

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/333406072>

A Model for Sibilant Distortion Detection in Children

Conference Paper · November 2018

DOI: 10.1145/3299852.3299863

CITATIONS

2

READS

728

6 authors, including:



Ivo Anjos

Fundação para a Ciência e a Tecnologia

7 PUBLICATIONS 22 CITATIONS

[SEE PROFILE](#)



Margarida Grilo

Escola Superior de Saude do Alcoitão

13 PUBLICATIONS 23 CITATIONS

[SEE PROFILE](#)



Mariana Ascensão

Alcoitão School of Healthy Sciences

9 PUBLICATIONS 10 CITATIONS

[SEE PROFILE](#)



Isabel Cristina Ramos Peixoto Guimarães

Escola Superior de Saude do Alcoitão

116 PUBLICATIONS 770 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



FraLusoPark [View project](#)



Establishment of an Interdisciplinary Clinical Master Program in Rehabilitation Sciences at JUST (JUST – CRS): Erasmus Plus Funded Project [View project](#)

A MODEL FOR SIBILANT DISTORTION DETECTION IN CHILDREN

Ivo Anjos¹, Margarida Grilo², Mariana Ascensão², Isabel Guimarães², João Magalhães¹, Sofia Cavaco¹

¹NOVA LINCS, Department of Computer Science, Faculdade de Ciências e Tecnologia
Universidade NOVA de Lisboa, 2829-516 Caparica, Portugal

²Escola Superior de Saúde do Alcoitão

Rua Conde Barão, Alcoitão, 2649-506 Alcabideche, Portugal

i.anjos@campus.fct.unl.pt, {margarida.grilo, mascensao, iguimaraes}@essa.pt, {jm.magalhaes, scavaco}@fct.unl.pt

ABSTRACT

The distortion of sibilant sounds is a common type of speech sound disorder in European Portuguese speaking children. Speech and language pathologists (SLP) use different types of speech production tasks to assess these distortions. One of these tasks consists of the sustained production of isolated sibilants. Using these sound productions, SLPs usually rely on auditory perceptual evaluation to assess the sibilant distortions. Here we propose to use an isolated sibilant machine learning model to help SLPs assessing these distortions.

Our model uses Mel frequency cepstral coefficients of the isolated sibilant phones and it was trained with data from 145 children. The analysis of the false negatives detected by the model can give insight into whether the child has a sibilant production distortion. We were able to confirm that there exist some relation between the model classification results and the distortion assessment of professional SLPs. Approximately 66% of the distortion cases identified by the model are confirmed by an SLP as having some sort of distortion or are perceived as being the production of a different sound.

Index Terms: machine learning, sibilant sounds, speech sound disorders, sigmatism assessment

1. INTRODUCTION

Many children suffer from speech sound disorders (SSD). As reported by Guimarães *et. al* for data on European Portuguese (EP), 8.8% of preschool-aged children suffer from some type of SSD [1]. These disorders can influence the children's quality of life, as they may affect the children's ability to communicate and may cause embarrassment, shame, frustration, among other negative feelings. In addition, when the SSD is not due to an oral structural problem, it can also have a negative impact on the children's future as it can affect the literacy acquisition and may cause literacy problems [2, 3]. If these problems are not detected and treated early, they may persist or lead to a worsening of the situation depending on the disorder.

Sigmatism is a type of SSD that consists of the distortion of sibilant sounds production. This SSD is very common in European Portuguese (EP) children [4, 5]. There are four different sibilant consonant sounds in EP: [z] as in zebra, [s] as in snake, [ʃ] as the *sh* sound in sheep, and [ʒ] as the *s* sound in Asia. The most usual sibilant mistakes committed by children are distortion errors [1, 2]. Distortion errors typically reflect a slight alteration in the production of a sound (e.g., a slight problem with tongue shape or placement). The resulting productions are in the correct phoneme category but lack phonetic precision or accuracy. It has been suggested that distortions (e.g., dentalized or lateralized [s]) may represent a breakdown in motoric processes [2].

Speech therapy plays an important role to overcome SSDs since it allows detecting the disorders as well as to correct and improve the children's speech. To assess children that may have a SSD that affects the production of sibilant consonants, the SLP analyzes both how the children react to the hearing and the production of the isolated sibilants. Here we propose to use a machine learning approach that uses the isolated sibilants to help the SLP identifying children who may suffer from sigmatism.

