WorldCALL

International Perspectives on Computer-Assisted Language Learning

Edited by Mike Levy, Françoise Blin, Claire Bradin Siskin and Osamu Takeuchi

ROUTLEDGE

WorldCALL

International Perspectives on
Computer-Assisted Language Learning

Edited by Mike Levy, Françoise Blin, Claire Bradin Siskin and Osamu Takeuchi

Routledge
Taylor & Francis Group
New York London

First published 2011
by Routledge
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

This edition published in the Taylor & Francis e-Library, 2011.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.

Library of Congress Cataloging in Publication Data

WorldCALL: international perspectives on computer-assisted language learning / edited by Mike Levy...[et al.].
p. cm.—(Routledge studies in computer assisted language learning; 5)
Includes bibliographical references and index.
1. Language and languages—Computer-assisted instruction. 2. Language and languages—Globalization. I. Levy, Mike, 1953–
P53.85.W68 2011
418.0078'5—dc22
2010031672

ISBN 0-203-83176-4 Master e-book ISBN

ISBN13: 978-0-415-88086-2 (hbk)
ISBN13: 978-0-203-83176-2 (ebk)

Contents

PART II: Developing Language Skills through Technology

PART III: Materials Design and Development

6 Improving Pronunciation via Accent Reduction and Text-to-Speech Software

Ferit Kılıçkaya

INTRODUCTION

There is a common belief that proper pronunciation is an indispensable part of successful communication, which reflects an important current direction of teaching foreign languages in schools. The ultimate goal of language learning in most cases is communication, both written and oral; however, many language teachers often neglect or are forced to neglect pronunciation throughout the entire teaching process. One of the reasons for this is certainly lack of time, resulting from a focus on grammar and lexical aspects of teaching English more than on its pronunciation. It may also be the case that less experienced teachers, who show a great interest in this subject, prioritize their aims in such a way that grammar and vocabulary remain at the top of their agenda. As a result, although pronunciation receives interest both on the part of the teacher and the learner, it is often neglected due to the stated reasons. Furthermore, there is no denying that teachers often lack proper preparation for teaching pronunciation and do not take any measures to receive extra training in this field.

Pronunciation is important not only to communicate ideas easily but also to understand other speakers, given that listening comprehension and pronunciation are interdependent, as contended by Gilbert (1995): Learners of a language can be "cut off" from language if they cannot understand what is being said and cannot be understood by others speaking or learning that language. It is particularly important to integrate pronunciation into beginner classes as it will, from the very beginning, help avoid the risks of fossilization and stabilization of pronunciation habits (Ritchie & Bhatia, 2008). This integration can be achieved through various activities such as drills, listen and imitate, and computer-assisted pronunciation teaching (Kılıçkaya, 2006; Levis, 2007; Seferoğlu, 2005).

The goal of teaching pronunciation is not to make learners sound like native speakers of English, as only few highly gifted and motivated individuals can achieve this (Jenkins, 2005). A more realistic approach is to enable the learners to pronounce the language so that s/he can be understood (Celce-Murcia, Brinton, & Goodwin, 1996). In order to increase this

ability, exposure to the target language is vital as learners acquire language mainly from the input they receive. A significant amount of exposure to the target language might help learners to practice the pronunciation and increase their ability to comprehend and express ideas (Tench, 1981). Consequently, the significance of the teacher's role is commonly acknowledged in maximizing their students' exposure to the target language as much as possible. In order to achieve this aim, most language teachers try to use mainly L2 during their lessons to promote the development of the communicative abilities needed to deal with real-life contexts requiring the target language or L2. Moreover, most try to take advantage of technology and provide their students with listening texts in the classroom via CD player, radio, TV, computer, and the Internet, as well as trying to encourage them to listen to English outside the classroom. They recommend English films to watch or even provide students with the DVDs, encourage them to visit particular EFL/ESL websites, or podcast the materials from the websites on their own and deliver them to students.

