Acoustic correlates of speech intelligibility: the usability of the eGeMAPS feature set for atypical speech

Wei Xue[1], Catia Cucchiarini[1], Roeland van Hout[2], Helmer Strik[1,2]

[1]Centre for Language and Speech Technology, Radboud University, Nijmegen, The Netherlands
[2]Centre for Language Studies, Radboud University, Nijmegen, The Netherlands

`{w.xue, c.cucchiarini, r.vanhout, w.strik}@let.ru.nl`

Abstract

Although speech intelligibility has been studied in different fields such as speech pathology, language learning, psycholinguistics, and speech synthesis, it is still unclear which concrete speech features most impact intelligibility. Commonly used subjective measures of speech intelligibility based on labour-intensive human ratings are time-consuming and expensive, so objective procedures based on automatically calculated features are needed. In this paper, we investigate possible correlations between a set of objective features and speech intelligibility. Specifically, we study the usability of acoustic features in the eGeMAPS feature set for predicting phoneme intelligibility by using stepwise linear multiple regression analysis. The results showed that the acoustic features are potentially usable for predicting intelligibility. This finding may help to boost the development of automatic procedures to measure speech intelligibility with the underlying relevant acoustic phonetic characteristics. Our analysis also covers the comparison between two speech types (dysarthric and normal), and between two different types of speech material (isolated words and running text). Finally, we discuss possible avenues for future research on speech intelligibility and implications for clinical practice.

Index Terms: speech intelligibility, dysarthric speech, eGeMAPS, acoustic features, objective measures of intelligibility

1. Introduction

Speech intelligibility is an important construct that has been studied in different fields like speech pathology [1], second language (L2) pronunciation [2], speech synthesis evaluation [3-4], speech perception [5-6], and telecommunication, but from different perspectives. For instance, in telecommunication, speech intelligibility is defined in relation to the lossless of the transmission channel, while speech pathology mainly focuses on speaker-related aspects such as the presence of a speech disorder. Hence, the approaches to measuring speech intelligibility are different as well. Telecommunication studies use measures such as the Speech Transmission Index [7], the diagnostic rhyme test [8] and the modified rhyme test [9]. Speech pathological research uses tests such as the Dutch Intelligibility Assessment (DIA) [10] and the Sentence Intelligibility Test [11]. In this paper, we are interested in exploring the relation between acoustic features and the intelligibility of speech in which the intelligibility may be reduced by speaker-related aspects such as dysarthria.

A common definition of speech intelligibility in the clinical practice of speech therapy was proposed by Hustad [1] "Intelligibility refers to how well a speaker's acoustic signal can be accurately recovered by a listener". Munro and Derwing suggested a similar definition in L2 pronunciation research "the extent to which listeners' perceptions match speakers' intentions" [12, pp. 14]. According to these definitions, speech intelligibility cannot "be evaluated without some sort of reference to listener data" [12, pp. 13]. As a result, measuring speech intelligibility has conventionally resorted to human listeners by asking them to provide ratings of intelligibility in different ways [13-15]. Common procedures are to ask listeners to express scalar judgments on the degree of intelligibility of speech samples by using such as an equal-appearing interval scales like the Likert scale [15], or by using a horizontal line, on which a point is placed to indicate intelligibility, like a visual analogue scale (VAS) [16]. Moreover, transcribing verbatim what they listeners hear [2, 17-18] and indicating how well individual phones in isolated words were realized [19-20] were also explored.

In order to alleviate the effect of subjectivity in the above methods, ratings are usually collected from multiple listeners and then averaged for further analyses. Besides, reliability measures are also needed. Research has shown that these operations can help obtain reliable ratings [21], and in fact, have been widely used in research and clinical practice. However, these operations are generally time-consuming and costly and the need for multiple raters makes this practice even more laborious and expensive. In addition, while it may be feasible to apply these rating procedures in a research context, they are still problematic in clinical practice, where easy-to-use tools are strongly preferred.

