

Patrizia Sorianello

Università degli Studi di Bari "Aldo Moro"

patrizia.sorianello@uniba.it

The contribution of certain acoustic vowel measures to the overall speech intelligibility

The primary aim of this research was to verify if the vowel space of Italian people with speech sound disorders was different from that of typically developing people. Several acoustic metrics derived from vowels, such as VSA-3, FCR-3, VAI-3 F1-Ratio and F2-Ratio, were applied to three different experimental groups: 1) individuals with Down Syndrome, 2) speakers with profound hearing loss, and 3) hearing subjects without any chromosomal condition. The results proved that impaired subjects displayed a reduced vowel space, both in the range of F1 and F2, a great overlap between vowels and low speech intelligibility when compared to typical speakers. All measures were able to differentiate subjects with speech disorders from controls. In contrast, FCR-3 and F1-Ratio failed to distinguish people with Down Syndrome from deaf individuals.

1. Introduction

Several studies agree in recognising the production of “distorted” vowels as one of the constant manifestations of speech intelligibility, regardless of the aetiology that induces the neurological condition, as emerges from the number of reports published on this specific topic. Results have proved that people with low speech intelligibility have a reduced corner vowel space, a significant overlap between vowels and a high degree of vowel centralisation¹. The studies have shown the presence of a close relationship between speech intelligibility and the size of the F1\F2 vowel system. In speakers with poor intelligibility, the vowel working space area is compressed with a remarkable articulatory undershoot of vowel targets. The representation of the F1\F2 formant frequencies through Cartesian diagrammes is able to

1 There is a wealth of knowledge available on this topic, among others, see Higgins and Hodge (2002); Kent et al. (1999); Kent and Rountrey (2020); Landsford and Liss (2014); Liu et al. (2005); Mou et al. (2018); Weismer et al. (2001).

highlight only the most general aspects of this process, but it is inadequate in quantifying the differences between typical and disordered speech. For this reason, to overcome this limitation, for several years now, the *Vowel Space Area* (VSA) is considered as an index of the accuracy of vowel articulation and motor control (refer to Kent et al. 1999; Liu et al. 2005; Weismer et al. 2001, among others). This metric represents the area of the irregular polygon defined by the frequencies of the first two formants (F1 and F2) of the corner vowels of a given language, typically /i/, /a/ and /u/; VSA is expressed in Hz² and is calculated according to the formula reported in (1):

$$(1) \text{ VSA: } [(F1i (F2a - F2u) + F1a (F2u - F2i) + F1u (F2i - F2a)] / 2$$

The underlying hypothesis is that the vowel space of people with disordered speech is narrower than that of control subjects, due to oral–motor, articulatory or phonological perception issues (Higgins and Hodge 2002; Lansford and Liss 2014). An undersized VSA was found in many primary or secondary neurological diseases, including motor speech disorders, in which the coordination, the synchronisation and the rate of articulatory gestures were found to be unstable and altered. This happens, for instance, in Parkinson’s disease, Amyotrophic Lateral Sclerosis (ALS), cerebral palsy, Down Syndrome, but also in cases of tracheotomy or glossectomy (or other diseases alike). The latter are specific clinical conditions in which the precision and the coordination of the tongue–jaw complex, as well as the control of the muscles of the larynx, are profoundly deviated. All this determines a significant narrowing of the articulatory–acoustic vowel space, with a clear overlapping of the vowels that tend to converge towards the centre of the system. VSA is generally used with the aim of evaluating the alterations that a vowel system undergoes in the presence of a speech disorder; several studies have shown that VSA can be considered a predictive metric of the degree of speech intelligibility; at the same time it is a valid discriminating factor between dysarthric and non–dysarthric speech (Higgins and Hodge 2002; Liu et al. 2005; Weismer et al. 2001). The effectiveness of this vowel measure has been demonstrated in clinical populations, which differed for speech disorders, language and age².

Nevertheless, the presence of an enlarged VSA is also possible in speech disorders, for instance, it was found in people with hearing–impairment as a sort of strategy to compensate the lack of auditory feedback, probably triggered by therapists (Baudonck et al. 2011), or in all those conditions causing speech hyperarticulation.

