Group 2; and 8 participants in Group 3. The participants did not receive any training nor had a time limit to complete the instrument.

The instrument was submitted for the analysis of an expert in English phonetics and phonology for face and content validity examination. The data collected through the responses of the participants were statistically analyzed to determine the degree of instrument validity for its purpose. Thus, tests of reliability and internal consistency of the instrument were performed: Intra-Class Correlation Coefficient (ICC); One-Way Analysis of Variance (ANOVA); Cronbach's alpha coefficient; and Factor Analysis. Some procedures of descriptive statistics were used to describe and summarize the results collected from responses sent by the 46 participants of this study when analyzing the software program *Pronunciation Power 2*.

Results

The results of this study are presented in three sections. In section one, the results for face and content validity of the instrument are described; in the second section, the results for the reliability of the instrument are presented; and in section three, the results for the internal consistency of the instrument are described.

Results for Face and Content Validity of the Instrument

The researchers also had the opportunity to submit the instrument for analysis by an expert[4] in English phonetics and phonology. The criteria that guided the

design of the instrument were given to this expert. After having analyzed and used the instrument himself, the expert assured that the instrument "on its face" seemed to be appropriate to analyze if a software program met the criteria.

Results for the Reliability of the Instrument

The first statistical procedure used to analyze the level of agreement among the participants' responses was the ICC. The ICC found for the three groups of participants was .983 and the level of confidence at 95% was .976 to .988. These results indicated a high degree of correlation among the ratings of the 46 participants despite the fact that they teach English in different contexts and have different levels of expertise in English phonetics and phonology (see Table 2).

Table 2
ICC of the Instrument

Intraclass Correlation	95% Confidence Interval	
	Lower Bound	**Upper Bound**
.983	.976	.988

The second statistical procedure performed was the One-Way ANOVA,[5] which was used to compare the responses given by the three groups of participants: Group 1 (Brazilian EFL teachers); Group 2 (Brazilian experts in English phonetics and phonology); and Group 3 (Foreign experts in English phonetics and phonology). These three groups form the independent variable while the total rates for the 67 questions[6] of the instrument are the dependent variable.

Despite the differences in the English teaching context and the level of expertise in English phonetics and phonology of the three groups, the results of the one-way ANOVA indicated no significant difference (*p* > .05) in the evaluation of 62 of the 67 questions for the three groups. This means that there was an agreement of 95%, *p* > .05, in the evaluation of 62 questions among the three groups of participants.

The five questions with *p* less than .05, that is, with the lowest level of agreement, are Questions 1.1, 1.2, 1.6,

2.14 and 2.26. Concerning Questions 1.1 and 1.2 (*p* = .008 and .004, respectively), some participants reported that the characteristics described in those questions were not found in the software program analyzed. Therefore, they had to assign zero to those questions. The objectives and the proficiency level proposed were only described on the developer's website and could only be found by navigating to the software program developer's website before logging in to the program. As these pieces of information were not available in the program itself, some participants may have not navigated to the developer's website. This may have led to a significant difference in the evaluation of the characteristics described in those two questions. As to the other three questions – 1.6, 2.14 and 2.26 – the *p* found, respectively .039, .048, .025, although less than .05, does not point to a large discrepancy in evaluating the characterisitics described in those three questions.

In the One-Way ANOVA, the Post Hoc Fisher's test comparing the least significant difference (LSD) was also performed to locate differences in the *p* value <.05. The results of the post-hoc test indicated that the significant difference for these five questions always occurred between the group of Brazilian EFL teachers and the group of Brazilian experts in English phonetics and phonology or the group of teachers and the group of foreign experts in English phonetics and phonology, but never between the two groups of experts (see Table 3).

