

Customizable Serious Speech Therapy Games with Dynamic Difficulty Adjustment for Children with Stigmatism

Sofia Martins, Sofia Cavaco

NOVA LINCS, Department of Computer Science
Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa
2829-516 Caparica, Portugal

Abstract

Children with speech sound disorders should attend speech and language therapy and should practice the speech exercises regularly to surpass their speech difficulties. Since doing the speech exercises often may be tedious, there is the need to motivate children to practice them. During the COVID-19 pandemic, speech and language pathologists had the need to adapt their procedures to others with less physical contact.

Here, we propose two serious games to motivate children with stigmatism on doing the speech exercises, which can be used at home and during face-to-face and online speech therapy sessions. The games use automatic speech recognition to classify speech productions. Visual and auditory feedback are used to help children understand their performance, and a hint system is used to help them perform the exercises correctly. A dynamic difficulty adjustment system is used to change the level of difficulty according to the child's speech performance in previous trials.

Keywords:

Serious therapy games, speech and language therapy, dynamic difficulty adjustment.

Introduction

As children grow, they learn how to speak through hearing and observing articulatory movements and trying to repeat them [1]. However, there are children who present difficulties in this process and, in some cases, speech sound disorders (SSD) may occur [2]. These difficulties can generate frustration, social isolation and changes in social relationships. Such situations can have a direct impact in the child's quality of life and academic success.

Unfortunately, it is very common for children to have these deviations. The best way to overcome them is by attending speech and language therapy. While children with SSD usually attend a weekly speech therapy session, this may be insufficient to reach speech improvements and children have to train outside the therapy sessions in order to overcome their difficulties.

In spite of the importance of practicing the speech exercises regularly, children may not have the responsibility and the awareness of the importance of such training. Thus, it is important to motivate them on practicing the exercises. A way of accomplishing this is through gamification. The gamification of speech exercises should include rewards, hints that help achieving what is expected in the exercises, activities with different levels and difficulties that adjust to the user's performance. Having different difficulty levels is important because it may be hard to find a challenging level adequate for all users [3, 4]. A system with these characteristics can break the

monotony of speech exercises and help the child accomplish what is intended.

Here we propose two serious games for speech therapy for European Portuguese (EP) speaking children with stigmatism. Stigmatism is a SSD that affects the production of sibilant phonemes, which in EP are: [s] (like the s phoneme in the word spell); [z] (like the z phoneme in the word zoo), [ʃ] (like the sh phoneme in short) and [ʒ] (like the s phoneme in casual) [5,6]. The games' goal is to motivate these children to practice speech exercises that help them master the production of sibilant consonants.

The proposed games integrate several speech exercises that are already used in speech therapy sessions and have been suggested by speech and language pathologists (SLPs) that we have interviewed. The games can be customized to meet the specific needs of each child. In addition, the games include visual and auditory feedback that help children understand when they have completed the exercises correctly or when they have failed. They also include a system of hints that helps children with more difficulties to perform the exercises. The games have several difficulty levels and can automatically adjust the level of difficulty to take into account the child's speech performance in previous trials.

An important aspect of this work is that the games use automatic speech recognition (ASR) systems. This allows children to use the games (and therefore practice the speech exercises) outside the speech therapy sessions, which, as mentioned above, is important for children to achieve speech improvements. Nonetheless, the games should always be introduced by a SLP to the child, and the SLP should customize the game according to the child's needs.

The games are suitable to be used during face-to-face therapy sessions but also during online sessions. This may help SLPs to keep children motivated to complete the exercises when face-to-face therapy sessions are not possible.

Methods

This paper discusses the development and validation of two computer games for speech therapy with children with stigmatism. The games are both integrated in an application for Windows, which gives access to the games and a shared rewards system. The games aim to help children with stigmatism to surpass their speech difficulties. In addition, this work aims to provide SLPs with a tool that helps them to motivate the children to practice the speech exercises at home, after the speech therapy sessions.

Several speech exercises for stigmatism were integrated in the games. These include exercises in which the child has to say

whole words that have sibilant phonemes, identify sibilant sounds and produce isolated sibilant phonemes.

These games are aimed at children with distinct difficulties and therefore are customizable to meet the specific needs of each child. In order to achieve this, the games are customizable. They include several general and several specific settings that the SLP can set according to the needs of each child. This functionality is extremely important, as children have different difficulties, requiring personalized training, in order to overcome their difficulties more quickly and easily.