However, in some studies, the descriptive effectiveness of VSA turned out to be rather weak. Sometimes, the metric was not able to distinguish between patients with different types of dysarthria (Multiple Sclerosis vs. ALS), at other times it did not clearly separate patients from the controls (Lansford and Liss 2014). The rea-

2 VSA was also used to distinguish read from spontaneous speech or to grasp the differences due to diatopic conditioning (Fox and Jacewicz 2009).

sons for this predictive failure were diverse. First of all, it must be underlined that most studies were carried out on English-speaking patients, in fact only a handful of studies have been published thus far on other languages³. Second, triangular VSA (tVSA or VSA-3) was very often calculated considering only /i a u/ and this probably underestimated the vowel space, particularly in languages like English, which is characterised by four-point vowels⁴. According to Sapir and co-workers (2010, 2011) and Sandoval and colleagues (2013), the speech samples examined differed not only for inter-speaker variables, such as gender, age and severity of the disorder, but also for anatomical differences or for differences connected to idiosyncratic strategies of posturing the articulators. These are essentially the reasons why in recent years VSA has been integrated with the measurement of other alternative acoustic metrics, such as the Formant Centralization Ratio (FCR)⁵, the Vowel Articulation Index (VAI)⁶, the F1-Ratio and F2-Ratio, whose formulas are reported in (2), (3), (4) and (5), respectively:

- (2) $FCR = (F2u + F2a + F1u + F1i) : (F2i + F1a)$
 (3) $VAI = (F2i + F1a) / (F2u + F2a + F1u + F1i)$
 (4) $F1\text{-Ratio} = F1a / (F1i + F1u)$
 (5) $F2\text{-Ratio} = F2i / F2u$

Unlike the VSA measure, FCR and its inverse metrics VAI allow to normalise the vowel space, minimising the inter-speaker variability. In normal speech, regardless of age and gender effects, the value of both metrics is near 1; FCR ratios greater than 1 or, vice versa, VAI ratios lower than 1 have been typically found in subjects affected by disordered speech. Skodda and colleagues (2011) noticed that the VAI measure distinguished both male and female subjects with Parkinson's disease from controls, while the triangular VSA only captured the difference between male speakers. Both the CR and VAI measures are based on the three-vowel systems /i a u/, but recently VAI was also calculated on four-vowel systems (VAI-4) in order to better capture the acoustic properties of quadrilateral vowel areas (Ajish et al. 2019).

The descriptive framework described in this paper was also enriched by two further vowel metrics: the F1-Ratio and the F2-Ratio. The first quantifies, as a ratio, the amount of movements made along the vowel height dimension; the value of this metric is directly proportional to the degree of jaw lowering and mouth opening. The second, also called *Vocalic Functional Ratio*, helps to quantify the phonetic distance between the two corner vowels /i/ and /u/; this measure is very sensitive to variations in front-back movements of the tongue and to differences between

3 For Persian, Slovenian and Croatian languages, see Naderifar et al. (2017).

4 In more recent years, some studies, mainly focused on American English, have represented a quadrilateral vowel system (qVSA or VSA-4), estimating VSA on four vowels, instead of three (Bunton and Leddy 2011; Carl et al. 2020; Higgins and Hodge 2002). See also Fougeron and Audibert (2011) for the French language.

5 Sapir et al. (2010).

6 Skodda et al. (2011).

rounding and un-rounding of the sounds of the lips; the ratio is low in the presence of vowel centralisation⁷.

2. The groups of speech sound disorders

This research examined the production of speech in speakers with Down Syndrome vs. speakers affected by congenital deafness. These are two very different conditions, characterised by specific manifestations: the first having cognitive and the second having sensorial repercussions. Despite their diversity, both deaf individuals and people with Down Syndrome show a significant speech impairment and are characterised by a scarce speech intelligibility.

Down Syndrome (DS) is a disorder with a genetic origin and is associated with an intellectual disability, ranging from mild to severe. The linguistic profile of DS is peculiar: imprecise production of consonants and vowels, syllable alterations with frequent intrusive schwa, stuttering, vowel prolongations, low rate, prosodic monotony and harsh quality of the voice are just some of the processes most frequently encountered (among all publications refer to: Do and Thompson 2001; Kent and Vorperian, 2013; Nadel 1988; Penz and Moran 1988; for Italian publications see: Contardi and Vicari 1995; Sorianello 2012, 2015). The speech sound disorder is due both to central causes, such as cognitive delay, short-term memory difficulties, motor programming deficits, and to peripheral causes, including muscular hypotonia, cranial and facial anomalies, macroglossia, and/or a reduced labial and mandibular range. However, it is important to keep in mind that people with DS acquire language spontaneously, although developmental stages are delayed over time if compared to typically developing individuals. People with DS show a low speech intelligibility that, although often attenuated by speech therapy, is never completely compensated during their life. With explicit reference to acoustic metrics derived from vowels, the empirical results proved that individuals with DS have a smaller acoustic vowel area and a reduced articulatory working space than control subjects. For Anglo-American individuals, Bunton and Leddy (2011) observed how quadrilateral VSA, calculated on a sample of monosyllables spoken in isolation, was about 25% lower than the control group; Kent and Vorperian (2013) reported a similar result a few years later. As regards Italian-speaking subjects with DS, Sorianello (2015) observed a relevant compression of the triangular VSA; in male individuals the vowel acoustic space was smaller of 60% than typical subjects. All this indicates the presence of a speech sound disorder that mostly invalidates the front-back tongue movements.