Table 3
Questions with p < .05 (ANOVA: Post-Hoc Fisher'sTest)

Dependent Variable			**Sig.**
1.1 the proposed objectives.	Teacher	Brazilian Expert	.002
		Foreign Expert	.242
	Brazilian Expert	Teacher	.002
		Foreign Expert	.148
	Foreign Expert	Teacher	.242
		Brazilian Expert	.148

1.2 the proposed proficiency level.	Teacher	Brazilian Expert	.120
		Foreign Expert	.001
	Brazilian Expert	Teacher	.120
		Foreign Expert	.074
	Foreign Expert	Teacher	.001
		Brazilian Expert	.074
1.6 syllabic constituents (e.g., initial consonant clusters, final consonant clusters).	Teacher	Brazilian Expert	.712
		Foreign Expert	.012
	Brazilian Expert	Teacher	.712
		Foreign Expert	.053
	Foreign Expert	Teacher	.012
		Brazilian Expert	.053
2.14 contrasts different vowel sounds.	Teacher	Brazilian Expert	.021
		Foreign Expert	.147
	Brazilian Expert	Teacher	.021
		Foreign Expert	.577
	Foreign Expert	Teacher	.147
		Brazilian Expert	.577
2.26 the perception of rhythm, stress and intonation.	Teacher	Brazilian Expert	.157
		Foreign Expert	.009
	Brazilian Expert	Teacher	.157
		Foreign Expert	.214
	Foreign Expert	Teacher	.009
		Brazilian Expert	.214

Perhaps Questions 1.6, 2.14 and 2.26 inquire about specific characteristics of the program, requiring a greater knowledge of English phonetics and phonology to be analyzed. This may explain the significant difference found between the evaluation of the group of teachers and the group of Brazilian and foreign experts in those three questions.

Results for the Internal Consistency of the Instrument

The first statistical procedure used to analyze the degree of internal consistency of the instrument was the Cronbach's alpha coefficient. The Cronbach's alpha coefficient obtained for the 67 questions in the instrument was equal to .918, indicating a high degree of internal consistency of the instrument.

The second statistical procedure performed was a factor analysis. Through factor analysis, we tried to determine the existing components (factors) among the 67 questions (variables) of the instrument and the correlation of each of the questions in those components (factors). A principal components analysis extraction and promax rotation with Kaiser normalization method was used and factors with eigenvalues > 1 were retained in this analysis to reduce the amount of observed variables into a smaller number of factors.

This extraction method resulted in 18 components with initial eigenvalues > 1 with the first five components explaining 50.48% of the total variance of the questions of the instrument. The first component – Factor 1 – with initial eigenvalue of 15.95 explained 23.81% of the total variance; the second component – Factor 2 – with initial eigenvalue of 5.63 explained 8.41%; the third component – Factor 3 – with initial eigenvalue of 4.72 explained 7.05% of the total variance; the fourth component – Factor 4 – with initial eigenvalue of 3.89 explained 5.81% of the total variance; and the fifth component – Factor 5 – with initial eigenvalue of 3.60 explained 5.37% of the total variance. The other 13 components explained less than 5% of the total variance of the questions of the instrument and therefore were not retained in this analysis.

After the retention of five components in the first stage of factor analysis, we calculated the factor loading of each item of the instrument regarding these factors to verify which questions would be grouped into each of them.

The few questions with factor loading > .30 in the 13 excluded components also loaded > 0.30 in one of five components retained. Thus, these five components with factor loading > 0.30 were analyzed in order to identify common themes among them and label them.

The first category – Factor 1 – was labeled "*Pedagogical Design*". The questions in this factor are related to the communicative approach of teaching EFL/ESL pronunciation. This factor includes 24 questions related to the way the program presents the content (Questions 2.14 to 2.21), the types of pronunciation activities (Questions 2.22 to 2.33) and attributes that facilitate navigation through the program (Questions 2.34 to 2.37). The second category – Factor 2 – was labeled "*Multimedia Design*". Its 19 questions (Questions 4.48 to 4.66) are related to the technological characteristics that the program offers.

The third category – Factor 3 – was labeled "*Assessment/Flexibility Design*". The 10 questions in this factor are related to the way feedback of activities is given (Questions 3.38 to 3.42) and how one can tailor the program to the users' needs (Questions 3.43 to 3.47). The fourth category – Factor 4 – was labeled "*Content Design*". In this fourth factor, the 13 questions are related to the content that the program addresses, including the objectives and the level of proficiency proposed (Questions 1.1 and 1.2), as well as segmental (Questions 1.3 to 1.9) and suprasegmental (Questions 1.10 to 1.13) aspects addressed. The fifth category – Factor 5 – was labeled "*ASR Design*", since the questions in this factor are related to the ASR mechanism that the program uses (Questions 5.67 to 5.72).

