PRAXIS PEDAGÓGICA

Artículo de investigación

Cómo citar: Samaniego C., Mendoza X., & Vera R. (2024). Voice modulator Talk It in English pronunciation construction: The case of Spanish young learners. PRA, 24(37), 146-167. https://doi.org/10.26620/uniminuto.praxis.24.37.2024.146-167

ISSN: 0124-1494

eISSN: 2590-8200

Editorial: Corporación Universitaria Minuto de Dios - UNIMINUTO

Recibido: 6 junio 2024

Aceptado: 12 noviembre 2024

Publicado: 29 noviembre 2024

Conflicto de intereses: los autores han declarado que no existen intereses en competencia.

Voice modulator Talk It in English pronunciation construction: The case of Spanish young learners

El modulador de voz Talk It en la construcción de la pronunciación inglesa: El caso de los jóvenes aprendices españoles

Modulador de voz Talk It na construção da pronúncia do inglês: O caso dos jovens aprendizes de espanhol

Cristhian F. Samaniego Andrade
Universidad Laica Eloy Alfaro de Manabí Extensión Chone
https://orcid.org/0009-0001-7273-544X
cristhian.samaniego@pg.uleam.edu.ec
Ecuador

Henry Xavier Mendoza Ponce
Universidad Laica Eloy Alfaro de Manabí Extensión Chone
https://orcid.org/0000-0003-2322-0383
xavier.mendoza@uleam.edu.ec
Ecuador

José Rafael Vera Vera
Escuela Superior Politécnica Agropecuaria de Manabí Manuel Félix López
https://orcid.org/0000-0003-1721-8770
jose_verav@espam.edu.ec
Ecuador

Abstract

Speaking has largely been one of the most challenging skills for language learners to master, especially in early stages of a foreign language learning process (Natasia & Angelianawati, 2022), and it demands a good command of pronunciation. Thus, the purpose of this study was to analyze the effectiveness of the voice modulator (VM) *Talk It* to help English as a Foreign Language (EFL) young learners enhance their pronunciation. This exploratory research used a quasi-experimental design with a mixed approach (quantitative-qualitative). The methodology consisted of three stages: (1) the analysis of the participants' current command of pronunciation; (2) the implementation of VM *Talk It* as a tool to promote pronunciation enhancement in the experimental group; (3) the comparison of pronunciation enhancement between the control and experimental groups. The results showed that Spanish EFL learners experienced special difficulties with phonemes entirely absent or differently pronounced in their mother tongue. The most affected phonemes regarding vowels were /ə/ /e/, /ɒ/, /ɪə/, /aʊ/, and /θ/, /ð/, /z/, /ʃ/ in the case of consonants. It was found that VM *Talk It* effectively contributed to the enhancement of pronunciation in the experimental group raising a significant gap compared to the control one; however, previous studies in literature demonstrate

that VMs are principally effective in early stages of EFL learning. In conclusion, while VMs offer significant benefits for young learners and can serve as valuable tools in helping them achieve native-like pronunciation, it is essential to further enhance these tools in order to support progression toward more advanced stages of language learning.

Keywords: Education technology, speaking skill, pronunciation, voice modulators, young learners.

Resumen

Hablar ha sido, en gran medida, una de las habilidades más desafiantes para los estudiantes de idiomas, especialmente en las primeras etapas del proceso de aprendizaje de una lengua extranjera (Natasia & Angelianawati, 2022), y requiere un buen dominio de la pronunciación. El propósito de este estudio fue analizar la efectividad del modulador de voz (VM) *Talk It* para ayudar a los estudiantes jóvenes de inglés como lengua extranjera (EFL, por sus siglas en inglés) a mejorar su pronunciación. Esta investigación exploratoria utilizó un diseño cuasi-experimental con un enfoque mixto (cuantitativo-cualitativo). La metodología consistió en tres etapas: (1) el análisis del dominio actual de la pronunciación por parte de los participantes; (2) la implementación del VM *Talk It* como herramienta para promover la mejora de la pronunciación en el grupo experimental; (3) la comparación de la mejora de la pronunciación entre los grupos de control y experimental. Los resultados mostraron que los estudiantes de EFL de habla hispana enfrentaron dificultades especiales con fonemas que están completamente ausentes o se pronuncian de manera diferente en su lengua materna. Los fonemas más afectados en el caso de las vocales fueron /ə/ /e/, /ɒ/, /ɪə/, /aʊ/, y /θ/, /ð/, /z/, /ʃ/ en el caso de las consonantes. Se encontró que *Talk It* contribuyó de manera efectiva a la mejora de la pronunciación en el grupo experimental, generando una diferencia significativa en comparación con el grupo de control; sin embargo, estudios previos en la literatura demuestran que los VM son principalmente efectivos en las primeras etapas del aprendizaje de EFL. En conclusión, aunque los VM ofrecen beneficios significativos para los estudiantes jóvenes y pueden ser herramientas valiosas para ayudarles a alcanzar una pronunciación nativa, es esencial seguir mejorando estas herramientas para apoyar el progreso hacia etapas más avanzadas del aprendizaje de idiomas.

Palabras clave: tecnología educativa, destreza de hablar, moduladores de voz, pronunciación, estudiantes jóvenes.

