Can TTS help L2 learners develop their phonological awareness?

Almir Anacleto de Araújo Gomes[1], Walcir Cardoso[2], and Rubens Marques de Lucena[3]

Abstract. Text-To-Speech synthesizers (TTS) have raised the interest of researchers and teachers for their ability to enhance foreign/second language (L2) learning, particularly with regards to the development of pronunciation skills (Liakin, Cardoso, & Liakina, 2017). Despite some optimistic results, there are no studies that investigate TTS's pedagogical potential to enhance L2 Phonological Awareness (PA), especially in *foreign* language contexts, where access to rich aural input is limited in terms of both quantity and quality. The present study examines TTS's pedagogical potential as a tool to assist English L2 learners develop their PA, focusing on the morphophonological alternations that characterize regular past tense marking in English (past -ed). Results show that TTS contributed positively for the auditory perception and controlled (but not spontaneous) production of the targeted phenomenon.

Keywords: Text-to-speech synthesizers, phonological awareness, L2 pronunciation.

1. Introduction

Research has shown that the pedagogical use of TTS has the potential to contribute to learning in the acquisition of L2 vocabulary and pronunciation (e.g. Liakin et al., 2017; Soler-Urzúa, 2011). Recently, there have been studies that evaluate the voice quality of TTS systems, corroborating previous hypotheses that current synthesized voices are ready for use in L2 education: they are not only perceived as appropriate

1. Universidade Federal da Paraíba, João Pessoa, Brazil; almir@ufcg.edu.br
2. Concordia University, Montreal, Canada; walcir.cardoso@concordia.ca
3. Universidade Federal da Paraíba, João Pessoa, Brazil; rubenslucena@yahoo.com

How to cite this article: De Araújo Gomes, A. A., Cardoso, W., & De Lucena, R. M. (2018). Can TTS help L2 learners develop their phonological awareness? In P. Taalas, J. Jalkanen, L. Bradley & S. Thouësny (Eds), *Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018* (pp. 29-34). Research-publishing.net. https://doi.org/10.14705/rpnet.2018.26.808

(or 'good enough') from a sound perspective (Bione, Grimshaw, & Cardoso, 2017; Cardoso, Smith, & Garcia Fuentes, 2015), but they are also more likely to enhance opaque features of the target language (e.g. regular past tense marking in English is often detected with higher accuracy in synthesized speech; Bione et al., 2017; John & Cardoso, 2017). One area that has not received consideration by researchers is TTS' ability to raise L2 learners' PA.

As such, this pilot study attempts to contribute to the literature on CALL-assisted PA. It focused on TTS's ability to provide the type of input necessary to develop learners' ability to become aware that past tense marking (past -ed) in English is characterized by three allomorphs: /t/ in the presence of a preceding voiceless segment (e.g. talk[t]), /Id/ when the targeted morpheme is preceded by homorganic /t/ or /d/ (e.g. add[Id]), and /d/ elsewhere (e.g. play[d]). To assess participants' PA development, we included tasks that evaluated their pre-/post-test awareness of past -ed pronunciation, their auditory perception of the phenomenon, and their controlled/spontaneous oral production. The general question that guided this research was: can English as a Foreign Language (EFL) learners benefit from exposure to synthesized voices for developing their awareness to past -ed allomorphy?

2. Method

Nine Brazilian Portuguese (BP) native speakers participated in this study. They were EFL students living in João Pessoa (Brazil, with ages ranging from 25 to 35. They had all completed (or were attending) secondary school.

The study consisted of two hour one-shot sessions with the participants. For the treatment, participants were asked to complete a set of learning activities using a popular NeoSpeech synthesized voice, *Julie* (via the NaturalReader software). They consisted of activities that encouraged them to listen and respond to question such as *listen and decide whether the action took place in the present or past*.

For data collection, participants were asked to complete a series of tasks designed to evaluate their PA before and after the use of TTS during the treatment phase. To examine learners' PA, participants completed two questionnaires (pre- and post-tests), in which they were asked to evaluate ten propositions about English pronunciation using an eight point Likert scale, according to Alves's (2012) five levels of PA: L2 phonemes, L2 allophones, identification of non-distinctive sounds in L1 and distinctive sounds in L2, syllabic patterns, and rhymes.

Based on Cardoso et al. (2015) and Bione et al. (2017), participants also performed an auditory identification task, in which they heard 16 sentences, 12 of them in the past (four distractors), but without lexical elements that could identify their tense.