The games are controlled by the user's voice, who must perform these speech therapy exercises in order to play them. In this way, we aim to motivate the children with sigmatism on performing the speech exercises integrated in the games.

In order to assess whether the speech exercises are correctly performed, the games use two ASR systems: a sibilant consonant classifier, used to determine if the child produced the expected sibilant consonant, and a keyword spotting system that determines if the child produced the expected word [7].

Speech exercises

Children with sigmatism can have difficulties producing one or more sibilant phonemes. Another common difficulty is on distinguishing these phonemes and using the correct phoneme.

The sibilant phonemes are produced using different places of articulation [5]. [s] and [z] are both alveolar sibilants. The difference between these two sibilants results from the use of the vocal folds: [s] is voiceless and [z] is voiced. On the other hand, [ʃ] and [ʒ] are palato-alveolar sibilants, and [ʃ] is voiceless while [ʒ] is voiced. A frequent mistake made by children is the exchange of sibilant phonemes that are produced with the same place of articulation but different use of vocal folds.

During the design phase of the proposed games, we interviewed SLPs who helped define the speech therapy exercises that are integrated in the games. The selected exercises are already used in traditional speech therapy sessions and include: (1) production of words with sibilant consonants, (2) identification and production of sibilant consonants, (3) minimal pairs exercise, and (4) recognition and distinction between sibilant consonants.

Production of word with sibilant consonants

As the name suggests, this exercise consists of producing a word that has one or more sibilant phonemes. This exercise is widely used by SLPs in therapy sessions who have children with sigmatism so that they can overcome their difficulties in the production of words with sibilant consonants.

Some aspects have to be taken into account when choosing the words for this exercise. The location of the sibilant phoneme(s) in the word is of great importance, as different locations present different production difficulties. Another important aspect to be considered is the structure of the word. Simple words, that is, those where syllables structure is always consonant-vowel, are easier to produce than words with sequences of two or more consonants. Similarly, common words that children are used to hearing in their daily lives, are also easier for children to produce than words they hear sporadically.

Identification and production of sibilant consonants

Another very common therapy exercise is to identify sibilant consonants and then utter them. This exercise, like the previous one, trains the correct production of these consonants. In addition, it also focuses on recognizing them.

This type of exercise presents the same word production difficulty aspects related to the structure of the words as well as the location of the sibilant phoneme in the word.

Minimal pairs

Minimal pairs are words that differ only on a phoneme. That is, these are similar words apart from the addition, substitution or exclusion of a single phonological element, such as in the words sit and seat. Given children often exchange words in minimal pairs, this exercise is frequently used in speech therapy [8].

Recognize and distinguish sibilant consonants

Recognizing and distinguishing sibilant consonants upon hearing them are also exercises performed during therapy sessions. Although children are not training the production of sibilant consonants, it is nevertheless important and challenging for children to carry out this type of exercise.

An exercise to recognize and distinguish sibilant consonants is to let children decide which consonants are the same as when they hear, for example, words. For this, words with different consonants can be used. However, in order to increase the difficulty, SLPs can also use words with the same consonants, but with different characteristics, such as frequency and/or intensity.

Serious games for sigmatism

Two different games have been developed: the **pairs game** and the **maze game**. In this way, when the child is tired of playing a particular game, he/she has the chance to play a different game while continuing to practice the speech exercises that help overcome difficulties in articulating sibilant sounds. On the other hand, not all children like the same kind of games, thus, having two different games will increase the number of children who will enjoy training with the proposed games.

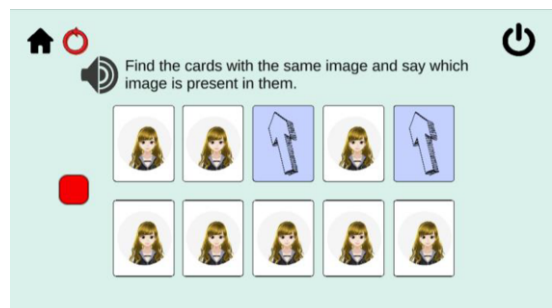


Figure 1- Pairs game.