Resumo

Falar tem sido, em grande parte, uma das habilidades mais desafiadoras para os estudantes de línguas, especialmente nas primeiras etapas do processo de aprendizagem de uma língua estrangeira (Natasia & Angelianawati, 2022), e exige um bom domínio da pronúncia. O objetivo deste estudo foi analisar a eficácia do modulador de voz (VM) Talk It para ajudar os estudantes jovens de inglês como língua estrangeira (EFL, sigla em inglês) a melhorar sua pronúncia. Esta pesquisa exploratória utilizou um desenho quase-experimental com uma abordagem mista (quantitativa-qualitativa). A metodologia consistiu em três

etapas: (1) a análise do domínio atual da pronúncia pelos participantes; (2) a implementação do VM Talk It como ferramenta para promover a melhoria da pronúncia no grupo experimental; (3) a comparação da melhoria da pronúncia entre os grupos de controle e experimental. Os resultados mostraram que os estudantes de EFL de fala hispânica enfrentaram dificuldades especiais com fonemas que estão completamente ausentes ou são pronunciados de maneira diferente em sua língua materna. Os fonemas mais afetados no caso das vogais foram /ə/ /e/, /ɒ/, /ɪə/, /aʊ/, e /θ/, /ð/, /z/, /ʃ/ no caso das consoantes. Constatou-se que Talk It contribuiu de maneira eficaz para a melhoria da pronúncia no grupo experimental, gerando uma diferença significativa em comparação com o grupo de controle; no entanto, estudos anteriores na literatura demonstram que os VMs são principalmente eficazes nas primeiras etapas da aprendizagem de EFL. Em conclusão, embora os VMs ofereçam benefícios significativos para os estudantes jovens e possam ser ferramentas valiosas para ajudá-los a alcançar uma pronúncia nativa, é essencial continuar aprimorando essas ferramentas para apoiar o progresso em direção a estágios mais avançados da aprendizagem de línguas.

Palavras-chave: Tecnologia educacional, habilidade de falar, moduladores de voz, pronúncia, estudantes jovens.

Introduction

Pronunciation is the basis of effective communication (Plailek & Essien, 2021). It implies a number of challenges, but difficulties grow bigger when learners start trying to sound like a native speaker (Djurayeva, 2021). Indeed, Whitehead & Ryu (2023) assert that there are several factors affecting the process of commanding a native-like pronunciation in teachers-related, students-related, and curriculum- and materials-related implications. In this context, previous findings suggest that English as a Foreign Language (EFL) learners have to be aware of their phonetic differences from native speakers (Plailek & Essien, 2021). However, despite the fact that pronunciation is not directly considered as one of the four language-learning skills, it is undeniable that independent attention to pronunciation will bring learners extra benefits (Quoc *et al*., 2021).

Native-like speaking training helps learners to reach a good command of pronunciation. It is one of the most useful skills in communication, yet also one of the hardest to master (Leong & Ahmadi, 2017; Levelt, 1995; Natasia & Angelianawati, 2022). Indeed, the articulation of words and sentences is crucial when conveying ideas orally as native listeners rely more on stress patterns rather than individual segmental pronunciation (Levelt, 1993 1995). Immersing in the understanding of pronunciation is necessary to analyze its nature, which is composed of several factors; in turn, those are grouped as segmentals and suprasegmentals. The first group is aligned to the sound of individual units of a word (individual consonants, vowels and syllables sounds), whereas the second involves aspects of the overall phrase sound such as word stress, phrase rhythm, construction stress, primary stress, and intonation.

Moreover, Deterding (2004) states that EFL learners may face uncertainty caused by two main factors: (1) different varieties of the language, and (2) different phonological production. In this regard, segmentals in the English language such as the vowels present uncertain variety of sounds; one of the most outstanding cases is the "schwa sound" [ə], which Ahn (2001) defines as the phenomenon of reducing the vowel sound in unstressed syllables. This variation in sound production affects EFL language acquisition as learners are used to articulating those graphemes in their native tongue.

On the other hand, EFL students frequently face significant challenges in pronunciation that can hinder their communicative effectiveness. A primary obstacle is the presence of phonemes absent in their native language (Hayes-Harb & Barrios, 2021), or differences in graphemes between languages that complicate precise sound production (Al-Hamzi *et al*., 2021; Nyarks & Hanson, 2023), which can eventually lead to confusion and poor articulation. Additionally, foreign language learners struggle with the nuances of intonation (Qizi, 2024) and rhythm (Baills, 2021), which are essential for conveying meaning and ensuring natural speech patterns. Furthermore, the blending of words in rapid speech, a common characteristic in many languages, poses difficulties in both comprehension and reproduction, making the acquisition of an appropriate accent particularly challenging, especially in the absence of regular exposure to native speakers. Addressing these pronunciation challenges is crucial, as they require deliberate practice and constructive feedback to achieve proficiency.

When it comes to training students to reach a native-like pronunciation in a non-English speaking country, beginners are forced to create their own learning environment. To this end, technology may become a great ally to reach their goal. Indeed, computer-assisted systems of words' sounds' articulation help young learners enhance their pronunciation in a short term of training compared to those traditional teacher-led methods (Neri *et al*., 2008). For instance, software designed to recognize voices can rate spoken words according to an individual's utterance and avoid mispronunciation (Liu *et al*., 27-29 December 2011).

Likewise, Yenkimaleki and Van Heuven (2022) prompt a large discussion about the different ways of learning, where inductive methods, comparisons between mother and second language production features, and the awareness of suprasegmentals are relevant contributors to the process of improving the skill. However, they have been analyzed separately, raising a gap about how learners could work together through a digital tool aimed at enhancing students' pronunciation.

As a result of the accelerated evolution of digital tools, teachers must be ready to examine their usefulness and implement those that remains effective, especially those oriented to help learners improve their language skills (Durham, 2023). It demands

trainers to keep up with the global changes emerging from the line of technology development. Indeed, Sriudomkij & Sopirak (2013) assert that young learners have significant knowledge of technological tools use, and teachers have to keep on a path of continuous preparation to face the learning environment they are constantly exposed due to global education changes.

In this regard, voice modulators (VMs) provide learners with the opportunity to promote foreign language pronunciation improvement (Felps *et al*., 2009). In other words, language learners can build their perception and production of oral phrases with confidence and fluency. In the context of the information and communication technologies (ICT) era, VMs have shown up to help learners equip their virtual learning environment (VLE) to expose themselves to different types of language exchanging or talking simulating contexts. In relation with the aforementioned situations, the scientific question of how VMs affect young learners' pronunciation is stated.