For these reasons, there is a need for valid procedures to obtain objective measures of intelligibility in an automated way that do not rely on intensive human efforts. Several researchers have employed Automatic Speech Recognition (ASR)-based algorithms and ASR-free features to obtain automatic measures of pathological speech quality, which have been shown to be strongly correlated with human-based measures of speech intelligibility [22-26]. In the language learning domain, such objective and automated approaches to measuring speech intelligibility have not been undertaken, although the need is felt there as well [27]. So far, it is not clear how exactly these ASR-based and ASR-free measures are related to properties of pathological speech that can be addressed in therapy, and whether these measures could be used to develop easy-to-use tools for clinical practice.

In this paper, we investigate the usability of a new set of acoustic features called the extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPS) [28], which was recently developed for research in different areas of speech analysis. Specifically, we apply it to prediction of speech intelligibility. The advantage of using this feature set is that it consists of a

standardized, limited set of features, which were chosen based on their demonstrated theoretical relevance, their potential to distinguish important aspects of speech production and their ease in automatic calculation. As the authors suggest, the idea is "to provide a common baseline for evaluation of future research and eliminate differences caused by varying parameter sets or even different implementations of the same parameters". The implementation is realized through the open-source openSMILE toolkit [29], which guarantees standardized calculation of all parameters. In addition, we also included speech rate-related features to investigate if these are complementary to the eGeMAPS feature set for predicting speech intelligibility.

The question we address in this paper is to what extent these automatically calculated features are related to intelligibility scores of pathological and normal speech. If this relation is strong, then it might be also worthwhile to investigate whether these features can be employed to measure intelligibility in other types of atypical speech like L2 speech and to provide the possibility of developing an easy-to-use tool for clinical practice.

An important question in this connection is which intelligibility scores should be taken as the point of reference. As explained above, different types of human ratings have been employed in literature, and these often vary with respect to their degree of detail [21]. In addition, not all sorts of human ratings are available for all speech databases. In other words, the choice of the speech database usually also determines which human ratings are to be used, because they are the only ones available in that specific speech corpus. In this study, we chose to use the COPAS database [30] because it has many advantages, as will be explained in Section 2.1. In this database, the intelligibility scores of a speaker were calculated based on phoneme level ratings and then averaged at speaker level. These intelligibility scores will be taken as the point of reference to evaluate the automatically predicted scores based on eGeMAPS acoustic features.

The rest of the paper is organized as follows. We first describe the database and our methods in Section 2. The experimental results are presented in Section 3, while their implications are discussed in Section 4 and the conclusions are provided in Section 5.

2. Method

2.1. Speakers and speech material

The speech materials used in the study were selected from the COPAS database which was collected and used to develop a reliable ASR-based speech assessment tool for pathological speech within the framework of the SPACE project. This speech database contains recordings, as shown in the COPAS document [30], collected from 197 pathological speakers with speech disorders such as dysarthria, cleft, articulation disorders, voice disorder, laryngectomy and glossectomy and from a control group of 122 normal speakers whose speech is not disordered. The speech materials include not only isolated words, but also isolated sentences and short passages. The word reading task used in this paper is the Dutch Intelligibility Assessment (DIA) [31] with isolated-word material which contains 35 versions of 50 consonant-vowel-consonant (CVC) words and pseudowords organized in three subsets. We also use an additional phonetically balanced reading text known as Text Marloes (TM) [32].

The subjects were selected on the basis of participation in both the DIA and TM tasks. There were 20 female and 29 male dysarthric speakers, and 48 female and 33 male normal speakers, giving a total of 49 dysarthric and 81 normal speakers.

2.2. Intelligibility ratings

The DIA task consists of three subsets containing different words designed for assessing specific phonemes [30]. The phoneme intelligibility on each subset is achieved by asking the therapist whether the target phoneme was correct or not. Then the overall phoneme intelligibility of each speaker was calculated by averaging the percentages of correctly perceived target phonemes in these three subsets and used as the intelligibility score on both the DIA and TM tasks.