A BRIEF REVIEW OF TECHNOLOGICAL DEVICES USED IN PRONUNCIATION ACTIVITIES

Looking back to the past, the tape recorder (latterly the audiocassette player) was the first foreign language classroom device through which a classroom might be called technology assisted. As claimed by Kiely (2005), it "brought sound to the classroom for nearly half a century" and was particularly favored during the era when the audio-lingual method was popular around the world. Audiocassettes were mainly used to repeat words or sentences provided in coursebook materials and the main aim was to imitate the way these words or sentences were spoken. Today, audiocassette players have been increasingly replaced by CD/DVD/mp3 players due to their sound quality and their having the functions that traditional audiocassette players lack, such as easily finding the recording or going backward to locate previous tracks.

Another very popular and useful device that provides students not only with the audio input but also with the visual support is a videocassette player. In pronunciation classes, teachers used or still use videocassette players to present movement of articulatory organs while pronouncing particular sounds. In such presentations, using the materials available on the videocassettes, students can be "guided to relax their vocal apparatus, practicing the tensing and flexing of various speech organs" (Celce-Murcia et al., 1996, p. 314). Unfortunately, few such programs are available and video is used mostly as a feedback tool, focusing either only on sounds, or on speech organs' movement. Additionally, the teacher may slow down the pace of speech delivery or freeze the picture to make students focus on the speech organs when pronouncing a particular sound or word.

Following the advances in the technology, videocassettes have been replaced by CD/DVD videos, due to the high picture quality and the functions enabling finding and/or viewing part of the video easily. In addition to this, accent reduction software such as Pronunciation Power has come into the market with full functions and features that videocassettes do not have. Although each piece of software has different features and functions, most provide the presentation of each sound through auditory and visual activities. They aim to identify the deviations of a learner's current speech, such as pronunciation, and to change the way s/he uses her/his mouth to produce the sounds together with intonation and stress. Pronunciation Power I, marketed by the Canadian company English Computerized Learning Inc. (ECL) and developed by Blackstone Multimedia Corporation, focuses on the practice of the 52 English sounds and contains S.T.A.I.R. (Stress, Timing, Articulation, Intonation, and Rhythm). Moreover, it provides real-time visual illustrations of articulatory movements (both a side and a front view) for the production of each sound to accompany the recordings of these sounds' pronunciation. The front view is presented by a video clip of a jaw, lip, and tongue movement of a real person. The side view uses animated drawings providing an X-rayed image of the complete articulatory mechanics (including manner and location of airflow, placement and movement of lips and tongue, velum movement, etc.). A graphic representation of the sound utterance as a waveform is also available. Furthermore, a phonetic transcription of the sounds that were either read on one's own or listened to is included. The types of activities include identifying the target sound in minimal pairs with text and audio files, listening and recording the difficult word, and identifying a word through audio and spelling as well as in sentences with examples of the target sound. Considering these features and the types of activities provided in Pronunciation Power and similar products, language learners have been provided with activities, which especially empower them with both audio and video representations.

However, the main problem with these tools is that they only deal with some fixed sounds or words in isolation and do not provide opportunities for practicing the words in sample sentences or context. In order to overcome the problem of pronunciation of a limited number of words, CD/DVD dictionaries, such as the *Longman Dictionary of Contemporary English* (LDOCE5—http://www.longman.com/ldoce) have come into use, providing fast access to words in hyperlinked word entries with advanced searching capabilities and a wide coverage of words together with their pronunciations. Moreover, Internet-based dictionaries come into the picture also to provide definitions and example sentences with sound files for pronunciation, such as *Cambridge Dictionaries Online* (http://dictionary.cambridge.org/) and other meta-dictionaries giving access to multiple dictionaries and thesauri. In spite of these resources, the problem of using the words in sentences still prevails and teachers have referred to computers with Internet access to provide their students with real uses. The Internet is a great source

of materials for meeting the challenge of leading pronunciation classes, presenting language, and benefiting from a source of teaching, listening, and pronunciation materials. The freely available websites such as Voice of America (http://wwww.voanews.com), Randall's ESL Cyber Listening Lab (http://www.esl-lab.com), English Listening Online (http://www.elllo.org), and Breaking News English (http://www. breakingnewsenglish.com) have been invaluable websites from which language teachers can download many audio files created for the purpose of teaching EFL. The huge amount of streaming audio and video materials on the Internet, CD/DVD/online dictionaries, and software for pronunciation has promoted pronunciation. However, as most EFL listening materials cover a limited range of topics and the level of materials is generally targeted towards learners with higher levels of English, selecting materials for the learners' level and interest has become another issue for teachers. Recently, text-to-speech technology, or speech synthesis, has drawn attention from some researchers such as Azuma (2008), Kataoka, Funakoshi, and Kitamura (2007) and Kılıçkaya (2006), as it can read and convert any document provided whether it comes from a website, a newspaper article, or a book.