Due to some comments made by participants about the meaning of the term "ASR mechanism", which was confused with visual displays such as spectrogram, waveform, and pitch tracing, we decided to add a new item (Item 4.56) to the "*Multimedia Design*" component. This new item asks if the program uses visual acoustic displays (e.g., waveforms, spectrograms, pitch tracings) to provide immediate feedback on the users' speech recordings.

Thus, after the factor analysis, the 72 questions of the instrument were grouped into 5 components (factors) and arranged in the following order: 1. Content Design; 2. Pedagogical Design.; 3. Assessment/Flexibility Design; 4. Multimedia Design; and 5. ASR Mechanism Design (see the final version of the EFL/ESL Pronunciation Teaching Software Evaluation Instrument in Appendix).

Discussion and conclusion

The analyses of the data collected suggest the following answers to the questions raised in this study.

1. How reliable is the EFL/ESL pronunciation teaching software program evaluation instrument? That is to say, does it generate similar results when used again in similar circumstances?

The ICC showed a high degree of correlation among the ratings of the 46 participants despite the fact that they teach English in different contexts and have different levels of expertise in English phonetics and phonology. The results of the One-Way ANOVA also showed that there was no significant difference in the evaluation of 62 of the 67 questions of the instrument for the three groups of participants (Brazilian EFL teachers, Brazilian experts and foreign experts in English phonetics and phonology). The Cronbach's alpha coefficient revealed a high degree of internal consistency of the instrument. The results of the Factor Analysis suggested five factors in which 72 questions were grouped. All these results ensure the instrument has a high degree of reliability.

2. How valid is the instrument in evaluating the extent to which an EFL/ESL pronunciation teaching software program follows the principles of the Communicative Approach (Celce-Murcia et al, 2010)?

The face and content validity of the instrument attested by the expert in English phonetics and phonology and the results from the statistical analysis of the data suggest that the instrument is potentially valid for evaluating the extent to which software programs designed for teaching English pronunciation to FL/SL learners follow the principles of the Communicative Approach (Celce-Murcia et al, 2010).

The results of this study indicate that it is possible to evaluate the degree to which an EFL/ESL pronunciation teaching software program uses the principles of the

Communicative Approach (Celce-Murcia et al, 2010). By analyzing each of the 72 questions of the instrument, it is possible to assess the extent to which the program presents pedagogical characteristics consistent with the principles of the Communicative Approach (Celce-Murcia et al, 2010) and technological characteristics effective for computer-mediated interactive pronunciation learning. Thus, this instrument can be widely used by teachers and researchers when analyzing and/or developing EFL/ESL pronunciation teaching software programs.

Some limitations of this study need to be mentioned. The participants of the study, who were Brazilian EFL teachers, Brazilian experts and foreign experts in English phonetics and phonology represent only a portion of the entire population of EFL/ESL teachers, and experts. EFL/ESL teachers from other countries were not represented in this sample. Further studies may include a more diverse group of participants.

In this study, only one EFL/ESL pronunciation teaching software program was used to evaluate the reliability and validity of the instrument. Future studies may use more programs to evaluate the reliability and validity of the instrument. Besides, this instrument was designed to evaluate only EFL/ESL pronunciation teaching software programs. Since there are also websites and cell phone and tablet applications designed for teaching English pronunciation to FL/SL learners, future studies may design instruments to evaluate websites and applications as well.

The data collected through the responses of 46 participants were statistically analyzed in this study. In order to evaluate the reliability and validity of the instrument, tests of reliability and internal consistency of the instrument were performed: ICC; One-Way ANOVA; Cronbach's alpha coefficient; and Factor Analysis. The same statistical tests as well as other statistical tests may be employed with data collected from a larger sample in additional studies.

Finally, additional studies may evaluate the extent to which a group of learners improve their pronunciation when using an EFL/ESL pronunciation teaching software program that follows the principles of Communicative Approach (Celce-Murcia et al, 2010),

then evaluate the same program using the instrument and contrast both results.