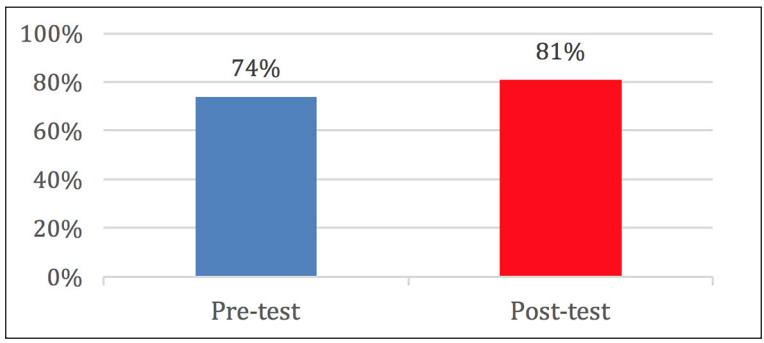
To evaluate the controlled production of past -ed, participants read a list of 26 words aloud (14 were regular past verbs, the others were distractors). To evaluate less controlled (*spontaneous* henceforth) production, participants were asked to answer questions about the past holidays of fictional characters, using a set of regular past verbs. The key difference between these tasks was the amount of cognitive effort to complete them: while the first relied on the participants' ability to produce ed-final words in isolation, the spontaneous task required them to focus on content, allowing us to obtain more natural speech (Gomes, 2015; Ortiz & Ivette, 2005; see Labov, 1972 for a sociolinguistic perspective).

3. Results

3.1. Auditory perception

Due to the limited scope of the study and the small number of participants, we report the results using descriptive statistics, in percentages. As Figure 1 indicates, there was a 7% improvement in past tense perception from pre-test to post-test.

Figure 1. Auditory identification of past -ed: pre-test and post-test

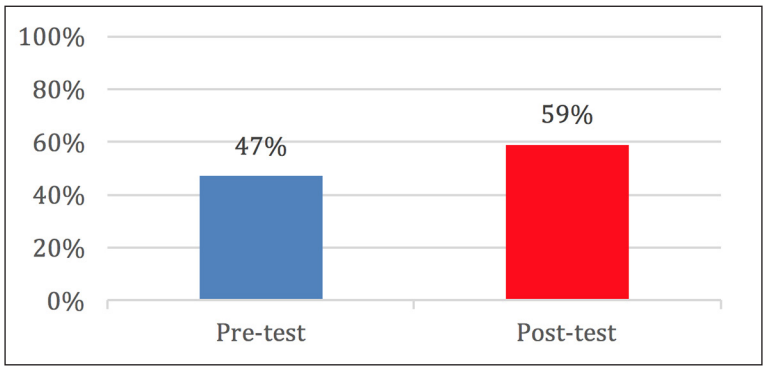


These results suggest that access to synthesized L2 input may contribute to a better perception of sentential verbal tense, as well as the identification of regular past -ed in English.

3.2. Production

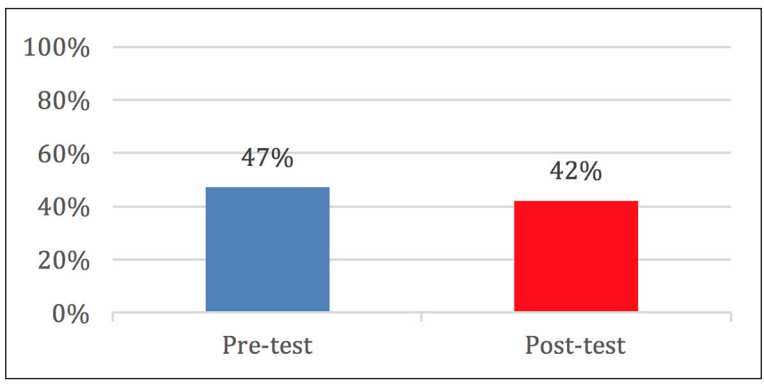
Comparing pre-test to post-test results in controlled production, we observed an overall improvement of 12% in past -ed pronunciation, as shown in Figure 2.

Figure 2. Controlled production of past -ed: pre-test and post-test



Interestingly, data from spontaneous production yielded a different pattern, showing a reduction in past -ed marking on the post-test, as illustrated in Figure 3.

Figure 3. Spontaneous production of past -ed: pre-test and post-test

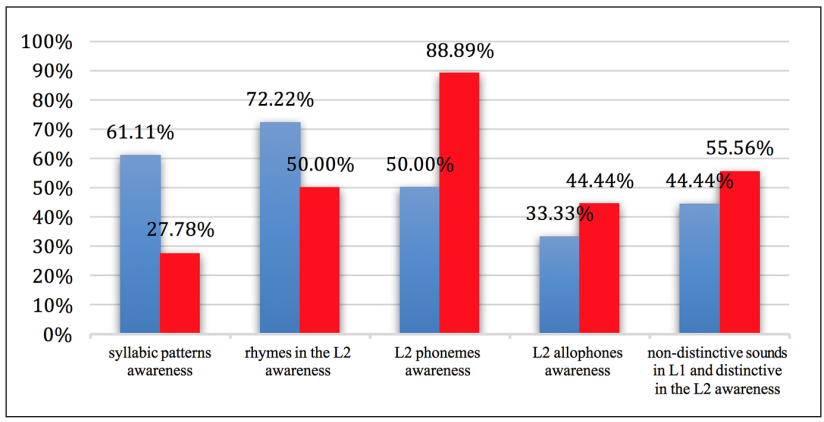


In sum, our findings for oral production yielded mixed results regarding the pedagogical effects of TTS: While an improvement in past -ed pronunciation was observed in controlled production, the reverse was found in spontaneous tasks.