Methodology

The purpose of this study was to determine the effectiveness of the voice modulator *Talk It* in helping EFL Spanish young learners enhance their pronunciation. For the depth of the study, this research was of an exploratory level due to the limited specific theoretical support, and consisted of a qualitative and quantitative approach in line with the data analyzed. The study followed a quasi-experimental type of research due to the manipulation of the independent variable (voice modulator *Talk It*) in one group compared to a control group to observe the behavior of the dependent variable (pronunciation).

This study considered the methodology implemented by Saed *et al*. (2021) in similar experimental research, which starts with the selection of participants and the establishment of two groups (control and experimental), then two speaking tests are administered (pre- and post-test) by certified experts according to international examination standards. In this study, that methodology was adapted to the context of the research, and comprised three main stages: (1) the analysis of the participants' current command of pronunciation; (2) the implementation of *Talk It* as a tool to promote pronunciation enhancement in the experimental group; (3) the comparison of pronunciation enhancement between the control and experimental groups.

Participants

The sample considered for this study was that of the upper basic students of the Unidad Educativa Fiscomisional Cinco de Mayo, located in the canton of Chone (province of Manabi), Ecuador. The sample was made of the Spanish students who reached an A2 level according to the CEFR (Common European Framework of Reference). They accredited that level in an official examination carried out by an American university. The total sample of the research was made up of 34 students which was split into two groups, the first half for the control group (17 students) and the second for the experimental one.

The present research was technically viable since the VM *Talk It* use is free, and there was the willingness of the Unidad Educativa Fiscomisional Cinco de Mayo to support the study with the participation of the upper basic education students who demonstrated a command of the English language at the A2 level. On the other hand, four EFL/ESL teachers of other educational institutions (who have a minimum certification of level B2 recognized by international English language evaluation organizations) participated in the study as peer-reviewers for the evaluation of learners' pronunciation.

Data collection and analysis procedure

The data collection consisted of two rounds: (1) diagnosing initial situation of the two groups, and (2) comparing the experimental group pronunciation enhancement compared to the control one. However, the procedure for the collection of data was the same for the two rounds. In both of them, the researchers demanded students under study to record a passage reading of an A2 students' book under the guidelines of the CEFR. In order to ensure the reliability of information, the voice note recording was carried out in the computing laboratory of the institution where participants were recorded under the supervision of the researcher and the audio track was anonymized before being summitted to peer-reviewers for their evaluation.

To answer the scientific question, the researcher performed 20 training sessions with the participants of the experimental group to practice their reading skill on their own by using the VM *Talk It* to improve their pronunciation when reading

passages within the A2 language threshold. In the first session, the researchers gave instructions on how to use the tool, but the other 19 were only supervised while the participants freely used the app. This stage was aimed at stimulating brain memory by the repetition method suggested by Zhan *et al*. (2018) as it implies metacognitive processes that necessarily deal with neuroscientific aspects; these repetitions address sensory processing concepts, which directly interfere with the variables manipulated in this study. For instance, Predictive Coding (PC) was used to expose individuals to familiar contexts that facilitate the assimilation of the content (Lupyan & Clark, 2015).

After a week, results of the evaluation were collected and analyzed on the Microsoft Excel program. Criteria considered in the pre- and post-test were organized into the two dimensions of pronunciation. The first one embodies the nature of segmentals (vowel and consonant utterance) for which the analysis method implemented by Gallego-Ortega *et al*. (2017) was used. On the other hand, there is limited quantitative evidence of suprasegmental evaluations to quantify mispronunciation as it may be observed in previous studies (Miller, 1978; Tong *et al*. 2008); for this reason, a Likert scale was prepared to evaluate this dimension (poor, fair and good). Suprasegmentals evaluated were pitch, sentence stress, articulation pattern, and intonation. Fluency and rhythm were not considered as they might have been scored as low due to the participants' early stage of learning. For the final stage, the evaluation of the learners' progression in pronunciation was conducted through the application of the assessment criteria applied in stage one.

Results

Pronunciation consists of a number of aspects, among which phonetics play a very significant role. According to The London School of English (16 August 2017), these are grouped into various sound categories. Learning about them (consonant and vowel) is substantial to analyze the blocks of sounds that may be affected the most by participants. Thus, in Table 1 these categories are shown.

Table 1. Phonetic alphabet: symbols, category of sound, and examples.

Symbol	Category of sound	Examples	Symbol	Category of sound	Examples
e	Short vowel	Went, intend, send, letter.	ð	Consonants Sounds: Fricatives	There, those, brothers, others.
æ		Cat, hand, nap, flat, have	z		Zoo, crazy, lazy, zigzag, nose
ʌ		Fun, love, money, one, London, come	ʃ		Shirt, rush, shop, cash
ʊ		Put, look, should, cook, book, look	ʒ		Television, delusion, casual
ɒ		Rob, top, watch, squat, sausage	h		High, help, hello.
ə		Alive, again, mother	p	Consonants Sounds: Plosives	Pin, cap, purpose, pause.
iː	Long vowel	Need, beat, team.	b		Bag, bubble, build, robe
ɜː		Nurse, heard, third, turn.	t		Time, train, tow, late.
ɔː		Talk, law, bored, yawn, jaw.	d		Door, day, drive, down, feed.
uː		Few, boot, lose, gloomy, fruit, chew.	k		Cash, quick, cricket, sock.
ɑː		Fast, car, hard, bath.	g		Girl, green, grass, flag.
ɪə	Diphthong Vowels	Near, ear, clear, tear, beer, fear	tʃ	Consonants Sounds: Affricates	Choose, cheese, church, watch.
eə		Hair, there, care, stairs, pear	dʒ		Joy, juggle, juice, stage.
eɪ		Face, space, rain, case, eight	m	Consonants Sounds: Nasals	Room, mother, mad, more.
ɔɪ		Joy, employ, toy, coil, oyster.	n		Now, nobody, knew, turn.
aɪ		My, sight, pride, kind, flight	ŋ		King, thing, song, swimming.
əʊ		No, don't, stones, alone, hole	r	Consonants Sounds: Approximants	Road, roses, river, ring, ride.
aʊ		Mouth, house, brown, cow, out	j		Yellow, usual, tune, yesterday, yard.
f	Consonants Sounds: Fricatives	Full, Friday, fish, knife.	w		Wall, walk, wine, world.
v		Vest, village, view, cave.	l and t		Law, lots, leap, long, pill, cold, chill, melt.
θ		Thought, think, bath.			