The clinical profile concerning deafness (D) is different. Deaf people present a sensorineural hearing damage, which can be of primary or secondary nature. The linguistic consequences of deafness are diversified and deeply linked to the type,

7 Kent and Vorperian (2013); Moura et al. (2008); Sapir et al. (2010).

degree, period of insurgency and treatment of the hearing loss (Caselli et al. 2006). More in general, hearing loss affects language development, both in the oral comprehension and in the production, with significant influences on the cognitive abilities⁸. If deafness is not mitigated at an early stage with hearing aids or, when possible, with cochlear implants, deaf people remain deprived of language. The lack of auditory input interferes with the capacity to plan and execute oral-verbal messages. Consequently, in the presence of profound prelingual hearing loss, the oral communication is strongly altered. The speech sounds are degraded, fluency, rhythm and intonation are inadequate, while the voice quality is characterised by excessive muscular tension and hypernasality (Monsen 1976; Rahilly 2013). With specific reference to vowels, both high and front vowels appear more damaged than low or back ones, the neutralisation of the more peripheral vowels is frequent too (Angelocci et al. 1964; Barzaghi and Mendes 2008; Nicolaidis and Sfakianaki 2016; Ozbič and Kogovšek 2008; Yasuyoshi and Yoshino 1989). Few studies have adopted VSA, or other acoustic metrics, thus far for the description of the D vowel system. The VSA value was mostly calculated in deaf subjects with a cochlear implant (Baudonck et al. 2011; Neumeyer et al. 2010). In this direction, Löfqvist and co-workers (2010) reported that the vowel area of Swedish deaf children with cochlear implants was smaller and deviated compared to that of normally hearing children. The alterations were clear even five years following cochlear implantation, as pointed out by a recent research conducted on Belgian deaf children during a vowel imitation task (Verhoeven et al. 2016). In Hindi deaf children VSA-4 and VAI-4 were lower than typically developing children, while FCR-3 was higher, indicating a greater degree of vowel centralisation (Ajish et al. 2019).

No experimental data dealing on this specific topic conducted on Italian deaf people is available, to the best of my knowledge.

3. The research

The main goal of this research was to investigate whether individuals with sound speech disorder and typical subjects differed in the way they performed stressed vowels. To achieve this goal, we considered the following acoustic metrics: VSA-3, FCR-3, VAI-3, F1-Ratio, F2-Ratio and the estimation of the Centroid. Three different groups of speakers participated in the research: individuals with Down Syndrome (DS), deaf subjects (D) and typically developing individuals as controls (CG). The working hypotheses were that in the speech of DS and D individuals the articulation of vowels was altered. We wanted to ascertain the descriptive effectiveness of the metrics in distinguishing speakers with disordered speech from typical speakers. The expected behaviours were essentially two:

8 Hearing loss by itself does not cause cognitive impairment. Nevertheless, if deaf children are deprived of language, in the absence of a specific rehabilitation, they will also have cognitive problems over time, since general knowledge about the world is acquired using language and by interacting with others.

hypothesis 1= the metrics will reflect the differences in accuracy of vowel articulation between deaf speakers and speakers with DS as two groups with reduced speech intelligibility, and hearing speakers without chromosomal condition as the group with typical vowel articulation, and will show their potential to overcome the constraints in speech intelligibility assessment imposed by the language or by the aetiology;

hypothesis 2= the metrics cannot discriminate speakers with a low speech intelligibility from subjects with a typical vowel articulation. It is important to underline that these vowel metrics have been tested for speech intelligibility assessment on a limited number of languages and never for Italian.

In addition, it should be noted that VSA has shown some weak aspects in previous studies (Lansford and Liss 2014; Sandoval et al. 2013; Sapir et al. 2010, 2011); however, we think that in a language like Italian, which has a typical triangular vowel system formed by the corner vowels /i a u/, the metric could have a high descriptive potential. For this reason, all the measures were based on the three cardinal vowels.

3.1. Participants

A total of three groups of Italian-speaking people took part in the research. All participants were first language speakers of the same variety of Southern Italian⁹. The first group consisted of eight people with Down Syndrome (DS), 4 males and 4 females, the mean chronological age was 22.2 years old (SD: 1.8). These were students or young graduates that had been subjected to a speech therapy treatment in the past. The second group was represented by eight deaf subjects (D), 4 males and 4 females; they had a mean age of 23.5 years old (SD: 2.1) and were affected by profound bilateral congenital deafness (> 70 dB HL). Deaf participants were educated following an oralist approach based on lip-reading and supported by speech therapy; three of them knew or used the LIS (Italian Sign Language), two subjects had deaf parents. None of them had a cochlear implant or a hearing aid. Finally, the control group (CG) was formed by eight hearing people, 4 males and 4 females, who were all students or graduates, without any form of speech disorder (mean age: 24.3 years old; SD: 1.2). The participants, or their parents in the case of DS individuals, gave written informed consent to take part in the research.