Text-to-Speech Technology

Text-to-speech technology or speech synthesis is the conversion of text to speech through special computer applications called Text To Speech software (TTS). This technology works with voices, digitally created and trained to read any text. Each sentence is individually generated instead of being played from a previously recorded sentence, text, or document. TTS systems are primarily developed to address the needs of companies and organizations and appear to be invaluable tools for the visually impaired, enabling them to read anything on the computer, such as newspapers. However, in recent years, TTS has been widely used in audiobooks, dictionaries, and computer voice interfaces and it has been given much attention from researchers (see, for example, Azuma, 2008; Kataoka et al., 2007; and Kılıçkaya, 2006). Studies of TTS as a tool for foreign language education have generally focused on the suitability of this technology in language classrooms (Azuma, 2008), and whether students noticed that the voices produced are artificial (Kataoka et al., 2007). According to Azuma (2008), TTS speech has enough quality to be used as audio materials especially for listening for Japanese EFL learners. Kataoka et al. (2007) stressed that the students in their studies did not notice that the dialogues or the sentences were produced by TTS speech. Furthermore, in a study conducted by Kataoka (2007, December), high school students in Japan better memorized the words and phrases with the voices produced by text-to-speech technology, and learned and memorized more English words, than without using this technology. Although there is little research on TTS technology, preliminary results show that it has an important role, especially in listening and pronunciation in language classrooms.

This role has been put forward by Kılıçkaya (2006), who lists the advantages of text to-speech technology as follows:

- You can listen to any text and any topic (most EFL listening materials cover a limited range of topics and some of them are rather expensive).
- You can adjust the speed of reading according to your own needs.
- You can create audio versions from any text (wav or mp3 files).
- You can create pronunciation exercises for yourself (a single word can also be read).
- You can create mini-dialogues (changing speakers at run times is possible).

However, as nothing is without pitfalls, Sha (2009), regarding the disadvantage of TTS, argues that "[a] vital disadvantage of TTS is its limitations in naturalness, pleasantness and expressiveness (TTS voices can never be used in automated dialogue replacement in film making, at least up to now) though intelligibility seems a lesser problem" (p. 640).

In summary, the pronunciation skill in any target language is a crucial part of communication and language learners should be provided with pronunciation practice that is well incorporated in the classroom materials or activities. To help learners achieve intelligible pronunciation for better understanding and for being understood by native speakers, learners, or speakers of the target language, tools provided by accent reduction software and TTS in addition to CD/DVD materials can be of significant help by allowing learners to study at their own pace and exploit visual elements. The purpose of this study is to find out whether integrating accent reduction and text-to-speech software in elementary language classes would result in improvements in learners' pronunciation and to explore their opinions on accent reduction and TTS technology. Hence, this study tries to answer the following research questions:

1. Do accent reduction and text-to-speech technology result in better improvements in learners' pronunciation considering the scores on the tests given?
2. What are the learners' opinions regarding the use of accent reduction and text-to-speech technology?

METHOD

The Turkish Context for English as a Foreign Language (EFL)

Turkish students, though there might be some exceptions in some schools or institutions, mostly follow an English curriculum, in which pronunciation is not regarded as an integral part of language teaching and learning.

Language exams conducted in Turkey lack the assessment of oral skills and focus mainly on grammar, vocabulary, and reading. In other words, the receptive skills are preferred over the productive skills. Therefore, teachers do not regard pronunciation as an integral part of language teaching. The result is that Turkish students have problems with pronunciation, especially with plural and past tense ending, and differentiating between final /σ/ and /ç/ both in perception and production. However, language acquisition does not depend on the teacher and the teaching process only. Kenworthy (1990) argues that there are also other factors that affect pronunciation acquisition such as the native language—the problem of L1 interference and divergence of L1 and L2, and the amount of exposure—how much the learner is "surrounded" by English, and motivation and concern for good pronunciation. In Turkish, the relationship between the phonemes and the letters are one-to-one and each phoneme consists of just one letter, whereas each English phoneme can be represented in many different ways. Therefore, it can be said that the difficulties of Turkish speakers, resulting from L1 interference, can also be attributed to the irregular spelling in English. When we add the lack of practice in pronunciation in the curriculum, pronunciation becomes a real problem.