Notes

1. Although in the literature the concepts of Second Language (SL) and Foreign Language (FL) are sometimes distinguished, in this paper they are used interchangeably.

2. The figures in this paper will show only some examples of the questions in the instrument. For a total perception of all questions, the whole Evaluation Instrument should be analyzed in the Appendix.

3. According to the publisher, *Pronunciation Power 2* is designed for intermediate to advanced English learners of all ages. It presents 52 English sounds which are considered necessary to speak English clearly. It also brings hundreds of hours of instruction and practice (See References for bibliographical information on this software).

4. The expert was Dr. John Michael Levis. Dr. Levis studies SL pronunciation and speech intelligibility, with a focus on how SL pronunciation research affects the teaching of pronunciation. He is the founder of the annual *Pronunciation in Second Language Learning and Teaching* conference and is the founding editor of the *Journal of Second Language Pronunciation*. He is co-editor for the Phonetics & Phonology section of the *Encyclopedia for Applied Linguistics*, and two books, *Social Dynamics in Second Language Accent* and the *Handbook of English Pronunciation*. Dr. Levis has also written articles about pronunciation teaching in a variety of journals such as *Applied Linguistics*, *TESOL Quarterly*, *TESOL Journal*, *ELT Journal*.

5. A One-Way ANOVA compares existing variations within groups, of participants with existing variations between groups checking if there is a significant difference between them or not. To determine the location of the difference if there is one, it is necessary to perform a post-hoc analysis besides the One-Way ANOVA.

6. As the participants analyzed an online version of the software program *Pronunciation Power 2*, Questions 4.48 to 4.51 were ignored, totalling 67 questions analyzed instead of 72.

References

Abreu, A. C. B. (2010). *Avaliação de usabilidade em softwares educativos*. Master's thesis, Universidade Estadual do Ceará, Fortaleza, Ceará, Brasil. Retrieved July 12, 2012 from http://www.uece.br/mpcomp/index.php/ dissertacoes/47-dissertacao/129-dissertacoes-2010.

Assumpção Filho, M. M. de. (2011). *O mercado de livros mudou?* Retrieved July 12, 2012 from http://www.abramark.com.br/o-mercado-de-livros-mudou/.

Benson, P. (2007). Autonomy in language teaching and learning. *Language Teaching*, 40(1), 21-40.

Best, C. (1995). A direct realist view of cross-language speech perception. In Strange, W (Ed.), *Speech perception and linguist experience: Theoretical and methodological issues in cross-language speech research* (pp. 171-203). Timonium: York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In Bohn, O.S., & Munro, M (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13-34). Amsterdam: John Benjamins.

Bybee, J., & Cacoullos, R. T. (2008). Phonological and grammatical variation in exemplar models. *Studies in Hispanic and Lusophone Linguistics, 1(2),* 399-413. Retrieved November 14, 2015 from https://www.unm.edu/~jbybee/downloads/BybeeTorresCacoullos2008ExemplarModels.pdf.

Bongaerts, T., Summeren. C. V., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in second language acquisition*, 19, 447-465. Retrieved July 12, 2012 from http://www.ling.ohio-state.edu/~skang/371/DownloadablePapers/BongaertsEtAl97.pdf.

Borges, V. M. C. (2006) *The incorporation of communicative language teaching into the elaboration of interactive software for ESL/EFL learning*. Unpublished doctoral dissertation, University of Rhode Island, Providence.

British Council. (2012).*The English language: Frequently asked questions*. Retrieved August 10, 2012 from http://www.britishcouncil.org/organisation/publications/the-english-effect.

Burston, J. (2003). Software selection: A primer on sources and evaluation. *CALICO Journal*, *21(1),* 29-40. Retrieved May 20, 2011 from https://calico.org/journalTOC.php.

Cardoso, C., Smith, G. Garcia Fuentes, C. (2015). Evaluating text-to-speech synthesis. In Helm, F., Bradley, L. & Thouësny, S. (Eds.), *Proceedings of the 2015 EUROCALL Conference*, Italy, 7-12.