The frequency plot of the total phoneme intelligibility of the dysarthric and normal speakers is shown in Figure 1. As can be seen in Figure 1, the normal speakers and the dysarthric speakers showed an overlap in phoneme intelligibility from 82 to 100, with most dysarthric speakers having lower scores than 82. The dysarthric speakers showed more variation in phoneme intelligibility than the normal speakers.

2.3. Automatic acoustic measures

In order to investigate possible acoustic correlates of speech intelligibility, we chose the eGeMAPS feature set because of its wide coverage of standardized relevant acoustic features that can be measured automatically. There are many acoustic features that might be correlated to speech intelligibility such as measures of voice (loudness, fundamental frequency (F0), jitter) and articulation (Mel-frequency cepstral coefficient 1-4, vowel formant frequencies). More detailed descriptions can be found in [28]. Specifically, the selected features were calculated for all the speech recordings using the default configuration file in openSMILE [29]. In addition, a Praat script [33] was used to calculate a set of speech rate-related features such as number of syllables, number of pauses, phonation time, speech rate, articulation rate and average syllable duration by automatically detecting syllable nuclei without using segmentation [34]. This complementary feature set may identify additional parameters related to intelligibility. All of the features in all subsets were averaged for each speaker in the regression analyses in predicting intelligibility.

2.4. Regression analysis

In order to investigate the correlation between acoustic features and speech intelligibility, a stepwise linear multiple regression (SLMR) algorithm [35] (p in 0.05, p out = .10) was applied to predicting phoneme intelligibility using the eGeMAPS features. This process was denoted as SLMR_1. Besides, we applied the second round of SLMR (SLMR_2) on the combination of eGeMAPS features and the speech rate-related features, to predicting the residual values of SLMR_1. This was aimed to investigate whether speech rate-related features can complement the eGeMAPS features for better predicting the intelligibility scores. For each of the two speaker groups, this procedure was applied to the whole DIA task and TM separately. For DIA, the experiments were also carried out for three subsets separately.

All of the features were normalized. The stepwise linear multiple regression was implemented by using the ols_step_both_p function in the R olsrr package.
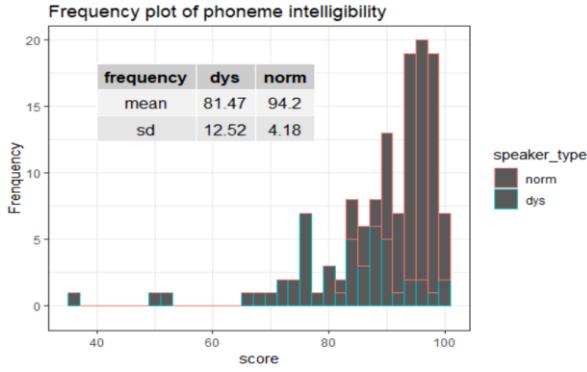
Figure 1: *Frequency plot of phoneme intelligibility of dysarthric (dys) and normal speakers (norm) with a frequency table.*

3. Results

3.1. Correlation results

To explore the correlation between acoustic features and speech intelligibility, we first calculated the correlations between phoneme intelligibility and the eGeMAPS features of the DIA and TM tasks, respectively.

Table 1: *Correlation ranges of the ten highest correlations between phoneme intelligibility and eGeMAPS features for DIA and TM tasks and the overlaps between each two speaker groups (dys.: dysarthric, norm.: normal, and comb.: combined).*

Correlation	DIA	TM
Range dysarthric	0.294-0.492	0.266-0.395
Range normal	0.311-0.409	0.334-0.404
Range combined	0.210-0.423	0.313-0.528
Overlap dys. – norm.	0	2 features
Overlap comb. – dys.	3 features	3 features
Overlap comb. – norm.	1 feature	0

Table 1 shows the range of the ten highest correlations between phoneme intelligibility and eGeMAPS features and the overlaps of them between each two pairs of speaker groups. The highest correlation on the DIA task was found for dysarthric speech. We also found that only a small number of features in the ten highest correlated features were shared between each two pairs of speaker groups.