Participants

The participants of this study were 35 Turkish EFL students enrolled in a General English course. Students from three elementary classes at a private language institution in Turkey, aged between 22 and 28 with an average age of 23.5, participated in the study. Of the 35 participants, 25 were female and 10 were male. All were freshman students at various universities in Ankara, the capital of Turkey. They were graduates of high schools where English is a compulsory subject. Age and gender were not taken into consideration. The participants were chosen using purposive sampling; it did not include random assignment as the availability of participants was limited. However, before the course began, the participants took a placement test to specify their level of English and participants having similar scores were assigned to three groups. They were also given a pre-test to ensure that there were no statistically significant differences between the groups. The sample consisted of 10 students in the control group, 13 students in experimental group one, and 12 students in experimental group two.

Research Design and Instruments

During the course, the Total English Elementary Student's Book and DVD Pack was used as the main coursebook, which is based on the objectives of the Common European Framework of Reference (CEFR) and includes teacher support material, DVD with authentic clips from film and TV, and a "catch up" CD-ROM with practice for students who miss lessons. DVD and

CD-ROM materials were used in class after classroom lectures and activities, using a laptop, data projector, and speakers. The pre/post-test method was used for the study. Pronunciation questions were developed from the book *Tree or Three?* by Baker (2006), a pronunciation course for beginner and elementary students of English that provides practice in the pronunciation of English sounds, word stress, and intonation through a variety of interesting exercises and activities. Pre- and post-tests consisted of three sections: The first section consisted of 20 most frequently mispronounced words, the second consisted of 10 affirmative and declarative sentences including these words, and the last section presented a picture that portrayed a picnic where some children are busy with activities such as playing with a football and eating. Answers for both the pre- and post-tests were rated according to a 5-point Likert scale for responses. Before administering the pre- and post-tests, three instructors in the ELT department were asked to comment on the items presented in the test in terms of clarity and content. These tests were later given to 30 randomly selected students who were in the preparatory school of Middle East Technical University. The questions were then analyzed taking the Kuder-Richardson scale into consideration. According to the results, *the reliability level* was .89 on the Kuder-Richardson scale, showing that these tests could be used as standardized tests. In addition to the quantitative analysis, semi-structured interviews were conducted with the participants in experimental group two, who were asked to provide feedback on their use of text-to-speech technology in pronunciation activities. The participants were interviewed one by one. The interviews were recorded, with the learners' consent, and the researcher took notes. The participants were interviewed in their mother tongue, Turkish. The main questions were related to whether the students enjoyed the activities undertaken during their class hours and elaborated on the reasons from their own perspectives.

Procedure

On the first day of class, an informed consent form was signed by the students agreeing to participate in the study. After students signed the form, the instructors administered the pre-test. One class (control group) followed the traditional instruction (using a CD player and a pronunciation textbook, *Tree or Three?*, by Baker, 2006); another class (experimental group one) followed the traditional instruction that integrated the use of accent reduction software (Pronunciation Power I); and the final class (experimental group two) followed the traditional instruction that integrated the use of accent reduction and text-to-speech software (Text Aloud mp3 with NeoSpeech voices—Paul and Kate). The study lasted for 16 weeks and the instructor met the groups for three hours each week. With the results obtained from the pre-test, and by means of a one-way ANOVA test, it was possible to establish whether or not there were significant differences between two groups of participants at the 0.05 alpha levels (see Table 6.1).

Table 6.1 Pre-Test Results

Group	*N*	*Mean*	*Std. Deviation*	*Std. Error*
control	10	5.60	.516	.163
experimental 1	15	5.73	.458	.118
experimental 2	10	5.60	.516	.163
Total	35	5.66	.482	.081

	Sum of Squares	*df*	*Mean Square*	F	*Sig.*
Between Groups	.152	2	.076	.315	.732
Within Groups	7.733	32	.242		
Total	7.886	34			

As can be seen, the significance level was higher than 0.05, $F(2,32) = .315$, which leads to the conclusion that there were no significant differences between the groups. Once this point became clear, the study was carried out with these three groups. On the last day of class, the instructor administered the post-test to all groups. In addition, experimental group two was interviewed regarding their views on accent reduction and text-to-speech software. The scores obtained in the pre- and post-tests were analyzed to see whether there was a statistically significant difference among the groups.