Carvalho, R. S. de. (2007). *Análise crítica do discurso publicitário na promoção de livros didáticos de língua inglesa.* Master's thesis, Universidade do Sul de Santa Catarina, Tubarão, Santa Catarina, Brasil. Retrieved July 14, 2012 from http://busca.unisul.br/pdf/90472_Richarles.pdf.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. Hong Kong: Cambridge University Press.

Chapelle, C. A., & Jamieson, J. (2008). *Tips for teaching with CALL: Practical approaches to Computer-Assisted Language Learning*. White Plains: Pearson Education.

Dale, P. (2001, October). English pronunciation: The 4th wheel. *New Routes*,15, 26-28.

Derwing, T. M., Munro, M. J., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly*, 34(3), 592-603.

Figueiredo, A. A. (2000). *Avaliação de software educativo: Geometer's sketchpad, um (micro)mundo da geometria.* Monograph, Universidade Lusíada de Lisboa, Lisboa, Portugal. Retrieved July 7, 2012 from http://users.prof2000.pt/agnelo/comunicar/avalsoft.htm.

Fino, C. N. (2003). Avaliar software "educativo". *Proceedings of the III Conferência Internacional de Tecnologia de Informação e Comunicação na Educação.* Portugal, 689-694. Retrieved April 07, 2012 from http://www.uma.pt/carlosfino/publicacoes/16.pdf.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In Strange, W (Ed.), *Speech perception and linguist experience: Theoretical and methodological issues in cross-language speech research* (pp. 233-277). Timonium: York Press.

_____. (2002). Interactions between the native and second-language phonetic systems. In Burmeister, P., Piske, T., & Rohde, A (Eds), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 217-244). Trier: Wissenschaftlicher Verlag.

_____. (2003). Assessing constraints on second-language segmental production and perception. In Meyer, A., & Schiller, N (Eds.), *Phonetics and phonology in language comprehension and production, differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

Harmer, J. (2007). *The practice of English language teaching*. Essex: Pearson Education.

Hubbard, P. (2006). Evaluating CALL software. In Ducate, L., Arnold, N. (Eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching* (pp. 313-338). San Marcos: CALICO. Retrieved May 12, 2011 from http://www.stanford.edu/~efs/calleval.pdf.

Jackson, G. B. (2000). How to evaluate educational software and websites. *TechKnowLogia*, *2(3),* 57-58. Retrieved April 17, 2012 from http://www.techknowlogia.org/TKL_active_pages2/CurrentArticles/main.asp?FileType=HTML&ArticleID=129.

Jamieson, J., Chapelle, C. A., & Preiss, S. (2005). CALL evaluation by developers, a teacher, and students. *CALICO Journal, 23(1),* 93-138.

Jenkins, J. (2000). *The phonology of English as an international language*. Hong Kong: Oxford University Press.

Kelly, G. (2000). *How to teach pronunciation*. Essex: Pearson Education Limited.

Lacerda, R. de A. (2007). *Proposta de um modelo para análise de requisitos de software educativo*. Master's thesis, Universidade de Brasília, Brasília, Distrito Federal, Brasil. Retrieved July 10, 2012 from http://repositorio.bce.unb.br/bitstream/10482/2951/1/2007_RafaeldeAlencarLacerda.pdf.

Levis, J. M. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics*, *27,* 184-202. Retrieved February 08, 2012 from http://journals.cambridge.org/action/displayIssue?iid=1820412.

Levis, J. M., Levelle, K. (2009). Rebuilding a professional space for pronunciation. *Proceedings of the 1ˢᵗ Pronunciation in Second Language Learning and Teaching Conference*, USA, 1-9. Retrieved February 08, 2012 from http://apling.public.iastate.edu/PSLLT/2009/levis_levelle.pdf.

Liakin, D., Cardoso, W., & Liakina, N. (2015). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *Computer-Assisted Language Instruction Consortium (CALICO),* 32(1), 1-25.

Mclaughlin, B. (1987). *Theories of second language learning*. London: Edward Arnold.

Navarro, E. (1999). Evaluación de materiales multimedia. *Comunicación Y Pedagogia*, *157,* 36-39.