Other machine learning models to assess SSDs or detect speech errors exist. These include models to detect vocal fold pathologies [6], paraphasias [7], and models to detect articulation impairments due to the vibration of the vocal folds [8]. Some of these models use MFCCs as input to their models. PEAKS is a system that automatically evaluates voice and speech disorders using Mel frequency cepstral coefficients (MFCCs) [9]. This type of assessment tools have also been used to detect Parkinson's disease [10, 11]

A few systems have been proposed to help detect sigmatism in children, but all previous work has some limitations. Valentini-Botinhao *et al.* have proposed a system to automatically detect sigmatism in children [12]. Even though they have high recognition rates at phone, word and speaker level, and they also discuss their sigmatism detection rates, their model is trained with data from speech therapy students simulating the sound distortions. Benselam *et al.* [13] also pro-

posed a system aimed at detecting sigmatism in the Arabic language. This system is trained with data from adults and uses only one word containing the sibilant [ʃ].

The distortions represented in the data set from Benselam *et al.*'s study are adults distortions. While the distortions in the data set from Valentini-Botinhao *et al.*'s study could have been well simulated, these are not from children's voices, which have quite different characteristics from adult voices. On the contrary, our model was trained with data from children and their natural distortions.

Miodońska *et al.* also use data from children. They developed a system for the classification of sibilant [ʃ] that was trained with data from 60 preschool children (5 and 6 year olds) from Poland [14]. Their model uses MFCCs together with other features. Although they have high accuracy rates with children's voices, they train the model to recognize only one sibilant. Also, although they mention that the results from the model could be used in systems to assess sigmatism, they do not explore how to do it.

We go a step further as we use a richer data set and also discuss how to use the results from our classification model to identify children that may have sigmatism. Our model was trained with data from 145 children from 4 to 11 years old, speaking normally. In this way, the model was trained with real data, that is, with real examples of correct and incorrect child speech productions, and is suited to recognize sibilants from preschool as well older children. In addition, it was trained with samples from all four EP sibilants.

We have previously developed a machine learning model that uses support vector machines (SVM) and MFCCs to classify the four isolated EP sibilants [15]. This model was initially proposed as an automatic classifier for a mobile game for training the EP sibilant consonants. Here we propose further developments of the model that uses its classification results to help the SLP detect cases of sigmatism. These developments use the false negative rate (f_{nr}) information. If a child's sibilants' productions are incorrect, it is possible that the child has sigmatism. Providing this information to the SLP, will help him/her to assess if the child has an SSD.

We were able to find some relation between the f_{nr} of our model, and the incorrect sound productions of children, which were confirmed by a professional SLP (section 5). Approximately 66% of the sibilants that the model identified as incorrect speech productions, were also identified as incorrect by an SLP, either because the SLP identified that the child had sigmatism or because the SLP perceived the sibilant production as a different sound. Thus, using the information provided by the model, SLPs will have more useful information that can help them making the assessments. Note that we are not proposing to use the model to identify types or causes of SSDs. We propose to use it to give information to the SLP that may help him/her to assess if the child has an SSD. To ensure the type of SSD it is necessary to do a clinical assessment by the SLP.

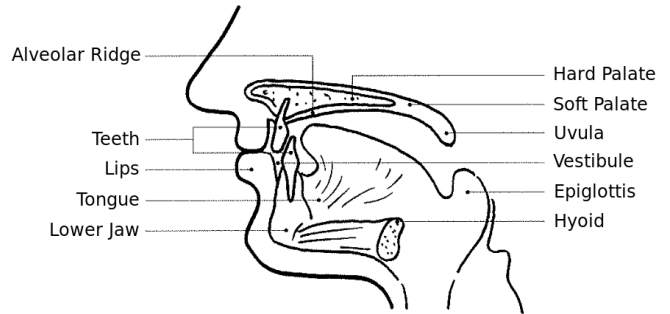


Fig. 1. Main places of articulation in the vocal tract, adapted from [16].

2. SIBILANTS

Different sibilant consonants are produced letting the air flow outwards in the mouth through a narrow channel [16]. As mentioned above, there are four sibilant consonants in EP: [z], [s], [ʒ], and [ʃ]. The sibilants [z], and [s] are produced when the tongue partially touches the alveolar ridge and air passes through the small space between the tongue and the alveolar ridge (figure 1). The sibilants [ʒ], and [ʃ], are produced moving the tongue towards the hard palate and letting air pass through that small formed space.

Sibilants can be voiced, when the vocal folds are used, or voiceless, otherwise. Both [z] and [ʒ] are voiced sibilants, and [s] and [ʃ] are voiceless sibilants.