3.2. Procedure

All participants were audio-recorded in a silent room with digital instruments (Tascam DR-07). Digital recordings in Wav audio format were digitised at the sampling frequency of 44Khz, with a 32-bit quantisation rate. Before starting, all subjects were put at ease, in a condition of spontaneous and relaxed cooperation. The

9 All the participants came from Cosenza (Italy) and outskirts and spoke a regional variety of Italian characterised by a five-vowel phonemic system, i.e. /i ɛ a ɔ u/.

participants were instructed on the speech task to conduct. In order to elicit connected free speech, they produced a description task of a picture-story. To this end, we used two different single pictures representing simple scenes of daily life.

Overall, the speech sample consisted of 1000 stressed vowels, around 300 for each group. The acoustic analysis was performed with *PRAAT* (vers. 6.0.31, Borsma and Weenink 2007); speech samples were transcribed orthographically and time-aligned by means of *Elan* software; *Text grid files were created as output and then processed within *PRAAT*.

The mean values and standard deviation of the first and second formant frequencies (F1, F2) of the stressed corner vowels /i/, /a/, /u/ were manually extracted using an FFT procedure on the 50 ms steady-state portion of the vowel. The formant values were used to calculate the following acoustic metrics: 1) the triangular Vowel Space Area (VSA-3), 2) the triangular Formant Centralization Ratio (FCR-3), 3) the triangular Vowel Articulation Index (VAI-3), 4) the F1-Ratio, and 5) the F2-Ratio. The metrics were computed using the formulas reported in points (1-5) and then compared across groups. Furthermore, we proceeded with the estimation of the Centroid based on the F1/F2 frequencies of the point vowels /i a u/¹⁰.

The statistical analysis was run using the *SPSS* software (IBM Software, vers. 20.0). We used one-way ANOVAs with Tukey post-hoc tests and Mann-Whitney U test to compare group differences; results were considered significant for p-values below .05.

4. Results

Relevant differences emerged between the three groups (DS, D, CG) in the rendering of the corner vowels. Overall, all the metrics explored were able to distinguish DS and D subjects from the CG. Figures 1 and 2 show the surface area of the vowel space (in Hz²) contained within the vowel triangle determined by the three corner vowels (VSA-3), for male and female subjects of the three groups, respectively. As regards this measure, we observed that the values computed for females were always higher than those of male subjects ($p < 0.001$), as was expected (compare Fig. 1 and Fig. 2). VSA-3 was not able to neutralise gender influences, but it was an acoustic metric capable of discriminating subjects with speech disorder from control individuals [all: $F(2.1)=3.66$, $p=.03$; males: $F(2.9)=13.59$, $p=.02$; females: $F(2.10)=9.2$, $p=.05$]. In the CG, VSA-3 reproduced a well-defined vowel space in which the position of the corner vowels respected the typical triangular shape of the Italian vowel system. In contrast, the DS and D groups showed a wide-spread compression of the vowel area. More specifically, the differences between

10 The Centroid is the grand mean of all formant values of each speaker (Koopmans-van Beinum 1983; Watt and Fabricius 2003), but it can be also calculated on three or four corner vowels. The Centroid may be considered a sort of neutral point within the vowel system of a specific group of speakers or variety of language, for this reason it has been used as a measure able to grasp the degree of vowel centralisation.

the DS and CG groups amounted to + 35% in males ($p=.04$) and to + 24% in females ($p=.04$); the distance between the D and CG groups was more conspicuous, being equals to + 87% in males ($p=.001$) and to + 68% in females ($p=.02$). DS and D subjects showed both a reduced vowel working space and a global shift of the vowel targets. In this respect, it was observed that in males with DS /u/ appeared centralised, with the consequent reduction of the vowel height axis, while in D subjects, regardless of gender, the corner vowels appeared contracted towards the inside, with a significant compression of the vowel distance in the front-back plane. Overall, VSA-3 of the D group was smaller to that obtained for speakers with DS, the difference was significant for both male ($p=.03$) and female speakers ($p=.01$). This proved the presence of a strongly degraded speech in which the phonological contrast among vowels, produced with a remarkable degree of centralisation, were often acoustically neutralised. Most vowels tended to converge towards the middle of the articulatory space, which is the area normally occupied by indistinct vowels. In D subjects, the movements of the tongue had, in fact, a limited range, especially with regard to the F2 axis.

This centralisation effect was also confirmed by the mean frequencies of the Centroid (see Figure 3). Comparing the data, we noticed that the Centroid values did not always differentiate the groups (F1: $p=.096$, F2: $p=.082$) and that the gender effect emerged only for the CG group [$p=.001$].