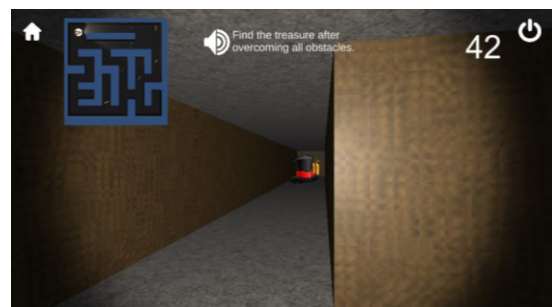


Figure 2- Maze game.

The **pairs game** is based on the well-known matching pairs card game, in which the cards are initially face down and the player has to find pairs of matching cards. Figure 1 shows a trial

in which most cards are face down (cards with the girl's face) and there is a matching pair for the word *seta* (which means arrow in EP). To select a pair, children have to click on the cards they want, but unlike in the original card game, here the child has to complete a speech exercise in order to win pairs of matching cards.

The **maze game** is based on the well-known 3D maze game. Since not all younger children are used to 3D games, and since older children may be used to games with more action, this game was especially designed for older children. As in the original game, here children have to discover and follow the path that leads to the end of the maze, where they will find a treasure chest (figure 2). While navigating in the maze, children have to overcome several obstacles that consist of images representing a particular word. To overcome them, children need to make a speech exercise, which consists of producing the word in the image, slightly lengthening the sibilant consonant.

Since the difficulty in performing the speech exercises depends on the child (more specifically on his/her age and degree of sigmatism), these games have several levels: Each level has a different speech exercise. (The levels are discussed in more detail below.) The pairs game has two modes and five levels. The maze game has three levels. The pairs game basic mode was designed for younger children and the advanced mode was designed for older children. Nonetheless, if the SLP finds appropriate, he/she can use any of the modes with his/her patients independently of their age.

The modes have some slight differences in gameplay, more specifically at the time of selecting the correct matching pair and in what the child has to utter to win the pair of cards. A pair in the basic mode consists of cards with the same image, while in the advanced mode it consists of two cards that represent words that contain the same sibilant sound.

The level of difficulty can be manually selected by the SLP, but it can also be changed automatically. The games include a dynamic difficulty adjustment (DDA) system that chooses the level of difficulty based on the child's performance in the previous trials.

The games give real-time visual and auditory feedback. The most basic auditory feedback consists of different sounds to indicate whether the speech exercises have been well accomplished or not. The most basic visual feedback in the pairs game consists of having the selected pair's background changed to a different color from the pairs previously found (figure 1) and showing which pairs have been won. As for the maze game the visual feedback is given by allowing the child to advance in the maze. (Upon reaching an obstacle, the child cannot move until he/she correctly utters the word. Only when the speech exercises are performed correctly, the child is able to advance in the maze.) A more sophisticated form of feedback implements a system of hints that helps the child perform the exercises. This system of hints is described below.

The instructions of both games are always available. This allows children to listen to the instructions whenever they feel the need. An important characteristic of the games is that the current state of the game is saved when exiting the game. In this way, when the child returns to the game after exiting he/she can resume it from where he/she left off. This is a very useful aspect, according to the SLPs, because during a therapy session time is limited. Having this functionality, allows the child to finish the practicing the exercises after the session has finished. Also, when resuming the game at the point where it was interrupted, the child is not discouraged by having to return to the beginning of the game.

Motivating systems

In order to motivate children to practice often, the proposed games include several tools that help the children perform the speech exercises and that aim to keep the level of difficulty adjusted to the child's performance. Apart from the simpler visual and auditory feedback described above, these tools include a hints system, several difficulty levels and a DDA system. A rewards system is shared by both games, comprising a store of characters and also costumes and items for the characters.

Hints system

The system of hints is intended to assist children with difficulties in accomplishing the games' speech exercises, without immediately offering them the answer. Thus, it forces them to think about what they have done and what is asked of them. In this way, this system only comes into effect if the child fails the exercises more than once. The implementation of this system was discussed with a SLP.