The most usual sibilant mistakes committed by children are distortion errors, which can be caused by [2]:

- Exchanging the voiced and voiceless sounds. For example, both [s] that is voiceless, and [z] that is voiced, are produced creating a narrow channel in the same location of the mouth. If the child does not use the vocal folds when trying to say “zip”, she will end up saying “sip” instead.
- Exchanging the local of the narrow channel in the vocal tract, which results in producing another sibilant.

3. DATA

In order to train our model we need sound samples for each of the four EP sibilant sounds. We collected data from children in three schools in the Lisbon area. We obtained an informed consent from all parents or legal guardians, and the ethics approval was provided by the ethics committee of *Escola Superior de Saúde do Alcoitão, Santa Casa da Misericórdia de Lisboa*.

We collected both short and long productions of the isolated sibilants (that is, a version that lasts less than a couple of seconds and another that last a few seconds). We have over 1500 sound productions, from 145 children from 4 to 11

Table 1. Age and gender of children participating in the recordings

Age	Girl	Boy	Total	Children with incorrect productions
4	0	1	1	1
5	8	3	11	8
6	8	8	16	3
7	19	9	28	6
8	21	20	41	7
9	23	19	42	4
10	3	2	5	0
11	1	0	1	0
Total	83	62	145	29

years old. From these there were 83 girls and 62 boys (table 1). Most of these sound samples are from children with no SSD but there were also children with SSD participating in the recordings. The last column in table 1 shows the number of children by age with incorrect sibilant productions. Table 2 shows the number of correct and incorrect productions of each sibilant.

The recordings were made in a quiet room at the schools (but where it was possible to hear the noise from the playground and corridors). Only one child at a time was present at the room, along with one SLP and possibly a couple of other adults (who could be SLP graduate students or one researcher from our team). For each sibilant, the SLP exemplified the sound and then asked the child to repeat it. For instance, the SLP could say “Now, say this sound *sss* as long as you can.” The SLP would first ask the child to produce all sibilant sounds, either in the short or long version, and then she asked the child to produce them in the other version. The order of the sibilants varied, and the order of the version (short or long) also varied.

We have eight sound samples from most children (a short and a long production for each sibilant), while from a few children we have more sound productions. Some children were not able to produce some sibilants. Thus, we also have children that lack the production of certain sibilants.

All the sound productions were labeled according to the sibilant that the SLP asked the child to produce, even when the production was incorrect. The labels also include information that indicates if the production is correct or incorrect. If the child produced the sibilant correctly, the sound would be marked with *R* for a correct production, otherwise it would be marked with *W* for an incorrect production. This analysis was not performed by an SLP. For instance, if the SLP asked the child to produce a sibilant s_1 and the child produced another sibilant s_2 , the sound sample is labeled with s_1W .

4. THE POTENTIAL DISTORTION

Table 2. Number of correct and incorrect sound productions for each sibilant

Sound	Correct productions	Incorrect productions
j	276	34
ʒ	257	58
s	278	39
z	265	49

DETECTION ALGORITHM

As mentioned above, we propose to use the information about false negatives detected by our sibilant classification model, to help SLPs detect sigmatism cases. In this section we describe the features used to train our model, and the classification algorithm.

4.1. Feature Extraction

The input features of our classification algorithm are vectors with the first 13 MFCCs of the children’s speech productions. For each child c and sibilant s , we concatenate all sound productions of that sibilant. Then we extract the MFCCs of the concatenated sound. To extract the MFCCs we use a 25 milliseconds window with a 10 milliseconds shift. So, for each child, each sibilant is represented by a matrix of length t , where t depends on the duration of the concatenated sound, and each column contains the first 13 MFCCs. We use the raw MFCCs values in the columns of these matrices, as input to the SVM classifier.

While the concatenated sounds last for a few seconds, we are not using the MFCCs of the whole sounds. From each concatenated sound, we use only a subset of the MFCCs matrix columns. These were selected randomly and are our feature samples that are used as our training samples. We are currently using 11 feature samples to represent each sibilant in the training set. This number was chosen for computing power reasons (more details in section 4.2).

4.2. Algorithm

We use support vector machines (SVMs) with a radial basis function (RBF) kernel to classify our sound samples. In order to find the best parameters for the RBF kernel, we use a grid search, to find the best combination of γ and C values.