The analysis of the other vowel metrics helped to better define the descriptive framework (see Figure 4). As expected, the value of FCR was around 1 only in the CG. In both DS and D subjects this parameter showed a clear centralisation tendency, although the difference between the DS and D groups was not significant [$F(1.9)=0.8$, $p=.93$]. We found a similar trend for VAI-3; in this case ratios lower than 1 revealed the presence of a phonetic process of vowel undershoot. D and CG groups behaved differently [$F(2.13)=4.22$, $p=.039$]. According to the VSA-3 findings, a greater degree of reduction in vowel space was found for D individuals and the alteration of vowel gestures was appreciable both along the vertical axis (F1) and, to a greater extent, along the horizontal axis (F2). The calculation of the F1-Ratio and F2-Ratio allowed us to better quantify the vowel space dispersion (refer to Fig. 4 for details).

The F1-Ratio index (degree of openness: high V vs. low V) was lower in the DS group compared to the other two groups. However, the difference was significant only when compared to the CG group [$F(2.17)=7.67$, $p=.04$], but not when compared to the D subjects, whose corner vowels, in fact, did not distance appreciably from the DS group [$F(1.9)=3.08$, $p=.11$]. In DS the reduction of the F1 range was due to several anatomical and functional impairments affecting both the movements of the jaw and the raising of the tongue within the oral cavity. Differences concerning the F2-Ratio index (place of articulation: front V vs. back V) were even stronger. The value was high in the CG group (Ratio: 2.52), low in the D group (Ratio: 1.39) and intermediate in the DS group (Ratio: 1.92); the differences between

groups were significant [$F(2.13)=16.1, p=.0001$], as well as those between DS and D individuals [$F(1.10)=7.8, p=.019$].