3.2. SLMR results on the DIA task and Text Marloes

The residual and fitted value plots of SLMR_1 for the dysarthric and normal speech on the DIA and TM tasks are shown in Figure 2. Normal speakers showed similar patterns on both tasks, while larger residual values were found for dysarthric speakers on TM.

Table 2 shows the multiple R-squared scores, which represent the proportion of the variance of the dependent variable that could be explained by the explanatory variables(s), of the final model in SLMR_1 on the whole DIA and TM tasks. The highest score was achieved for dysarthric speakers on DIA. The multiple R-squared score for the combined speakers on TM was larger than that of each separate speaker type.
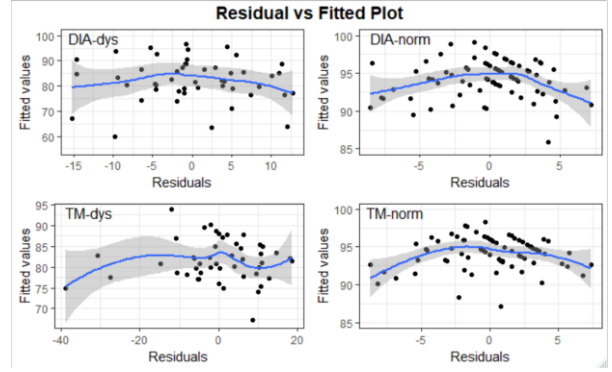

Figure 2: *Residuals and fitted values plot of SLMR_1 for dysarthric (dys) and normal speech (norm) on the DIA and TM tasks.*

The selected features of the final model in SLMR_1 are shown in Table 3. The largest number of selected features was achieved for dysarthric speakers on DIA. The overlap between any two speaker groups was very small regardless of the type of speech material.

Table 2: *The multiple R-squared scores of the final model in SLMR_1 for three groups of speakers (dysarthric, normal and combined speakers) on DIA (whole and three subsets A, B and C) and TM (higher is better).*

Multiple R-squared		Dysarthric	Normal	Combined
DIA	Subset A	0.6238	0.3712	0.2880
	Subset B	0.5011	0.2828	0.2669
	Subset C	0.7565	0.2821	0.6029
	Whole	0.6823	0.3297	0.3539
TM		0.1559	0.3318	0.4309

Besides, we found that speech rate-related features were selected as predictors of the final model in SLMR_2 for both dysarthric and normal speakers on TM, but not on DIA. So, we applied SLMR with the combined feature set of eGeMAPS features and speech rate-related features for directly predicting the phoneme intelligibility on TM. The multiple R-squared scores of the resultant model were 0.3239, 0.4615 and 0.4714 for dysarthric, normal and combined speakers, respectively. They were all higher than the scores in the models without speech rate-related features.

Table 3: *The number of selected features of the final models in SLMR_1 for dysarthric, normal and combined speakers on DIA (whole and three subsets A, B and C) and TM tasks plus the overlaps between each two of three speaker groups (o_dys-norm, o_comb-dys, and o_comb-norm).*

Selected features	DIA				TM
	A	B	C	Whole	
Dysarthric	4	5	9	7	1
Normal	4	5	3	4	3
Combine	5	4	14	4	5
o_dys-norm	0	0	0	1	0
o_comb_dys	2	1	3	1	0
o_comb_norm	0	1	1	0	0

3.3. SLMR results on subsets of the DIA task

The results of SLMR_1 on each subset of the DIA task are shown in Table 2. The multiple R-squared scores of dysarthric speakers were all higher than those of normal speakers and showed a larger variation across three subsets. The scores for the combined speakers were smaller than those for each speaker type except in subset C.