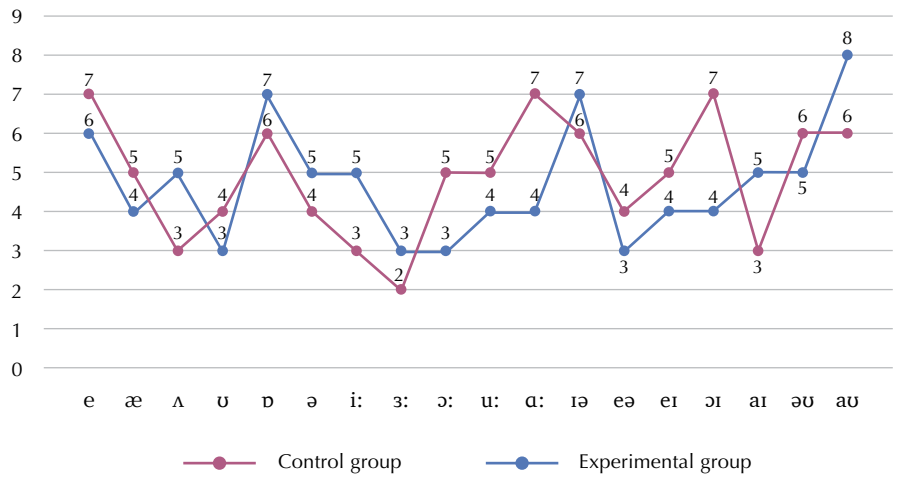
Note. Prepared by the authors based on The London School of English (16 August 2017

Initial diagnosis of pronunciation in both groups

In the first round of pronunciation assessment, the two groups demonstrated significant shortcomings in the production of phonemes. The results obtained from round 1 were segmented in two groups of analysis: (1) vowel sound in Figure 1, and (2) consonant sound in Figure 2.

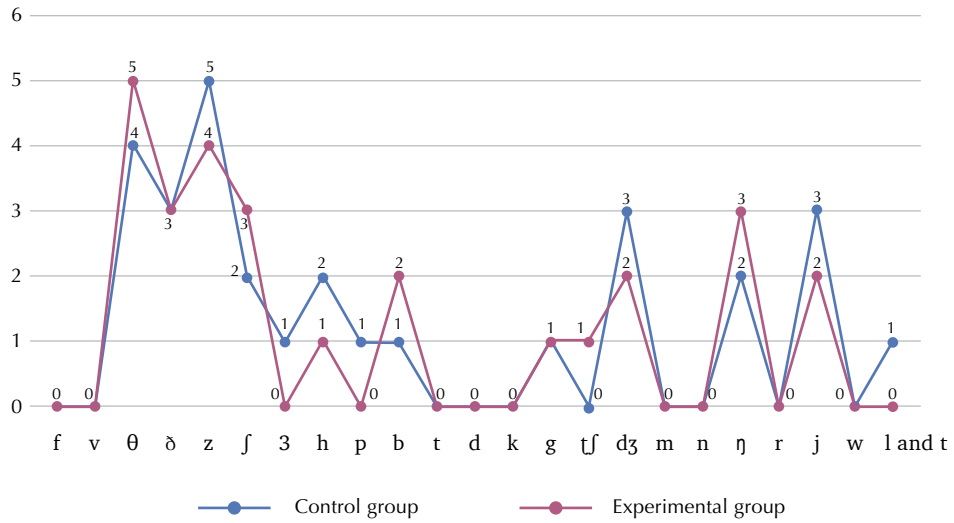
As shown in Figure 1, the number of errors made by learners in both groups fall within a moderate range. The greatest discrepancy in pronunciation errors was evident in the /ɔɪ/ phoneme, with a difference of three errors between them. Additionally, the most affected phonemes by both groups were: /ə/ /e/, /ɒ/, /ɪə/, and /aʊ/. Other phonetic affectations were detected in both groups separately, such as in the case of /ɑ:/ in the control group and /əʊ/ in the experimental one. On the other hand, the results of the consonant sound produced by the two groups are shown in Figure 2.

Figure 1. Vowel phonetic affectations.



Note. Prepared by the authors. Data obtained from the peer-reviewers assessment of the voice recording note. There was no discrepancy on the peer-reviewers' criteria about pronunciation errors.

Figure 2. Consonant phonetic affectations.