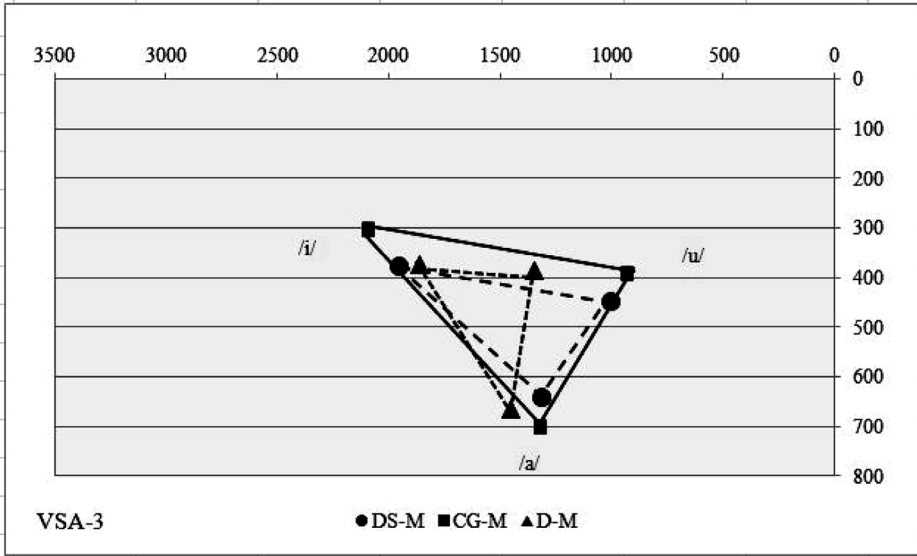


Figure 1. Mean acoustic VSA-3 (Hz2) for male participants

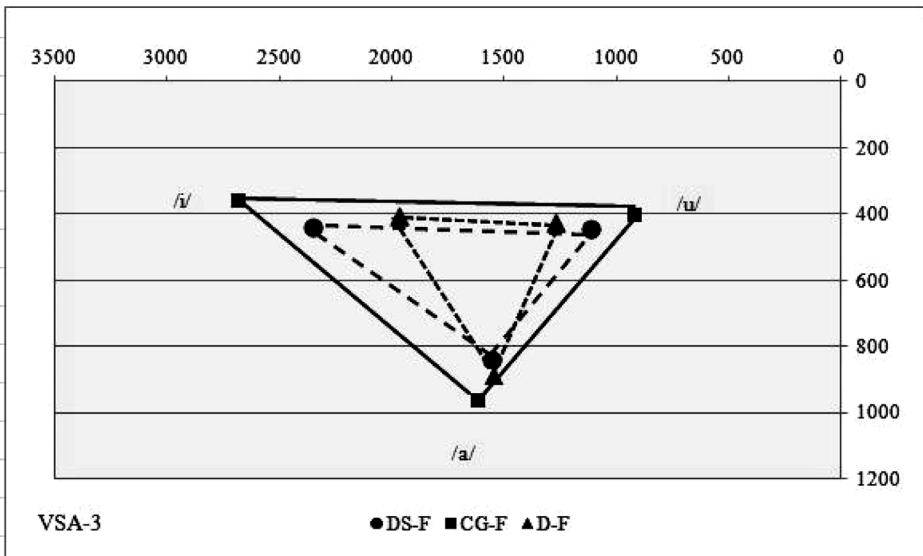


Figure 2. Mean acoustic VSA-3 (Hz2) for female participants

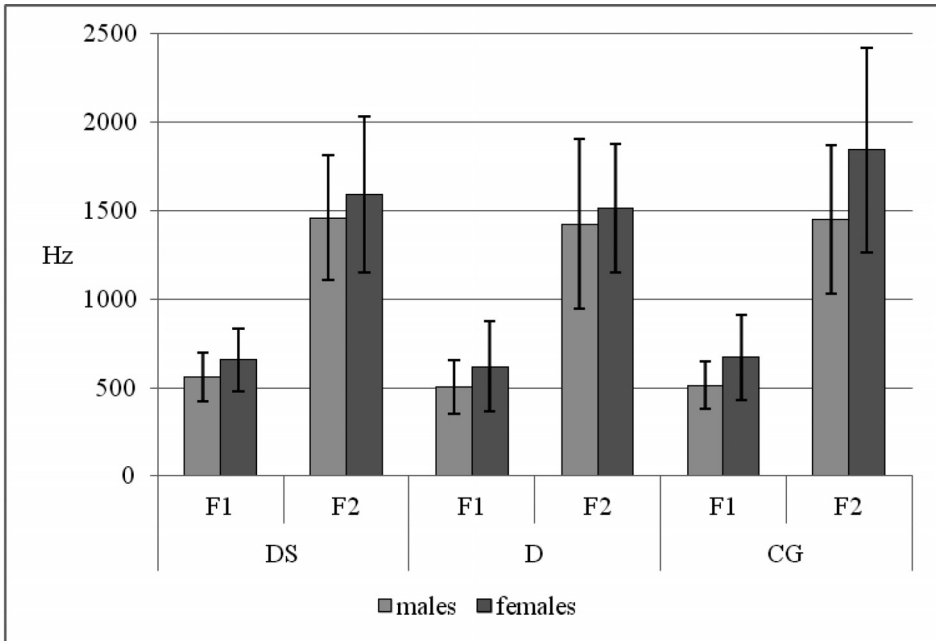


Figure 3. Mean values in Hz of the Centroid for each of the three groups (DS, D and CG), separated for gender (blue bars: males; red bars: females); the error bar represents the Standard Deviation

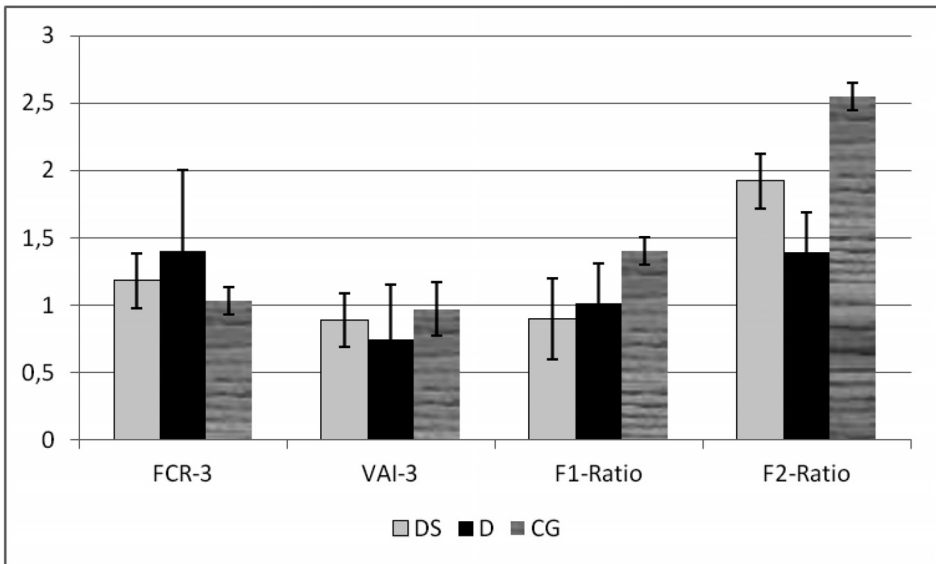


Figure 4. Mean ratios of metrics for the three groups (DS, D and CG); the error bar represents the Standard Deviation

5. Discussion

In this research we considered two groups of speakers with disordered speech and five acoustic metrics frequently used to estimate the vowel alterations of impaired subjects. The experimental procedure showed that all the metrics were able to discriminate the DS and D groups from the CG; the only exception was represented by VAI-3 that was at the limit of the critical threshold of statistical significance (5%). The first working hypothesis was thus confirmed. The vowel working space of DS and D individuals displayed a high degree of reduction compared to the CG. In both DS and D subjects, the vowels lost much of their articulatory distinctiveness, the point vowels were in fact projected towards a central area, causing a sensible compression of the vowel space, as proved by the VSA-3 values. The centralisation of the corner vowels emerged from the FCR-3, VAI-3 and Centroid values. Many phonological targets underwent a deep articulatory shift, which made them difficult to distinguish on a perceptual level, thus generating a certain confusion and a significant degree of unintelligibility.