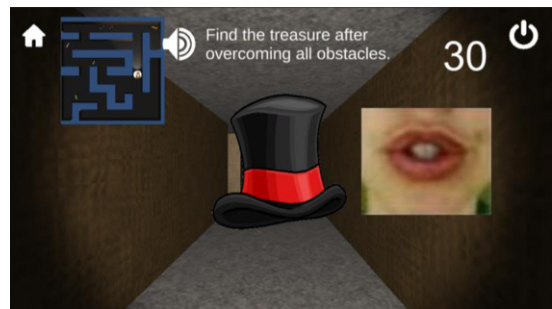


Figure 3- Maze Game - visual cue with an illustration on how to produce the [s] phoneme.

After the child has made several unsuccessful attempts to complete the speech exercise, a visual cue is provided to help him/her. The simpler cue consists of shaking or zooming on the selected cards so that the child pays attention to the images in the cards or in the maze. Also another cue consists of showing the written word under the image. If these cues are not enough, and the child keeps failing the speech exercise, a more direct visual support is provided, which consists of an image of lips exemplifying the movement needed to make the desired sound (figure 3). The lip images were obtained from a speech therapy materials website¹.

In addition, the game also uses audio cues in the basic mode: when the written word appears under the images, the game provides a button for the sound reproduction of the word. When the images with the lips appear, the game will play the word simultaneously.

Difficulty levels

The maze game contains three distinct levels. They differ not only in the size and complexity of the maze as well as in the number of obstacles present in the game. Thus more advanced levels have larger and more complex mazes, and include more obstacles that have to be surpassed in order to reach the treasure chest, which implies that the child will have to complete additional more speech exercises in more advanced levels.

The game of pairs has a more complex system of levels. It comprises five different levels of difficulty that, in some cases, may slightly modify the game play. It should also be mentioned that the number of levels of the two existing game modes are

¹ www.metododasboquinhas.com.br

not the same. The advanced mode has a greater variety of levels than the basic mode.

In the **first level**, the words with sibilants present in the cards are simple and common, that is, words with which the children should be familiar and with a consonant-vowel structure, such as in the words sun and some. In addition to those words, the **second level** includes words that may be less common and whose composition is not always based on a vowel-consonant structure, such as in smile and singer. The pairs in the **third level** consist of minimal pairs. In this level, the child has to select the cards whose words differ in a single phoneme and then the child has to produce the words of both cards, slightly prolonging the sibilant phoneme.

These first three levels are common to both game modes (basic and advanced). The remaining levels are only accessible in the advanced mode. The **fourth level** is intended to practice the auditory identification of sibilant sounds. When the child selects a card, instead of having the card turning upwards showing the image it contains, the game will simply play a sound. In order to win pairs of matching cards, children need to find the pair of cards that contain the same sound, not needing to utter anything for it to be considered correct. The **fifth level** adds some difficulty to the task of the fourth level: children have to choose the cards with the same sound, the same frequency and the same intensity. In this level the same sound can be present in several pairs, but with different characteristics.

Dynamic difficulty adjustment

In order to keep children motivated on playing the games, it is important that children do not become discouraged because the game is too easy for them, nor that they become frustrated due to excessive difficulty. Thus, it is important that the games' difficulty adapts to the child's speech performance.

In order to achieve this adaptation, both games have a DDA system that automatically changes the game level taking into account the child's speech performance in previous trials. The level of difficulty can increase or decrease, depending on the child's speech performance. In order to avoid abrupt level changes, children are required to play at least five trials on the same level before the level is automatically changed.

Since different EP sibilants are produced with different points of articulation and can be voiced or voiceless, the speech performance is measured in terms of:

1. performance for point of articulation and
2. performance for use of vocal folds.

We also have to be aware that younger children, as a general rule, may have greater difficulties in achieving good performances. Thus, the DDA system takes into account the game mode for the pairs game. When the DDA system detects a poor performance in the first two levels of the basic mode, instead of immediately lowering the game level, it adds some pairs of cards with easy words. These consist of words without sibilant consonants.

Validation

Validating this work is crucial to assure that it fulfils its goals and it is adjusted to its target audience. The COVID-19 restrictions imposed during the development of the study have unfortunately hampered attempts to undertake a longitudinal study on the children with sigmatism using the games. Therefore, we validated the games with questionnaires to SLPs.

After contacting several SLPs, we managed to have four SLPs validating the proposed games. These were all female SLPs and they all worked with children.

The validation protocol consisted of:

1. watching a video with a tutorial about the games and their features, and reading a written user manual,
2. exploring the games freely after installing them in their computers, and
3. answering a questionnaire.