To better identify the incorrect productions of the sounds and avoid introducing bias in the learning process, we use a one-child-out approach. This consists of a k -fold approach, where we are careful to have all the data from each child c either in the training or the test set. In other words, since there are recordings from 145 children, we have 145 experiments. In experiment c , we put all data from child c in the test set, and use the remaining children’s productions in the training

set. This allows us to have the most amount of training data possible, without inducing any type of bias.

The training set from each experiment contains only features from the correct productions of the sibilant sounds. In addition, for training we only use the data from children that have produced all sibilant sounds correctly. So, our training sets are composed of feature samples from 116 or 115 children, even though we have recordings from 145 children. For example, let us suppose that child c is one of the 29 children with incorrect productions (rightmost column in table 1). The test set of experiment c contains the feature samples from all sound productions of this child, and the training set contains the data from the productions of the 116 children with no incorrect productions. On the other hand, if child c has no incorrect productions, the training set of this experiment contains the data from 115 children.

Our solution uses four SVMs, one for each sibilant sound. So in reality, we create several models: four models, one for each sound and for each child. In total, we need to train over 500 models, and considering that we use a grid search to find the best parameters for every model, training all models requires a significant computing power.

As explained in section 4.1, for each concatenated sound, we extract 11 feature samples. In total the training set for each model consists of 5104 (or 5060) feature samples, which correspond to 116 (or 115) children \times 11 features \times 4 sibilants. This reduced number of training samples, was chosen due to the time needed to train our models.

To identify the concatenated sounds that are incorrect or contain incorrect portions that may be caused by sigmatism, we use the false negative rate, fnr . For each sibilant, s , the fnr is the percentage of test samples from the concatenated sibilant s that are classified as *not s*:

$$fnr = fn / (fn + tp), \quad (1)$$

where fn , is the number of false negatives, and tp is the number of true positives. Note that the false negatives, that is, the test samples that are classified as *not s*, may be test samples that are misclassified by the model, or test samples that belong to incorrect productions of the sibilant s because of some distortion error (which occur when the child is trying to produce sibilant s but ends up producing another sibilant or another sound).

In each experiment c and for each sibilant s , we have all the concatenated sound samples as the test set. When fnr (in experiment c and for sibilant s) is higher than a predetermined threshold, we consider that sibilant s from child c is worth further inspection by the SLP, and therefore, we mark it as a *possible distortion*. This does not exactly mean that the child has sigmatism. In more general terms, it can be one of the following different scenarios: the signal to noise ratio in the recordings is not high enough for the classifier to correctly identify the sound, the classifier itself might be incorrectly classifying the sibilant, the child has in fact some sort of SSD,

or, for some reason, the child may have produced the sound incorrectly (possibly because he/she did not understand what he/she was supposed to do).

5. RESULTS

In order to assess if the productions classified as incorrect by the algorithm are well identified, we compared the algorithm’s classification with the classification of a professional SLP. Due to the huge number of sound samples, the SLP did not label all sound samples. She analyzed a subset of sound samples that was chosen following these criteria: for each child c , if the concatenated sibilant s was marked as a *possible distortion* by the algorithm, we selected all original sound samples from sibilant s from child c to be analyzed by the SLP.

This process is done for all sibilants of every child. In addition, in order to avoid any type of bias due to only listening to incorrect sound productions, we added around 80% more sound samples. For these we chose samples that were not marked as distortions. This way, the SLP listened to a combination of correct and incorrect sound samples.

We compared the classifications given by the SLP to the classifications obtained by the algorithm with different fnr threshold values. The threshold values were selected by running several experiments and analyzing which values would give us the necessary amount of sound samples to be classified by the SLP and give us confidence on the results. We kept the fnr threshold above 50%, as going this low was giving us a huge increase in the number of sound samples to be classified by the SLP. Above 80% the fnr was too high and was not identifying all the incorrect productions correctly. Thus, the fnr threshold values used in this study were 60%, 70%, and 80%. The SLP classified a total of 245 samples.

For each sound sample analyzed by the SLP, we compared the classification given by the SLP to the algorithm’s classification. Since the algorithm uses the concatenated sounds, for each sibilant s marked as *possible distortion*, the algorithm also marked all sound samples from that child and sibilant s as incorrect.

Figure 2 shows the comparison of the SLP’s classifications with the model’s classifications for the three fnr threshold values. The solid line represents the percentage of sound samples that were classified as incorrect by both the SLP and the algorithm. The dashed line represents the percentage of occurrences that were classified as correct by the algorithm and as incorrect by the SLP.