4. Discussion

The present study was aimed at investigating the usability of the eGeMAPS feature set for predicting phoneme intelligibility. The medium-sized correlations between the features and phoneme intelligibility evidenced by the results in Table 1 but in particular the higher values of the multiple R-squared scores of Table 2 suggest that, in principle, these features could be employed to derive objective and automatically calculated measures of phoneme intelligibility for normal and dysarthric speech.

When comparing the results for the two types of speech, we see higher correlations for dysarthric speech than for normal speech in the case of isolated words. As shown in Table 2, the multiple R-squared scores for the dysarthric and normal speech on the DIA task are 0.6823 and 0.3297, respectively. In addition, we tried a fairer comparison by only considering the dysarthric speakers (totally 28) whose phoneme intelligibility was between 82 to 100, which was the same range as that of the normal speakers. In this case, the multiple R-squared score was still higher than that of the normal speakers (0.5477 vs 0.3297). This suggests that on isolated words, the eGeMAPS feature set is effective for predicting phoneme intelligibility in dysarthric speech, but probably not so effective in normal speech. It is conceivable that phoneme intelligibility in these two types of speech relies on different features and that further research is needed to gain more insights into these features.

As for the two different types of speech material, we can see a generally higher correlation for the isolated words than for the running text by a large margin (0.6823 vs. 0.1559), as shown in Table 2. From Figure 2, we can also see that the model for dysarthric speech achieves the largest residual values on TM. These results are not surprising since the applied phoneme intelligibility scores, which were calculated based only on isolated words, do not generalize to running text. It could be the case that phoneme intelligibility only reveals one aspect of general speech intelligibility for dysarthric speech and that additional measures are required to get a more comprehensive picture.

For normal speech, the variation of the correlation results between these two materials as shown in Table 2 is small (0.3297 vs. 0.3318). This indicates that the representation of speech intelligibility in normal speech does not vary much across different materials, as opposed to dysarthric speech.

The results of separate analyses for the three subsets of isolated words which addressed initial, central and final phonemes provided further evidence of the substantial differences between normal speech and dysarthric speech. Specifically, we found that the central phonemes (vowels) play the most important role for dysarthric speech with a correlation score of 0.6238, while for normal speech, the initial consonants contribute most with a score of 0.3712. This difference between the two speech types is also supported by the results in Table 1 and Table 3, where the overlap between the features appears to be very small.

After incorporating additional speech rate-related features, the multiple R-squared scores increased from 0.1559 to 0.3239 for dysarthric speech, and from 0.3318 to 0.4615 for normal speech, in the SLMR experiments on TM. This reveals that in the case of running text, speech rate is an important explanatory factor which could be employed to complement the eGeMAPS feature set, while this does not seem to be the case for isolated words.

On the one hand, these results might be seen as an indication that the eGeMAPS feature set does not contain features that are general predictors of phoneme intelligibility. On the other hand, it seems too early to draw such a conclusion. It is important to bear in mind that the measure of speech intelligibility adopted in the COPAS database is a very narrow measure of phoneme intelligibility. Before drawing conclusions about the usability of the eGeMAPS feature set, it is necessary to conduct further research with other measures of intelligibility at a higher level such as word and/or sentence.

This study has shown, once again, that speech intelligibility is a complex construct and that further research is needed to get a better understanding of both the human-generated measures of intelligibility, as well as their more objective acoustic correlates. We have seen that it is important to include different types of speech materials, to explore the substantial differences between pathological and normal speech in more detail, and to employ various rating procedures (i.e., Likert, VAS, and orthographic transcriptions). The important insights that derive from this more comprehensive research will not be limited to the clinical domain, but may help us analyze speech intelligibility in the field of L2 pronunciation. The performance of the relations between acoustic features and speech intelligibility of L2 learners' speech may be different from the results in this paper, but they may also reveal interesting overlaps and generalizations. Again, as we mentioned above, further exploring how the acoustic features are correlated with speech intelligibility is needed and may help researchers to understand the underlying processes and mechanisms that affect speech intelligibility.