The questionnaire included questions about the different modes, the diversity of levels and having levels with different exercises, about the system of hints, about the visual and auditory feedback, the customizable parameters, the DDA system and the games' design. The questionnaire also asked SLPs if the games are suitable to be used at home, and if SLPs would use the game with their patients.

Most questions used a 5-point Likert scale. Some questions included a free text box to allow SLPs to insert suggestions or more details on their answers.

Results

The SLPs who participated in the study did not regularly use serious games created specifically for speech therapy. However, not only would they like to use the proposed games in their therapy sessions, but they would also recommend children to use them at home.

All SLPs considered the graphical aspects of the application adequate to the age of the children that usually attend speech therapy (three to nine years old). Also, all SLPs considered the use of the application adequate, easy or even intuitive, for the children as well as for themselves. These two aspects are extremely important, especially for children, because an intuitive and appealing application motivates them to use it more frequently.

All SLPs gave positive feedback about the customizable parameters, system of rewards, hints, and the existence of several levels of difficulty. In the same way, all SLPs considered that adapting the difficulty dynamically according to the child's performance is a good strategy.

Another element that was assessed was the difficulty of the levels, which should be challenging, but not excessively, to keep the children from becoming demotivated. The SLPs respondents considered the difficulty levels adequate for the children.

The results show that the responding SLPs found that the games, the variety of levels and rewards can potentially motivate children with SSDs on practicing the speech exercises included in the games. Also the feedback and system of cues, can help the child understand how to improve his/her speech performance in these exercises.

Conclusions

Here we proposed two games developed and aimed at children who have difficulties in articulating sibilant sounds. These games integrate speech exercises and are controlled by the child's voice. They have several levels of difficulty and include a dynamic difficulty adjustment system that changes the level of difficulty according to the child's speech performance. The automatic difficulty takes into account the child's accuracy in performing the exercise (vocal and articulatory performance) in each of the sibilant sounds. In this way, the games become more complex and impel the child to improve in order to pass the level.

The proposed games aim to break the monotony of training and aim to motivate children on practicing the speech exercises often. They can be used in face-to-face and online sessions but also at home without adult supervision. The games were

validated with SLPs that found the games appropriate to be used for intensive training at home. In addition the SLPs agreed that the motivational tools may help keep the children engaged in playing the games often, and thus practicing the speech exercises integrated in the games.

Acknowledgements

This work was supported by the Foundation for Science and Technology under the BioVisualSpeech projects (CMUP-ERI / TIC / 0033/2014) and NOVA-LINCS (PEest / UID / CEC / 04516/2019). We thank Helena Alves and Aurora Ribeiro for their support during the design phase of the proposed work, and all SLPs who participated in the validation phase.

References

- [1] Bremner, A. J., Lewkowicz, D. J., & Spence, C. (Eds.), *Multisensory development*. Oxford University Press, 2012.
- [2] McLeod, S. *The international guide to speech acquisition*; Thomson Delmar Learning, 2007.
- [3] Oermann, M. H., *Annual Review of Nursing Education: Strategies for Teaching, Assessment, and Program Planning*, Volume 3, 2005.
- [4] Smith, R. The Long History of Gaming in Military Training. *Simulation & Gaming*. 2010;41(1):6-19.
- [5] Guimarães, I., *Ciência e Arte da Voz Humana*, Escola Superior de Saúde de Alcoitão, 2007.
- [6] Cruz-Ferreira, M. Portuguese (European). In *Handbook of the International Phonetic Association, A guide to the use of the international phonetic alphabet*; Cambridge, University Press, 1999.
- [7] Cavaco S, Guimarães I, Ascensão M, Abad A, Anjos I, Oliveira F, Martins S, Marques N, Eskenazi M, Magalhães J, Grilo M. The BioVisualSpeech Corpus of Words with Sibilants for Speech Therapy Games Development. *Information*. 2020; 11(10):470.
- [8] Fonética e Fonologia: Que diferença? Last visited: 23/12/2019. Url: http://cvc.instituto-camoes.pt/cpp/acessibilidade/capitulo3_3.html.

Address for correspondence

Sofia Martins, str.martins@campus.fct.unl.pt,
Sofia Cavaco, scavaco@fct.unl.pt.