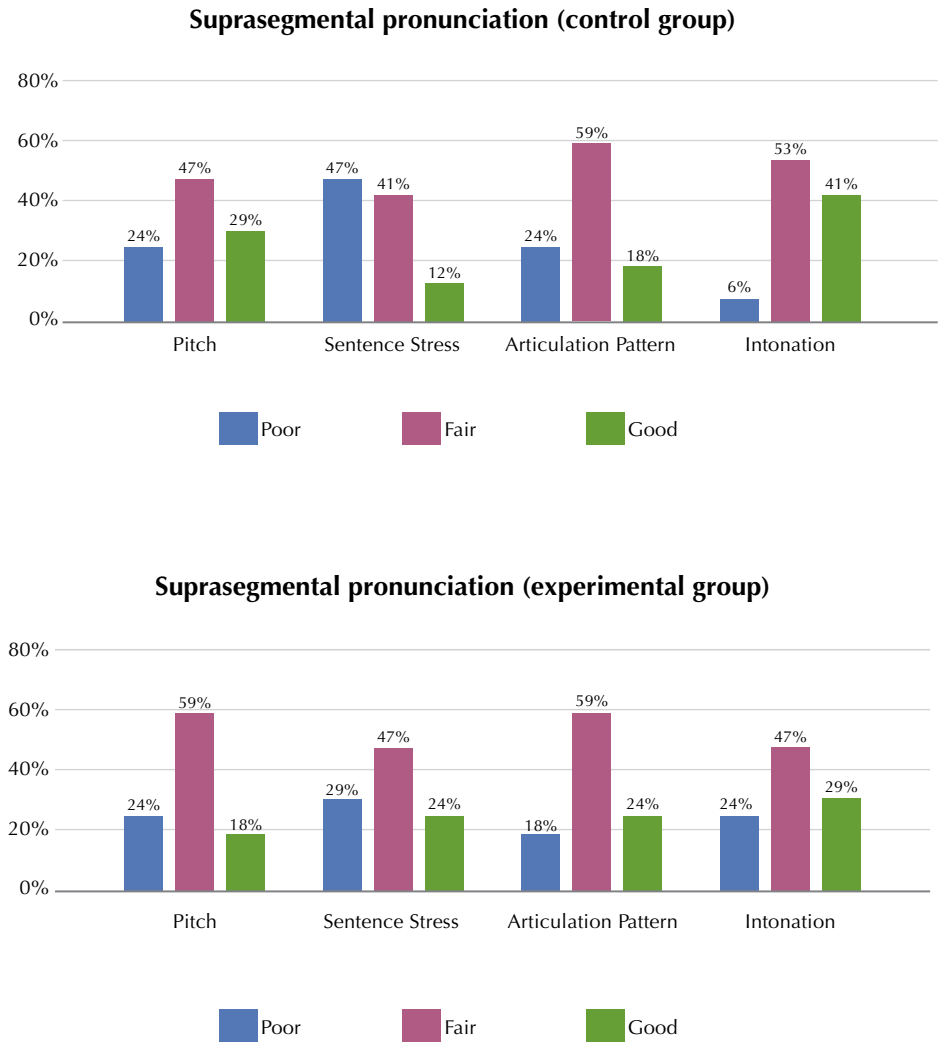


Note. Prepared by the authors. Data obtained from the peer-reviewers assessment of the voice recording note. There was no discrepancy on the peer-reviewers' criteria about pronunciation errors.

As it may be seen in Figure 2, the learners faced greater difficulties in the pronunciation of the phonemes of the fricative sound /θ/, /ð/, /z/, and /ʃ/ — this being the largest block affected. Yet, a significant impairment was found in the affricate consonant sound group (/dʒ/); a nasal (/ŋ/) and an approximant (/j/). It can be inferred that these effects occur due to the total absence or differences of these sounds in Spanish (the mother tongue of the students under study).

In line with the pronunciation of suprasegmentals, they were assessed for the aspect of pronunciation (pitch, sentence stress, articulation pattern, and intonation) appropriate to the passage, using a Likert scale to qualitatively indicate the level of command of each aspect as poor (1), fair (2), and good (3). The results show a comparison between two groups (Figure 3).

Figure 3. Evaluation of suprasegmental pronunciation performed by the two groups.



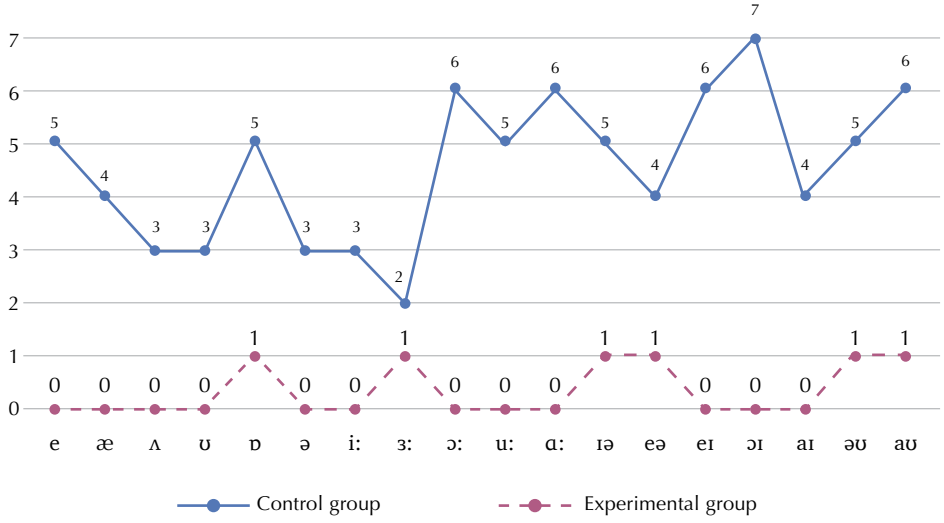
Note. Prepared by the authors. Data obtained from peer-reviewed evaluation of young learners' speech recording.

In the pitch block, both groups primarily demonstrated fair pronunciation, while the control group reflected greater limitations in sentence stress. However, the experimental group also presented marked shortcomings in terms of sentence stress, while the articulation pattern and intonation blocks did not represent major challenges for the students of both groups. Moreover, the prior assessment showed that, despite the similarity in relation to students' level, they still have remarkable differences in terms of suprasegmental pronunciation, raising a research gap that requires scientific observation. Though, these important differences within the group did not cause a bias for this research, as the two assessment moments were developed on the basis of the same individuals in each group.

Evaluation of segmentals and suprasegmentals pronunciation after VM Talk It implementation

The results of the pronunciation evaluation following the implementation of the VM *Talk It* made it viable to assess its effectiveness. Figures 4 and 5 depict the recorded numbers of affected phonemes (both vowel and consonant sounds, respectively) by individuals in each group.