5. Conclusions

Our experimental results showed moderate to medium correlations between phoneme intelligibility scores and the acoustic features in the eGeMAPS feature set, which seems promising for automatic speech intelligibility prediction. Our analyses also revealed important differences between dysarthric speech and normal speech, and between different types speech material (isolated words and running text). These results indicate new avenues for future research that are likely to benefit both the clinical and the language learning domain.

6. Acknowledgements

7. References

[1] K. C. Hustad, "The relationship between listener comprehension and intelligibility scores for speakers with dysarthria," in *Journal of Speech, Language, and Hearing Research*, vol. 51, pp. 562–573, 2008.

[2] M. J. Munro and T. M. Derwing, "Foreign Accent, comprehensibility, and intelligibility in the speech of second language learners," *Language Learning*, vol. 45, no. 1, pp. 73–97, 1995.

[3] C. Benoit, M. Grice, and V. Hazan, "The SUS test: A method for the assessment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences," *Speech Communication*, vol. 18, pp. 381–392, 1996.

[4] D. Gibbon, R. Moore, and R. Winski, Handbook of standards and resources for spoken language systems. *Berlin: Mouton de Gruyter*, 1997.

[5] A. Cutler, M. L. Garcia Lecumberri, and M. Cooke, "Consonant identification in noise by native and non-native listeners: Effects of local context," *Journal of the Acoustical Society of America*, vol. 124, no. 2, pp. 1264–1268, 2008.

[6] M. Cooke, C. Mayo, C. Valentini-Botinhao, Y. Stylianou, B. Sauert, and Y. Tang, "Evaluating the intelligibility benefit of speech modifications in known noise conditions," *Speech Communication*, vol. 55, no. 4, pp. 572–585, 2013.

[7] H. J. M. Steeneken and T. Houtgast, "A physical method for measuring speech‐transmission quality," *The Journal of the Acoustical Society of America*, vol. 67, no. 1, pp. 318-326, 1980.

[8] W. D. Voiers, A. D. Sharpley, and C. J. Hehmsoth, *Research on diagnostic evaluation of speech intelligibility*. No. TRACOR-T73-AU-9010-U. TRACOR INC AUSTIN TX, 1973.

[9] A. House, C. Williams, M. Hecker and K. Kryter, "Articulation testing methods: Consonantal differentiation with a closed response set," *Journal of the Acoustical Society of America*, vol. 37, pp. 158-166, 1965.

[10] C. Middag, J. P. Martens, G. Van Nuffelen, and M. de Bodt, "Dia: a tool for objective intelligibility assessment of pathological speech," In *6th International workshop on Models and Analysis of Vocal Emissions for Biomedical Applications*, pp. 165-167. Firenze University Press, 2009.

[11] K. Yorkston, D. R. Beukelman, and R. Tice, Sentence intelligibility test [Measurement instrument]. Lincoln, NE: Tice Technologies, 1996.

[12] M. J. Munro, and T. M. Derwing, "A prospectus for pronunciation research in the 21st century: A point of view," *Journal of Second Language Pronunciation* vol. 1, no. 1, pp. 11-42, 2015.

[13] S. D. S. Barreto and K. Z. Ortiz, "Intelligibility measurements in speech disorders: a critical review of the literature," *Pró-Fono Revista de Atualização Científica*, vol. 20, no. 3, pp. 201–206, 2008.

[14] N. Miller, "Measuring up to speech intelligibility," *International Journal of Language & Communication Disorders*, vol. 48, no. 6, pp. 601–612, 2013.

[15] K. M. Yorkston and D. R. Beukelman, "A comparison of techniques for measuring intelligibility of dysarthric speech," *Journal of Communication Disorders*, vol. 11, no. 6, pp. 499–512, 1978.

[16] C. Finizia, J. Lindstrom, and H. Dotevall, "Intelligibility and perceptual ratings after treatment for laryngeal cancer: laryngectomy versus radiotherapy," *Laryngoscope*, vol. 108, no. 1, pp. 138–143, 1998.