RESULTS

The post-test scores obtained by experimental and control groups were analyzed using the SPSS software package and the one-way ANOVA test to establish whether there were significant differences among the three groups of participants at the 0.05 alpha levels (see Table 6.2).

Table 6.2 shows that the significance level was higher than 0.05, $F(2,32) = 22.156$, which leads to the conclusion that there were significant differences between the groups. The analysis used to address the first research question (Do accent reduction and text-to-speech technology result in better improvements in learners' pronunciation considering the scores on the tests given?) revealed significant practice effects on listening comprehension.

Considering the scores in the post-test, experimental group two, which was exposed to accent reduction and text-to-speech software, did better than the other groups. However, the study revealed that there were no statistically significant differences in the pronunciation of single words as all the groups did equally well. Nevertheless, there were statistically significant differences between the groups in the pronunciation of sentences, with experimental group two doing significantly better than the other groups.

Table 6.2 Post-Test Results

Group	*N*	*Mean*	*Std. Deviation*	*Std. Error*
control	10	32.50	4.859	1.537
experimental 1	15	70.33	6.651	1.717
experimental 2	10	78.10	2.283	.722
Total	35	70.31	7.851	1.327

	Sum of Squares	*df*	*Mean Square*	F	*Sig.*
Between Groups	1216.810	2	608.405	22.156	.000
Within Groups	878.733	32	27.460		
Total	2095.543	34			

During the semi-structured interview session, in order to answer the second research question (What are the learners' opinions regarding the use of accent reduction and text-to-speech technology?) the participants in experimental group two provided their opinions on accent reduction and the text-to-speech software. Their responses have been categorized as follows:

- **Addition of visual support.** The participants highly valued the visual activities provided by Pronunciation Power and the real-time visual illustrations of articulatory movements.
- **No stress or fear.** The participants enjoyed "playing" with the activities during sheltered practice sessions in which the participant could take risks without stress or fear of error.
- **Self-pacing.** The participants had the opportunity to study at their own pace, without having to try to keep up with the teacher or the classmates who were fast learners.
- **Immediate feedback.** The participants were provided with immediate feedback by the pronunciation software and the text-to-speech software.
- **Pronunciation of any word or sentence.** Text-to-speech software provided pronunciation of any words or any utterance entered by the participants. This provided a quick and efficient access without searching the words in their paper-based or CD/DVD dictionaries or any other resource.
- **Improved spelling and listening.** The participants also stated that they practiced spelling and writing while creating sentences and listening to these sentences created by TTS technology.
- **Lack of emotion and naturalness.** The participants, while listening to the longer sentences created by the TTS software, had the feeling that

the speaker lacked emotion or naturalness though his/her pronunciation was clear and understandable.

DISCUSSION

The integration of accent reduction and text-to-speech software into classrooms can help learners of English improve their pronunciation due to factors such as practice sessions in which the learner can take risks without stress and fear of error and can receive immediate feedback. However, this does not mean that only TTS technology and/or accent reduction software resulted in the participants' improved results in the tests reported in the study. By adding a variety of resources and enriching the classroom activities with audio and video materials, the classroom environment provided "optimal" conditions for learning, which can be noticed in the responses provided by the participants to the interview questions. The classroom atmosphere, teachers' guidance, and the somewhat relaxed state of mind facilitated the retention of material to its maximum potential. Thus, TTS technology and supplementary activities using appropriate software under the teachers' guidance seemed to be an invaluable resource not only in classroom activities but also beyond the classroom, exposing the learners to the language as much as possible. Although TTS technology is not designed for language teaching, it can be used for different purposes, as in the current study, in an EFL class, including:

1. **Practicing mispronounced words.** The teacher can note down the words generally mispronounced by the students and these words, with the help of TTS technology, can be listened to. This can be done by the students and/or the teacher. Instead of directly correcting the mistakes in the class, as it is probably more effective for students to correct their own mistakes, students can study on their own or in groups. Moreover, these words can be posted on blogs or wikis for later use.
2. **Listening to authentic materials/short stories/articles.** The Internet is an invaluable source for finding short stories, articles, authentic materials, and the materials designed for pedagogic purposes. Reading these materials can be accompanied by a listening activity, which will lead to longer retention of vocabulary and pronunciation of the words.
3. **Creating dialogues.** Using TTS, learners can create dialogues with different voices and accents, using their imagination. This will, hopefully, help them practice both writing and listening, and pronunciation.