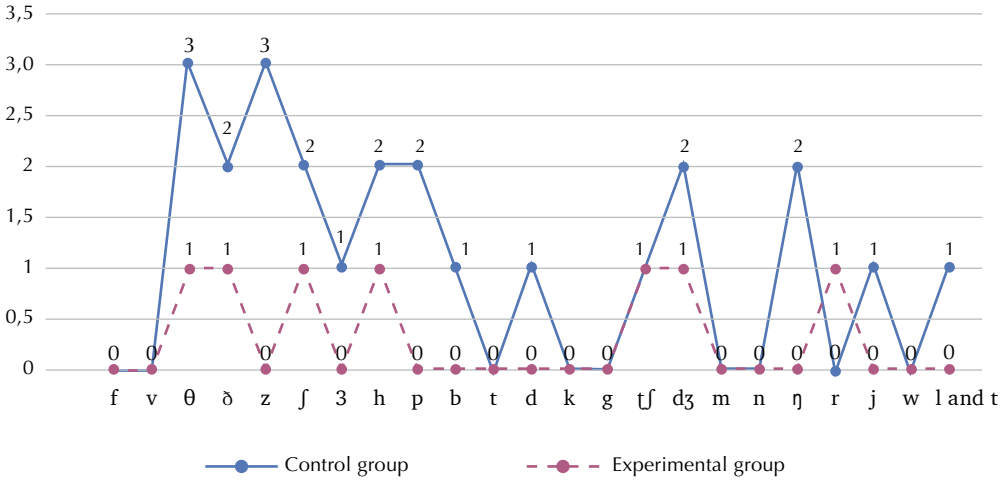
Figure 4. Vowel phonetic affectations.



Note. Prepared by the authors. Data obtained from the peer-reviewers assessment of the voice recording note. There was no discrepancy on the peer-reviewers' criteria about pronunciation errors.

As it can be seen in Figure 4, errors in vowel phonemes (short, long and diphthong sounds) decreased significantly in the experimental group, demonstrating substantial improvements in pronunciation, and allowing the students to express themselves better in EFL. On the other hand, the phonetic affectations of consonants by both groups are compared in Figure 5.

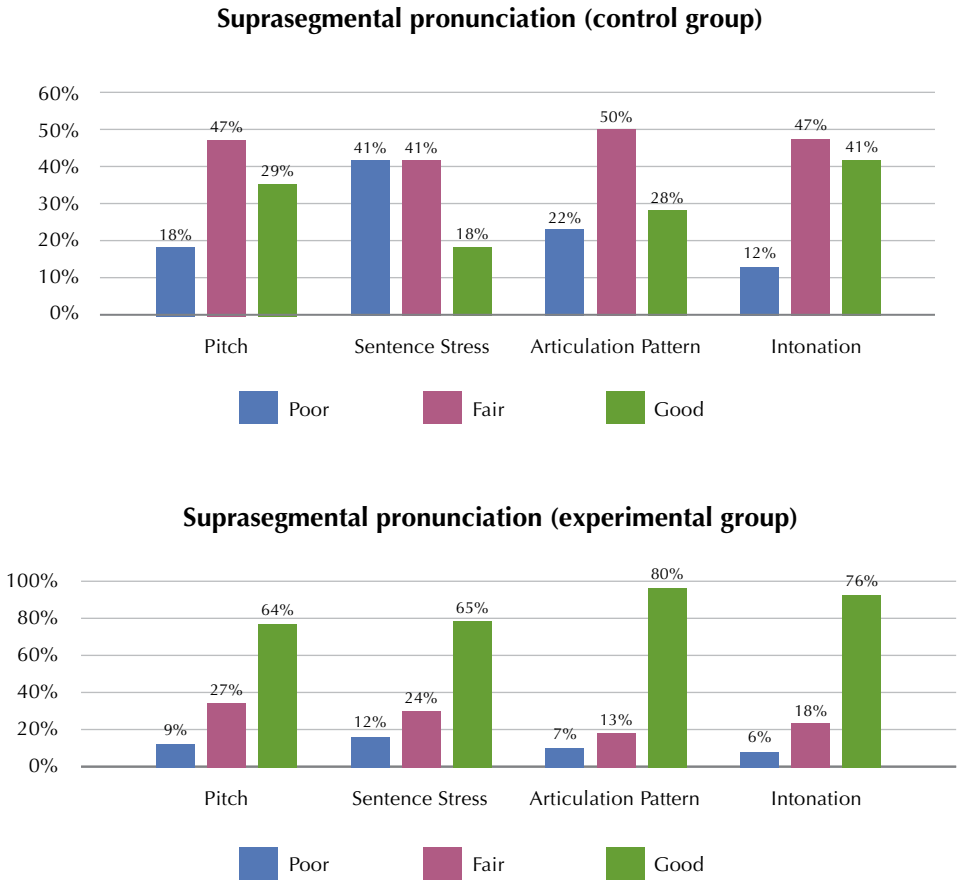
Figure 5. Consonants phonetic affectations.



Note. Prepared by the authors. Data obtained from the peer-reviewers assessment of the voice recording note. There was no discrepancy on the peer-reviewers' criteria about pronunciation errors.

Similar to what is shown in Figure 4, the phonetic affectations of the consonants (Figure 5) decreased significantly in terms of the number of students who mispronounced them. However, it is necessary to note that the effects raised are grouped into the same phonemes mispronounced in the evaluation prior to the implementation of the Talk It tool. In some way, this particular event suggests that it is caused by the root of the sounds of their native language or the absence of them in it. Likewise, the second evaluation also made it possible to establish a qualitative comparison of both groups in the face of their respective initial situation in relation to the pronunciation of the suprasegmentals. The results of this assessment are shown in Figure 6.

Figure 6. . Results of the evaluation of suprasegmental pronunciation performed by the two groups after the VM *Talk It* implementation.



Note. Prepared by the authors. Data obtain from peer-reviewed evaluation of young learners' speech recording.