With a fnr threshold of 80%, when our algorithm detects any type of anomaly with the sound samples, we have a percentage of agreement by the SLP of at least 61.6%, that increases to 66.0% when using a threshold value of 70%, and then, slightly decreases to 65.7% when using 60%. As mentioned above, for the concatenated sounds marked as *possible distortion*, the algorithm also marked the original sound sam-

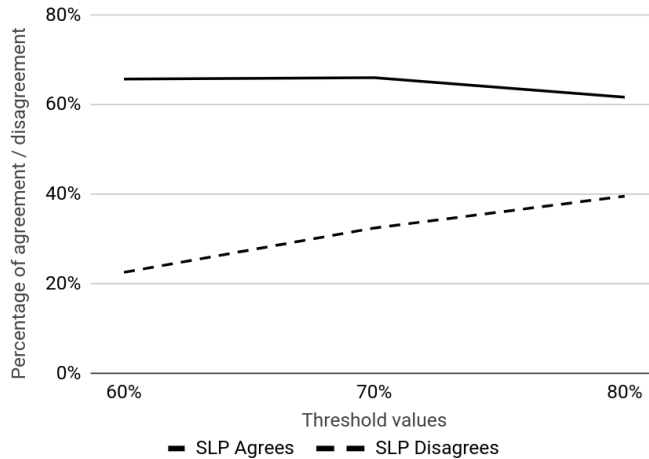


Fig. 2. Results chart. (—) Percentage of agreement between the SLP and the algorithm for sibilants marked as *possible distortion*. (- -) Percentage of disagreement for sibilants classified as correct by the algorithm.

ples as incorrect. Note that a child can produce the sibilant correctly some times while failing other attempts. If the algorithm had analyzed the sound samples from the same sibilant separately, the results could have been different and the percentage of agreement between the SLP would probably have been higher. Nonetheless, we opted to use the concatenated sounds because for the purpose of this work, it does not matter if the child produces the sibilant correctly some times while failing other attempts, that is, the classification of individual samples is not important. The important result is to detect the cases with anomalies when inspecting all productions from the same child.

An important value that must be analyzed is the number of cases for which the SLP detected a distortion that was not detected by the algorithm. This number should be as low as possible, so that not many children who can benefit from further analysis from a SLP go undetected by the algorithm. Figure 2 shows that with a *fnr* threshold of 80%, the algorithm does not detect 39.5% of the sound samples identified as incorrect by the SLP. This value decreases to 22.5% when using a 60% threshold.

While that result may seem high for what would be desirable, when we analyze the results in terms of the number of children, we see that the results are actually quite good. The SLP analyzed data from 89 children. From those, there were sound productions from 18 children (20.2%) that were identified as incorrect by the SLP but not by the algorithm. The percentage of children with incorrect productions missed by the algorithm varies from 6.5% to 14.3% depending on the sibilant. Table 3 shows the results for each sibilant. Note that the last line in the table is not the sum of the lines above because the intersection of the different sets is not empty.

A factor that must be taken into account is that the per-

Table 3. Data analyzed by the SLP and results with a 60% *fnr* threshold

Sibilant	Total number of children	Children missed by algorithm	Percentage of missed children
ʃ	31	2	6.5%
ʒ	42	8	19.0%
s	29	3	10.3%
z	42	6	14.3%
ʃ, ʒ, s, z	89	18	20.2%

centage of children with incorrect speech productions in the data set analyzed by the SLP (49.4%) is much higher than the percentage of children with SSD in the country. Thus, these results are actually very positive.

6. CONCLUSION AND FUTURE WORK

In this paper, we proposed an algorithm that given recordings of the isolated sibilants from children, signals those children who can possibly benefit from further analysis from a SLP to assess if the children have sigmatism. The algorithm inspects the false negative rate obtained by an isolated EP sibilant classifier trained with the MFCCs of the children’s sibilant productions. The novelty of this work is the use of the *fnr* given by a machine learning algorithm to detect cases of potential sigmatism.

The comparison of the algorithm’s results with the classification from a professional SLP shows that the *fnr* is a good measure for identifying productions with anomalies. The SLP agrees to the classification given to 66.0% of the samples marked as incorrect by the algorithm when using a *fnr* threshold of 70%, and 65.7% when using a 60% threshold.

The results show that a *fnr* threshold of 60% gives a good balance between the number of sibilant productions marked as incorrect and the number of correct productions that are not detected by the algorithm. With this threshold, 22.5% of the samples identified as incorrect by the SLP are not detected by the algorithm. While this may seem a high value, the number of children missed by the algorithm varies from 2 out of 31 to 8 out of 42 depending on the sibilant.