Regarding the limitations of the study, it was carried out for two hours each week for 16 weeks, and with only a small number of participants due

to the time constraint and the availability of the participants. It is suggested that similar experiments with a large number of participants through random assignment should be replicated.

CONCLUSION

In the past, sound and video materials were difficult to create and put into use in educational settings. However, with the fast development of technology, software, and equipment, it is now much easier to manage new audio and video materials because they are digital and offered in high quality. The quicker access to materials, and the use of these materials as learning and teaching objects, has made it possible to create a new learning and teaching experience in educational settings, especially in language classrooms. Considering the results in this study, the integration of accent reduction and text-to-speech software into classrooms has paved the way for helping learners of English improve their pronunciation, offering practice sessions in which the learner can take risks without stress and fear of error and providing the advantage of immediate feedback. TTS technology together with pronunciation practice software such as Pronunciation Power, with their potential for providing a valuable resource of pronunciation and enriching classroom activities, can greatly contribute to the field of language teaching and learning not only in the Turkish context but in others where the learner can go beyond the materials provided by course books. The combination of a human teacher together with opportunities for human–human interaction in the classroom or outside the classroom via text, and face-to-face or online communication supported by the technology, constitute the most desirable solution for both educators and students.

ACKNOWLEDGMENTS

I would like to gratefully and sincerely thank Dr. Françoise Blin for all her invaluable comments and suggestions.

REFERENCES

Azuma, J. (2008). Applying TTS technology to foreign language teaching. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced language acquisition and learning* (pp. 497–506). New York: Information Science Reference.

Baker, A. (2006). *Tree or three? An elementary pronunciation course* (2nd ed.). Cambridge: Cambridge University Press.

Celce-Murcia, M., Brinton, D., & Goodwin, J. (1996). *Teaching pronunciation*. Cambridge: Cambridge University Press.

Gilbert, J. (1995). Pronunciation practices as an aid to listening comprehension. In D. J. Mendelson & J. Rubin (Eds.), *A guide for the teaching of second language learning* (pp. 97–111). San Diego: Dominic Press.

Jenkins, J. (2005). Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *TESOL Quarterly, 39*(3), 535–543.

Kataoka, H. (2007, December). *Preparation for university entrance examinations: How to learn frequently-appearing English words with the TTS audios*. Paper presented at the Multimedia & Internet Seminar, Japan Association for Language Education and Technology (LET), Kansai Chapter, Osaka, Japan.

Kataoka, H., Funakoshi, Y., & Kitamura, Y. (2007). Utilizing text-to-speech synthesis for English education. In *Proceedings of the 23rd Annual Conference of Japan Society for Educational Technology* (pp. 429–430). Osaka: JSET.

Kenworthy, J. (1990). *Teaching English pronunciation*. New York: Longman.

Kiely, R. (2005). The role of television and televisual literacy in language teaching and learning. Retrieved from http://www.developingteachers.com/articles_tchtraining/tvpf_richard.htm

Kılıçkaya, F. (2006). Text-to-speech technology: What does it offer to foreign language learners? *CALL-EJ Online, 7*(2).

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*(1), 184–202.

Ritchie, W. C., & Bhatia, T. K. (2008). Psycholinguistics. In B. Spolsky & F. M. Hult (Eds.), *The handbook of educational linguistics* (pp. 38–52). Singapore: Blackwell Publishing.

Seferoğlu, G. (2005). Improving students' pronunciation through accent reduction software. *British Journal of Educational Technology, 36*(2), 303–316.

Sha, G. (2009). Using TTS voices to develop audio materials for listening comprehension: A digital approach. *British Journal of Educational Technology, 41*(4), 632–641.

Tench, P. (1981). *Pronunciation skills*. London and Basingstoke: Macmillan.