[17] K. C. Hustad, "Estimating the intelligibility of speakers with dysarthria," *Folia Phoniatrica et Logopaedica*, vol. 58, no. 3, pp. 217–228, 2006.

[18] J. S. Laures and G. Weismer, "The effects of a flattened fundamental frequency on intelligibility at the sentence level," *Journal of Speech Language and Hearing Research*, vol. 42, no. 5, pp. 1148–1156, 1999.

[19] G. van Nuffelen, C. Middag, M. de Bodt, and J.-P. Martens, "Speech technology-based assessment of phoneme intelligibility in dysarthria," *International Journal of Language & Communication Disorders*, vol. 44, no. 5, pp. 716–730, 2009.

[20] C. Middag, J.-P. Martens, G. van Nuffelen, and M. de Bodt, "Automated intelligibility assessment of pathological speech using phonological features," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, pp. 1–9, 2009.

[21] M. Ganzeboom, M. Bakker, C. Cucchiarini, and H. Strik, "Intelligibility of disordered speech: Global and detailed scores," in *Proc. INTERSPEECH*, pp. 2503–2507, 2016.

[22] M. Schuster, A. Maier, T. Haderlein, E. Nkenke, U. Wohlleben, F. Rosanowski, U. Eysholdt, and E. Noeth, "Evaluation of speech intelligibility for children with cleft lip and palate by means of automatic speech recognition," *International Journal of Pediatric Otorhinolaryngology*, vol. 70, no. 10, pp. 1741–1747, 2006.

[23] C. Middag, J.-P. Martens, G. van Nuffelen, and M. de Bodt, "Automated intelligibility assessment of pathological speech using phonological features," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, pp. 1–9, 2009.

[24] V. Berisha, R. Utianski, and J. Liss, "Towards a clinical tool for automatic intelligibility assessment," in *2013 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, May 26–31, Vancouver, BC, Canada, Proceedings, pp. 2825–2828, 2013.

[25] T. Pellegrini, L. Fontan, J. Mauclair, J. Farinas, C. Alazard-Guiu, M. Robert, and P. Gatignol, "Automatic assessment of speech capability loss in disordered speech," *ACM Transactions on Accessible Computing*, vol. 6, no. 3, pp. 8:1–8:14, 2015.

[26] M. J. Kim, Y. Kim, and H. Kim, "Automatic intelligibility assessment of dysarthric speech using phonologically-structured sparse linear model," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 4, pp. 694–704, 2015.

[27] O'Brien, M. Grantham, T. M. Derwing, C. Cucchiarini, D. M. Hardison, H. Mixdorff, R. I. Thomson, H. Strik et al., "Directions for the future of technology in pronunciation research and teaching," *Journal of Second Language Pronunciation*, vol. 4, no. 2, pp. 182-206, 2019.

[28] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, J. André et al., "The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing," *IEEE Transactions on Affective Computing*, vol. 7, no. 2, pp. 190-202, 2015.

[29] F. Eyben, F. Weninger, F. Gross, and B. Schuller, "Recent developments in openSMILE, the Munich open-source multimedia feature extractor," in *Proc. ACM International Conference on Multimedia*, pp. 835–838, 2013.

[30] C. Middag, "Automatic analysis of pathological speech," Dissertation. Ghent University, 2012.

[31] M. de Bodt, C. Guns, G. van Nuffelen, *NSVO: handleiding. Vlaamse Vereniging voor Logopedie: Herentals*, 2006.

[32] J. van de Weijer, and I. Slis, "Nasaliteitsmeting met de nasometer," *Logop Foniatr*, vol. 63, no.5, pp. 97-101, 1991.

[33] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International* 5, pp. 341-345, 2001.

[34] N. H. de Jong and T. Wempe, "Praat script to detect syllable nuclei and measure speech rate automatically," *Behaviour research methods*, vol. 41, no. 2, pp. 385-390, 2009.

[35] M. A. Efroymson, "Multiple regression analysis," *Mathematical methods for digital computers*, pp. 191-203, 1960.