Neri, A., Cucchiarini, C., & Strik, H. (2002). Feedback in computer-assisted pronunciation training: technology push or demand pull? *Proceedings of the ICSLP 2002.* USA, 1209-1212. Retrieved February 08, 2012 from http://lands.let.kun.nl/literature/neri.2002.2.pdf.

Oliveira, C. C., Costa, J. W., & Moreira, M. (2001). *Ambientes informatizados de aprendizagem: Produção e avaliação de software educativo*. Campinas: Papirus.

Pennington, M. (1996). *Phonology in English language teaching: An international approach*. London: Longman.

_____. (1998). The teachability of phonology in adulthood: A re-examination. *International Review of Applied Linguistics in Language Teaching*, 36(4)*,* 323-342.

_____. (1999). Computer aided pronunciation pedagogy: Promise, limitations, directions. *Computer Assisted Language Learning*, 72(5)*,* 427-440.

Pronunciation power 2 [Computer software]. (2010). Edmonton: English Computerized Learning Inc.

Soler-Urzua, F. (2012). *The acquisition of English /I/ by Spanish speakers via text-to-speech synthesizers: A quasi-experimental study*. Master's thesis, Concordia University, Montreal, Quebec, Canada. Retrieved November 14, 2015 from http://spectrum.library.concordia.ca/15159/1/Soler-Urzua_MA_F2011.pdf.

Thompson, R. I. (2011). Computer Assisted Pronunciation Training: Targeting second language vowel perception improves pronunciation. *CALICO Journal,* 28(3)*,* 744-765. Retrieved November 14, 2015 from https://www.equinoxpub.com/journals/index.php/CALICO/article/download/22985/18991.

Walker, R. (2010). *Teaching the pronunciation of English as a lingua franca*. Oxford: OUP.

Zardini, A. Sales. (2009). *Software educativo para ensino de inglês: Análise e considerações sobre seu uso*. Master's thesis, Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais, Brasil. Retrieved July 10, 2012 from http://www2.et.cefetmg.br/permalink/a0b03d88-14cd-11df-b95f-00188be4f822.pdf.

Appendix: EFL/ESL Pronunciation Teaching Software Evaluation Instrument

Title of the program: _____ **N. of CDs:** _____

1. Technical Description.

1.1 **Minimum system configuration**
1.2 **Program update**
1.3 **Technical support**
1.4 **Program installation**
1.5 **Program data storage regarding the activities done**
1.6 **Tools provided by the program**

2. Pedagogical Description.

2.7 **Syllabus**
2.8 **Target audience**
2.9 **Supplementary materials that accompany the program**
2.10 **Types of activities**
2.11 **Purpose of the activities**
2.12 **Recreational activities proposed**
2.13 **Results of the proposed activities**
2.14 **Forms of analysis of the speech recognition**
2.15 **Language teaching and learning approach of the program**
2.16 **Teacher's role**

Use the rating scale of 0-4, where 0 indicates absence of the characteristic and 4 indicates a fully satisfying exploration of the characteristic.

1. Content Design.	
The program addresses the content according to:	(Click on the circle with your mouse to fill it in)
1.1 **the proposed objectives.**	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 (barely) (completly)
1.2 **the proposed proficiency level**.	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 (barely) (completly)
The program addresses:	
1.3 **the vowel sounds.**	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 (barely) (completly)
1.4 **the diphthongs.**	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 (barely) (completly)
1.5 **the consonant sounds.**	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 (barely) (completly)

1.6 **syllabic constituents (e.g., initial consonant clusters, final consonant clusters).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.7 **word stress.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.8 **stress in compound words.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.9 **stress in words with suffixes.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.10 **sentence stress.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.11 **prominence (e.g., contrastive stress, emphatic stress).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.12 **connected speech phenomena (e.g., linking, reductions, assimilation).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
1.13 **intonation.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2. Pedagogical Design.	
The program contrasts:	
2.14 **different vowel sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2.15 **different consonant sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
The program presents and distinguishes:	
2.16 **grammatical intonation patterns (e.g., Yes/No questions, Wh- questions, statements, complex sentences, tag questions).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2.17 **semantic intonation patterns (for emphasis and contrast).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
The program uses:	
2.18 **phonetic symbols to present and practice vowel and consonant sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2.19 **phonetic transcriptions to present and practice rhythm, stress and intonation.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)