The results of the evaluation demonstrated a substantial improvement in the management of suprasegmentals within the pronunciation in the experimental group (Figure 6). A symmetrical qualitative scale could be observed in all the blocks evaluated, where the "poor" criterion presents a highly symbolic reduction in this group in terms of its initial situation. In addition, the decrease in the "fair" rating tends to group

together the errors of students whose pronunciation had been evaluated as "poor" in the first round. By contrast, the results obtained in the control group did not show significant variations, but did show a wide gap in ineffectiveness compared to the experimental group, which demonstrates that the VM Talk It is a powerful tool for improving young learners' pronunciation.

Discussion

Speaking is one of the most complex skills in language communication, and learners often strive for native-like pronunciation (Losavio, 2023). The results in this study revealed that students struggled mainly with segmentals containing phonemes absent in Spanish. This finding aligns with previous studies, such as that of Noviyenty and Putri (2021), who examined the influence of the mother tongue on English pronunciation. Their research concluded that students often pronounce English words differently due to this interference. Similarly, multiple authors in the scientific literature corroborate these findings (Rohach & Kishko, 2023; Septianasari, 2019; Shanmugasundaram & Noble-Jebakumar, 2022;).

In addition, learners in both groups experienced difficulties at the initial stage, due to the phonetic variations in the English language that resulted atypical in Spanish, such as the schwa sound in unstressed syllables and consonant sounds, which are uncommon or entirely absent in their mother tongue. This difficulty is not exclusive to Spanish speakers; for instance, Islam (2020) reported similar issues among Bengali speakers.

On the other hand, findings revealed that one of the most common affectations in line with suprasegmentals was sentence stress. This result is consistent with the findings by Kucukoglu (2012), who noted that this issue also extends to EFL teachers. In this regard, pitch and intonation command remained poor in most cases, as learners tended to read the message without internalizing it. Yet, peer reviewers' criteria about articulation pattern, pitch and intonation resulted as fair in most cases.

Finally, the scientific literature on the usefulness of VMs as tools that help language acquirers enhance their oral skill remains limited. However, their contribution to the development of accurate pronunciation and, consequently, proficient speaking

skills is noteworthy, as evidenced in the study by Baradaran and Davvari (2010). Despite that the results of this research demonstrate the positive impact of using the VM *Talk It* in EFL learning, new technological advancements in the field of VMs demands special attention regarding their effectiveness in EFL speaking construction, raising a gap for future research.

Conclusions

Scientific evidence has demonstrated that developing the speaking skill in the English language is heavily influenced by the pronunciation patterns of learners' mother tongue. This not only makes it challenging for them to produce clear speech, but also impacts the recipient's understanding of the message, contributing to early learner insecurities when engaging in EFL conversations. Scientific evidence has consistently shown that learners face difficulties in imitating native speakers to produce clearer and more understandable oral messages. Consequently, there is a need for tools that enable learners to recreate simulation environments to enhance their speaking ability.

From the pronunciation errors identified at both segmental and suprasegmental levels, possible causes related to pronunciation patterns rooted in learners' mother tongue can be inferred. Within the context of the sample included in this study, it was observed during the initial stage that vowel and consonant phonemes, which either did not exist or were pronounced differently in the participants' mother tongue, presented similar challenges in both groups. This observation suggests the identification of a new research gap, indicating potential avenues for further studies concerning sound variations among phonemes represented by written codes, but pronounced differently in various languages.

While the scientific literature on the effectiveness of VM for improving EFL pronunciation is still limited, the findings of this study demonstrate their usefulness at early stages of learning, as it allows learners to sound like English native speakers. Therefore, it may be concluded that VMs may bring together important benefits to young learners' and become a strategic ally for them to accurately develop their speaking skill.

References

Ahn, S.-W. (2001). The phonetics and phonology of English schwa. *Korean Journal of English Language and Linguistics, 1*(2), 311-329. https://koreascience.kr/article/JAKO200111922008576.pdf

Al-Hamzi, A. M. S., Al-Shrgabi, A. h. A. S., Al-Haidari, A. S. A., Faraj, M. M. A., & Al-Housali, T. A. A. (2021). Pronunciation errors of English front vowels by Yemeni EFL learners. *PAROLE: Journal of Linguistics and Education, 11*(1), 41-56. https://doi.org/10.14710/parole.v11i1.41-56

Baills, F. (2021). *Using the hands to embody prosody boosts phonological learning in a foreign language* [doctoral dissertation, Universitat Pompeu Fabra]. Institutional repository UPF. http://hdl.handle.net/10803/673777

Baradaran, A., & Davvari, Z. (2010). The Impact of Utilizing Computer Assisted Language Learning on EFL Learners' Foreign Accent Reduction. *JELS: Journal of English Language Studies, 1*(4), 41-62. https://journals.iau.ir/article_510805_1e9b97b2c6d819118a17596b088aa9e6.pdf

Deterding, D. (2004). How many vowel sounds are there in English. *STETS Language & Communication Review, 19*(10), 19-21. https://fass.ubd.edu.bn/staff/docs/DD/STETS-vowels.pdf

Djurayeva, Y. A. (2021). Enhancing English pronunciation in learning process. *Academic Research in Educational Sciences. Volume 2* (CSPI Conference 2; pp. 302-306). https://cyberleninka.ru/article/n/enhancing-english-pronunciation-in-learning-process/viewer

Durham, C. (2023). Teacher educator toolbox: Strategic mediation in teacher preparation coursework focused on technology. *System, 114*. https://doi.org/10.1016/j.system.2023.103031

Felps, D., Bortfeld, H., & Gutierrez-Osuna, R. (2009). Foreign accent conversion in computer assisted pronunciation training. *Speech Communication, 51*(10), 920-932. https://doi.org/10.1016/j.specom.2008.11.004

Gallego-Ortega, J. L., Gómez-Pérez, I. A., & Ayllón-Blanco, M. F. (2017). Valoración de un programa escolar para el desarrollo de habilidades fonológicas en niños españoles. *Revista Brasileira de Educação, 22*(77). http://doi.org/10.1590/S1413-24782017227167

Hayes-Harb, R., & Barrios, S. (2021). The influence of orthography in second language phonological acquisition. *Language Teaching, 54*(3), 297-326. https://doi.org/10.1017/S0261444820000658