In light of the results gathered, it can therefore be postulated that the vowel metrics, with reference to the groups considered, are reliable indexes capable of distinguishing typical from disordered speech. However, VSA-3 and F2-Ratio were the only parameters that were always able to distinguish DS from D subjects. In the DS and D groups, FCR-3, VAI-3, and more clearly the F1-Ratio, displayed in fact similar values, which were not statistically significant. The second working hypothesis was therefore only partially confirmed. This behaviour led to some reflections. In our opinion, the weakness of VAI-3 and the F1-Ratio to clearly discriminate between DS and D individuals depended on the experimental groups considered. Although the aetiology and the language acquisition modalities of DS and D individuals are different, their linguistic production shares at least two aspects: they both present articulatory impairment and low speech intelligibility. DS and D subjects show a damaged language. In our research, only VSA-3 and the F2-Ratio were predictive metrics of an impaired speech, since they were able to distinguish not only the DS and D groups from the CG, but also DS from D individuals. In contrast to what was observed in some past studies (Lansford and Liss 2014; Sandoval et al. 2013; Sapir et al. 2010), VSA-3 was shown to work rather well, although calculated only on three vowels. It is also interesting to point out that VSA-3 was able to differentiate DS and D subjects from typical speakers in spontaneous speech samples. Likewise, the F2-Ratio appeared to be a very sensitive index in the distinction between groups of impaired speech. A confirmation in this direction comes from the fact that the vowel system of deaf subjects, including those that are implanted, undergoes a deep deviation, which involves F2 more specifically; the distance between front and back vowels is generally reduced (Monsen 1976; Verhoeven et al. 2016). The F2 lowering, the first to have been documented in a sample of Italian deaf subjects, is a typical feature of the vowel production of individuals who are hearing-im-

paired. As reported by Monsen (1976), the articulatory gestures that occur along the horizontal dimension (F2) are less clear at a visual inspection, which is the only one possible for deaf people, if compared with the vowel gestures that occur along the vertical axis (F1). For this reason, in deaf people, the value of the second formant is systematically placed around 1700 Hz, determining the so-called ‘immobility of F2’, an aspect that affects the perceptual distinction of both vowels and consonants.

As it is known, the linguistic difficulties of deaf people are to be ascribed to the lack of an auditory feedback. With a compromised hearing, the attention of deaf subjects is directed towards other resources, such as lip-reading. With specific reference to the vowels, deaf people focus on what is easier to recover through sight, for instance the lowering of the jaw, performing a sort of visual feedback. Therefore, the movements concerning the vowel height (F1-Ratio) are better controlled than those performed along the anterior-posterior dimension (F2-Ratio); the latter are in fact more problematic for deaf individuals, given their inability of controlling the gestures relating to the retraction of the tongue through sight.

The situation that emerges in people with DS is different. In this case, the precision of the articulatory gestures is conditioned by those anatomical and functional specificities that directly affect their phonation, respiratory and auditory systems (Do and Thompson 2001; Kent and Vorperian 2013; Nadel 1988). The oral cavity is small, while the palate has an ogival and fissured shape; the movements of the tongue are both limited and inaccurate, because it is too large compared to the oral cavity. The dynamics of phonation are also conditioned by a poor muscle tone. Nevertheless, when compared to D subjects, people with DS show less impairment in vowel articulation, as is highlighted by both a wider VSA-3 and F2-Ratio. The oral language of D subjects is therefore more degraded: the absence of auditory input not only affects the ability to produce oral utterances adequately, but also the ability to mentally represent the sounds by classifying them into phonemic classes, a process that allows to capture their distinctive features by abstraction. The absence of an early natural language, which is spontaneously acquired from birth, has a strong impact on human neurocognitive development, being determined both by language and by sensory experiences; thus, deaf people who do not succeed in mastering a spoken language will have some difficulties with the mental representation of spoken words and, therefore, with phonology and articulation (Landau 2010; Neville et al. 1998).