These results are very promising but can still be improved. We still have some sound samples that are not detected by our algorithm and were marked as incorrect by the SLP, and also some that were marked as incorrect but the SLP considered them correct. We are exploring using other features that will help classifying correctly these cases.

7. ACKNOWLEDGEMENTS

This work was supported by the Portuguese Foundation for Science and Technology under projects BioVisualSpeech

(CMUP-ERI/TIC/0033/2014) and NOVA-LINCS (PEest/UID/CEC/04516/2013).

We would like to thank all the 3rd and 4th year SLP students from *Escola Superior de Saúde do Alcoitão* who collaborated in the data collection task. Finally, we would like to thank the schools from *Agrupamento de Escolas de Almeida Garrett*, and all the children who participated in the recordings.

8. REFERENCES

- [1] I. Guimarães, C. Birrento, C. Figueiredo, and C. Flores, *Teste de articulação verbal*, Ocina Didáctica, Lisboa, Portugal, 2014.
- [2] J. Preston and M. L. Edwards, “Phonological awareness and types of sound errors in preschoolers with speech sound disorders,” *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 1, pp. 44–60, 2010.
- [3] L. Nathan, J. Stackhouse, N. Goulandris, and M. J. Snowling, “The development of early literacy skills among children with speech difficulties: A test of the critical age hypothesis,” *Journal of Speech, Language, and Hearing Research*, vol. 47, pp. 377–391, 2004.
- [4] A. C. Figueiredo, “Análise acústica dos fonemas /s, z/ produzidos por crianças com desempenho articulatorio alterado,” 2017.
- [5] M. Rua, “Caraterização do desempenho articulatorio e oromotor de crianças com alterações da fala,” 2015.
- [6] R. J. Moran, R. B. Reilly, P. de Chazal, and P. D. Lacy, “Telephony-based voice pathology assessment using automated speech analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 468–477, 2006.
- [7] Duc Le, Keli Licata, and Emily Mower Provost, “Automatic paraphasia detection from aphasic speech: A preliminary study,” in *Proceedings of Interspeech*, 2017.
- [8] J. C. Vásquez-Correa, J. R. Orozco-Arroyave, and E. Nöth, “Convolutional neural network to model articulation impairments in patients with Parkinson’s disease,” in *Proceedings of Interspeech*, 2017.
- [9] A. Maier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster, and E. Nöth, “PEAKS – a system for the automatic evaluation of voice and speech disorders,” *Speech Communication*, vol. 51, no. 5, pp. 425–437, 2009.
- [10] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, “Novel speech signal processing algorithms for high-accuracy classification of Parkinson’s disease,” *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1264–1271, 2012.
- [11] A. Pompili, A. Abad, P. Romano, I. P. Martins, R. Cardoso, H. Santos, J. Carvalho, I. Guimarães, and J. J. Ferreira, “Automatic detection of parkinson’s disease: An experimental analysis of common speech production tasks used for diagnosis,” in *Text, Speech, and Dialogue*, Kamil Ekštejn and Václav Matoušek, Eds., Cham, 2017, pp. 411–419, Springer International Publishing.
- [12] C. Valentini-Botinhao, S. Degenkolb-Weyers, A. Maier, E. Nöth, U. Eysholdt, and T. Bocklet, “Automatic detection of sigmatism in children,” in *Proceedings of the International Workshop on Child Computer Interaction*, 2012.
- [13] Z. A. Benselama, M. Guerti, and M. A. Bencherif, “Arabic speech pathology therapy computer aided system,” *Journal of Computer Science*, vol. 3, no. 9, pp. 685–692, 2007.
- [14] Z. Miodońska, M. Krecichwost, and A. Szymańska, “Computer-aided evaluation of sibilants in preschool children sigmatism diagnosis,” in *Information Technologies in Medicine*, E. Pietka, P. Badura, J. Kawa, and W. Wiclawek, Eds. 2016, pp. 367–376, Springer International Publishing.
- [15] I. Anjos, M. Grilo, M. Ascensão, I. Guimarães, J. Magalhães, and S. Cavaco, “A serious mobile game with visual feedback for training sibilant consonants,” in *Advances in Computer Entertainment Technology*, A. D. Cheok, M. Inami, and T. Romão, Eds. 2018, pp. 430–450, Springer International Publishing.
- [16] I. Guimarães, *Ciência e Arte da Voz Humana*, Escola Superior de Saúde de Alcoitão, 2007.