3.3. PA test

Figure 4 shows an improvement in past -ed PA in three of the five levels of phonological awareness considered: L2 phonemes (from 50% to 88.89%), L2 allophones (from 33.33% to 44.44 %), and identification of non-distinctive sounds in L1 (allophones) and distinctive in L2 (from 44.44% to 55.56%). Regarding the other two levels (syllabic patterns and rhymes), PA decreased from 61.11% to 27.78% and from 72.22% to 50.00%, respectively.

Figure 4. PA tests: pre-test and post-test



Overall, considering the five levels of PA adopted, the participants' phonological awareness only increased by a mere 1.11% on the post-test (from 52.22% to 53.33%).

4. Discussions and conclusions

This study aimed to answer whether EFL learners could benefit from exposure to synthesized voices for developing their awareness to past -ed allomorphy. Our findings suggest that the pedagogical use of TTS can be effective in at least two of the measures encompassed by our definition of PA: auditory -ed perception and its controlled production. This can be explained by the fact that these two tasks required a focus on language (e.g. in comparison with spontaneous production). These findings suggest that, considering these two measures, TTS constitutes a good source of aural input (as reported Cardoso et al., 2015 and Bione et al., 2017), and that it may be used to enhance the opaque alternations found in past -ed marking in English (as hypothesized, based on Bione et al., 2017, and John &

Cardoso, 2017). We acknowledge some of the limitations of our study, particularly the limited number of participants and lack of statistical rigour in our analysis. In future research, we would like to mitigate these limitations and verify if, given a larger sample, there will be a more significant improvement at post-test.

5. Acknowledgements

We would like to thank the participants for their invaluable contributions. We would also like to acknowledge the financial support from CAPES (Brazilian Federal Agency for the Support and Evaluation of Graduate Education, Ministry of Education of Brazil). PDSE - 88881.131607/2016-01.

References

Alves, U. K. (2012). O que é consciência fonológica. In R. Lamprecht, A. P. Blanco-Dutra et al. (Eds), *Consciência dos sons da língua* (pp. 31-46). Edipucrs.

Bione, T., Grimshaw, J., & Cardoso, W. (2017). An evaluation of TTS as a pedagogical tool for pronunciation instruction: the 'foreign' language context. In K. Borthwick, L. Bradley & S. Thouësny (Eds), *CALL in a climate of change: adapting to turbulent global conditions – short papers from EUROCALL 2017* (pp. 56-61). Research-publishing.net. https://doi.org/10.14705/rpnet.2017.eurocall2017.689

Cardoso, W., Smith, G., & Garcia Fuentes, C. (2015). Evaluating text-to-speech synthesizers. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 108-113). Research-publishing.net. https://doi.org/10.14705/rpnet.2015.000318

Gomes, A. A. A. (2015). *A epêntese vocálica inicial em clusters sC por aprendentes brasileiros de inglês como LDE*. Edufcg.

John, P., & Cardoso, W. (2017). A comparative study of text-to-speech and native speaker output. In J. Demperio, E. Rosales & S. Springer (Eds), *Proceedings of the meeting on English language teaching* (pp. 78-96). Université du Québec à Montréal Press.

Labov, W. (1972). *Sociolinguistic patterns*. University of Pennsylvania Press.

Liakin, D., Cardoso, W., & Liakina, N. (2017). The pedagogical use of mobile speech synthesis (TTS): focus on French liaison. *Computer Assisted Language Learning, 30*(3-4), 348-365. https://doi.org/10.1080/09588221.2017.1312463

Ortiz, E., & Ivette, C. (2005). *The development of sC onset clusters in Spanish English*. Doctoral dissertation, Concordia University.

Soler-Urzúa, F. (2011). The acquisition of English /ɪ/ by Spanish speakers via text-to-speech synthesizers: a quasi-experimental study. Master's thesis. Concordia University.

Published by Research-publishing.net, a not-for-profit association
Contact: info@research-publishing.net

Future-proof CALL: language learning as exploration and encounters – short papers from EUROCALL 2018
Edited by Peppi Taalas, Juha Jalkanen, Linda Bradley, and Sylvie Thouësny

Publication date: 2018/12/08

Typeset by Research-publishing.net
Cover theme by © 2018 Antti Myöhänen (antti.myohanen@gmail.com)
Cover layout by © 2018 Raphaël Savina (raphael@savina.net)
Drawings by © 2018 Linda Saukko-Rauta (linda@redanredan.fi)

ISBN13: 978-2-490057-22-1 (Ebook, PDF, colour)
ISBN13: 978-2-490057-23-8 (Ebook, EPUB, colour)
ISBN13: 978-2-490057-21-4 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, UK: British Library.
Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: Décembre 2018.