2.20 **listening comprehension activities.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
The program presents:	
2.21 **the contents organized sequentially.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
The ACTIVITIES work on:	
2.22 **the perception of sounds.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.23 **the production of sounds.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.24 **the perception/production of sounds in a contextualized way.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.25 **the distinction of sounds.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.26 **the perception of rhythm, stress and intonation.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.27 **the production of rhythm, stress and intonation.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.28 **the perception/production of rhythm, stress and intonation in a contextualized way.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.29 **the distinction of intonation patterns.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
Regarding the ACTIVITIES:	
2.30 **the instructions are clear and objective.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.31 **the audio scripts are available.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.32 **the activities are varied.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
2.33 **the order in which the activities are presented is flexible.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)
Regarding the PROGRAM:	
2.34 **a "help" option is given on the screen.**	○ 0 ○ 1 ○ 2 ○ 3 ○ 4 (barely) (completly)

2.35 **it is easy to go in and out of a section.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2.36 **it allows the user to restart from where he left off.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
2.37 **the information it gives is complete.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3. Assessment / Flexibility Design.	
Regarding the feedback of the ACTIVITIES:	
3.38 **it gives explanations about the user's errors.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.39 **it asks the user to remake the production.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.40 **it remakes the user's production.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.41 **there are comments on the right and wrong answers.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.42 **the errors are signaled so that the user may have the option to redo the activity.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
Regarding the ACTIVITIES:	
3.43 **the user can choose their difficulty level.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.44 **the user can choose their content.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.45 **they can be redone by the user.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.46 **the results can be printed.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
3.47 **they allow the user to choose between a male or female voice.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4. Multimedia Design.	
Regarding the program: ***NOTE: If you are using an online version of the program, ignore Questions 4.48 to 4.51.**	

*4.48 **the installation instructions are clear.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
*4.49 **it is easy to install.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
*4.50 **it starts fast.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
*4.51 **it runs without freezing.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.52 **it uses animation to demonstrate the production of sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.53 **the animations are appropriate to what they want to show.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.54 **it uses videos to show the lip movements when producing sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.55 **the video images are clear.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.56 **it uses visual acoustic displays (e.g., waveforms, spectrograms, pitch tracings) to provide immediate feedback on the user's speech recordings.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.57 **it contrasts the user's mother tongue sounds to the English sounds.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.58 **it incorporates highly variable input (e.g., regional phonological varieties, male and female voices).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.59 **the sound is clear (free of noise).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.60 **it uses voice recording activities.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.61 **it gives explicit feedback of the activities by signaling the user's errors.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.62 **the screen layout is clean (icons, colors and images do not cause visual pollution).**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)
4.63 **it is easy to navigate.**	$\bigcirc_0$ $\bigcirc_1$ $\bigcirc_2$ $\bigcirc_3$ $\bigcirc_4$ (barely) (completly)

4.64 **it is easy to access the menu.**	0 1 2 3 4 (barely)　　(completly)
4.65 **the icons are easily comprehensible.**	0 1 2 3 4 (barely)　　(completly)
4.66 **the system configuration and required accessories (microphone, earphones) are easily found in the market.**	0 1 2 3 4 (barely)　　(completly)
5. Automatic Speech Recognition Design.	
The program uses:	
5.67 **Automatic Speech Recognition to provide immediate feedback on the user's pronunciation.**	0 1 2 3 4 (barely)　　(completly)
The program allows the user to calibrate the Automatic Speech Recognition according to:	
5.68 **his/her nationality.**	0 1 2 3 4 (barely)　　(completly)
5.69 **his/her region.**	0 1 2 3 4 (barely)　　(completly)
5.70 **his/her gender.**	0 1 2 3 4 (barely)　　(completly)
The Automatic Speech Recognition:	
5.71 **does not consider disfluencies/hesitations.**	0 1 2 3 4 (barely)　　(completly)
5.72 **is not sensitive to external noise.**	0 1 2 3 4 (barely)　　(completly)