Islam, S. M. (2020). Segmental errors in English pronunciation of non-native English speakers. *Journal of Education and Social Sciences, 16*(1), 14-24. https://www.jesoc.com/wp-content/uploads/2020/12/JESOC16-011.pdf

Kucukoglu, H. (2012). Sentence stress and learning difficulties of ELT teachers: A case study. *Procedia - Social and Behavioral Sciences, 46*, 4065-4069. https://doi.org/10.1016/j.sbspro.2012.06.198

Leong, L.-M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English speaking skill. *International Journal of Research in English Education, 2*(1), 34-41. https://ijreeonline.com/files/site1/user_files_68bcd6/sma1357-A-10-26-1-fefa0eb.pdf

Levelt, W. J. M. (1993). *Speaking: From intention to articulation*. The MIT Press.

Levelt, W. J. M. (1995). The ability to speak: From intentions to spoken words. *European Review, 3*(1), 13-23. https://doi.org/10.1017/S1062798700001290

Liu, Z., Wen, F., Lu, Z., & Chen, M. (2011, 27-29 December). Research on pronunciation scoring technology in language training system. *Proceeding of the International Conference on e-Education, Entertainment and e-Management* [paper]. Bali, Indonesia. https://doi.org/10.1109/ICEEEM.2011.6137788.

Losavio, A. (2023). Late starts leading to native-like pronunciation in second language acquisition [thesis dissertation, University of Central Florida]. Institutional repository UCF. https://stars.library.ucf.edu/etd2020/1805

Lupyan, G., & Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science, 24*(4), 279-284. https://doi.org/10.1177/0963721415570732

Miller, J. L. (1978). Interactions in processing segmental and suprasegmental features of speech. *Perception & Psychophysics, 24*(2), 175-180. https://doi.org/10.3758/BF03199546

Natasia, G., & Angelianawati, L. (2022). Students' perception of using storytelling technique to improve speaking performance at SMPN 143 Jakarta Utara. *Journal of English Teaching, 8*(2), 282-292. https://doi.org/10.33541/jet.v8i2.4063

Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning, 21*(5), 393-408. https://doi.org/10.1080/09588220802447651.

Noviyenty, L., & Putri, M. I. (2021). Mother tongue interference towards students' English pronunciation: A case study in IAIN Curup. In: *Proceedings of the International Conference on Educational Sciences and Teacher Profession (ICETeP 2020)* (pp. 283-290). Atlantis Press. https://doi.org/10.2991/assehr.k.210227.049

Nyarks, A., & Hanson, G. I. (2023). An assessment of sound system in Ibibio language: the clarity and phonological challenges. *International Journal of Educational and Scientific Research Findings, 5*(1), 81-89.

Plailek, T., & Essien, A. M. (2021). Pronunciation problems and factors affecting English pronunciation of EFL students. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(12), 2026-2033. https://turcomat.org/index.php/turkbilmat/article/view/7734

Qizi, K. M. A. (2024). Common intonation mistakes when developing English as an overseas language for lower-intermediate adults. *American Journal of Social Sciences and Humanity Research, 4*(01), 20-34. https://doi.org/10.37547/ajsshr/Volume04Issue01-04

Quoc, T. X., Thanh, V. Q., Dang, T. D. M., Mai, N. D. N., & Nguyen, P. N. K. (2021). Teachers' perspectives and practices in teaching English pronunciation at Menglish Center. *International Journal of TESOL & Education, 1*(2), 158-175. https://i-jte.org/index.php/journal/article/view/62

Rohach, L., & Kishko, O. (2023). Mother tongue interference on students' acquiring English pronunciation. *Contemporary Research on Foreign Philology, 1*(23), 289-298. https://doi.org/10.32782/2617-3921.2023.23.289-298

Saed, H. A., Haider, A. S., Al-Salman, S., & Hussein, R. F. (2021). The use of YouTube in developing the speaking skills of Jordanian EFL university students. *Heliyon, 7*(7). https://doi.org/10.1016/j.heliyon.2021.e07543

Septianasari, L. (2019). Mother tongue issues and challenge in learning English as foreign language. *IJIET (International Journal of Indonesian Education and Teaching), 3*(2), 204-214. https://doi.org/10.24071/ijiet.v3i2.1941

Shanmugasundaram, R., & Noble-Jebakumar, A. (2022). Mother tongue influence on English pronunciation: A case study in college students. *Journal for Educators, Teachers and Trainers, 13*(4), 312-316. https://jett.labosfor.com/index.php/jett/article/view/1206

Sriudomkij, P., & Sopirak, S. (2013). Factors influencing the development in English pronunciation skills training in primary students of schools in the Primary Education Services Area Office in Bangkok. *Procedia - Social and Behavioral Sciences, 103*, 396-399. https://doi.org/10.1016/j.sbspro.2013.10.351

The London School of English. (2017, 16 August). *Phonetic alphabet - Examples of sounds* [online]. https://n9.cl/xr1va

Tong, Y., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes, 23*(5), 689-708. https://doi.org/10.1080/01690960701728261

Whitehead, G. E. K., & Ryu, Y. (2023). "I am not a native speaker…": Exploring the perceived pronunciation teaching difficulties faced by Korean public elementary school English teachers. *System, 115*. https://doi.org/10.1016/j.system.2023.103056

Yenkimaleki, M., & Van Heuven, V. J. (2022). Comparing the nativeness vs. intelligibility approach in prosody instruction for developing speaking skills by interpreter trainees: An experimental study. *Speech Communication, 137*, 92-102. https://doi.org/10.1016/j.specom.2022.01.007

Zhan, L., Guo, D., Chen, G., & Yang, J. (2018). Effects of repetition learning on associative recognition over time: Role of the hippocampus and prefrontal cortex. *Frontiers in Human neuroscience, 12*. https://doi.org/10.3389%2Ffnhum.2018.00277