The research presented here has provided several clues concerning vowels realisation, allowing to outline a sort of hierarchy of the descriptive force of the metrics investigated. In this perspective, VSA-3 and the F2-Ratio were always distinctive, since they were able to capture the differences in vowel accuracy among groups. On the contrary, VAI-3 played an irrelevant role. All this produces phonetic and phonological consequences. On the articulatory level, the greater the reduction of VSA-3, the greater the probability that the vowels overlap towards the centre of the

system. On the perceptual level, on the other hand, a higher degree of vowel overlapping and centralisation corresponds to a progressive loss of the phonological distinctiveness of the segments. This happens, for example, when the phonological categories, being too close together, lose their acoustical and perceptual identity.

The results obtained in this research are promising considering the scarcity of Italian studies carried out in this direction. This study also has the merit of having considered connected speech, instead of isolated words, although the corpus of the data was still elicited in a controlled way. Our research opens up to other lines of investigation, such as the observation of other speech pathologies, the perceptual effects induced by a reduced VSA–3 or the application of other vowel measures and can also find a useful application in the field of speech therapy.

6. Conclusions

We measured certain vowel metrics in two groups of impaired speakers with low speech intelligibility (DS, D) and in typically developing individuals (CG). Compared to controls, all metrics, except for VAI–3, showed significantly different values in DS and D individuals. Among all, VSA–3 and the F2–Ratio appeared as the metrics with the greatest differentiating force, since they were always able to distinguish speech-impaired groups from the control group. For this reason, they can be considered good predictors of speech sound disorders.

Acknowledgements

The author would like to thank the DS, D and CG participants who kindly took part in this experimental research and two anonymous reviewers for their helpful comments and suggestions.

References

- Ajish, Abraham K., Mariswamy Pushpavathi, Nataraja Sreedevi, and Attuluri Navya (2019). Exploring acoustic measures of vowels (VSA, FCR3, VAI4, VFR) in children with hearing impairment. Calhoun, Sasha, Paola Escudero, Marija Tabain, and Paul Warren, eds. *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019: 1064–1068
- Angelocci, Angelo A., George A. Kopp, and Anthony Holbrook (1964). The vowel formants of deaf and normal-hearing eleven-to fourteen-year-old boys. *Journal of Speech and Hearing Disorders* 9(2): 156–170
- Barzaghi, Luisa, and Beatriz Mendes (2008). Stressed and unstressed vowels production in hearing-impaired. Barbosa, Plinio L., Sandra Madureira, and Cesar Reis, eds. *Proceedings of the International Conference on Speech Prosody*, Campinas, Brazil, 6–9 May: 199–212

- Baudonck, Nele, Kristiane Van Lierde, Ingeborg Dhooge, and Paul Corthals (2011). A comparison of vowel productions in prelingually deaf children using cochlear implants, severe hearing-impaired children using conventional hearing aids and normal-hearing children. *Folia Phoniatrica et Logopaedica* 63(3): 154–160
- Boersma, Paul, and David Weenink (2007). *Praat: doing phonetics by computer* (version 5.0), University of Amsterdam, The Netherlands. Online: <http://www.praat.org>
- Bunton, Kate, and Mark Leddy (2011). An evaluation of articulatory working speech area in vowel production of adults with Down syndrome. *Clinical Linguistics and Phonetics* 25(4): 321–334
- Carl, Micalle, Raymond D. Kent, Erika S. Levy, and D. H. Whalen (2020). Vowel acoustics and speech intelligibility with Down Syndrome. *Journal of Speech, Language and Hearing Research* 63(3): 674–687
- Caselli, Cristina, Simonetta Maragna, and Virginia Volterra (2006). *Linguaggio e sordità*. Bologna: Il Mulino
- Contardi, Anna, and Stefano Vicari (eds.) (1995). *Le persone Down. Aspetti neuropsicologici, educativi e sociali*. Milano: FrancoAngeli
- Do, Barbara, and Laura Thompson (2001). Speech disorder in children with Down's syndrome. *Journal of intellectual disability research* 5(4): 308–316
- Fougeron, Cécile, and Nicolas Audibert (2011). Testing various metrics for the description of vowel distortion in dysarthria. *Proceedings of the XVII International Congress of Phonetic Sciences (ICPhS)*, Hong Kong, China, 687–690
- Fox, Robert Allen, and Eva Jacewicz (2009). Cross-dialectal variation in formants dynamics of American English vowels. *The Journal of the Acoustical Society of America* 126(5): 2603–2618
- Higgins, Carolyn, and Megan M. Hodge (2002). Vowel area and intelligibility in children with and without dysarthria. *Journal of Medical Speech and Language Pathology* 10(4): 271–277
- Kent, Raymond D., and Carrie Rountrey (2020). What acoustic studies tell us about vowels in developing and disordered speech. *Journal of Speech, Language and Hearing Research* 29(3): 1749–1778
- Kent, Raymond D., and Hourii K. Vorperian (2013). Speech impairment in Down Syndrome: a review. *Journal of Speech Language and Hearing Research* 56(1): 178–210
- Kent, Raymond D., Gary Weismer, Jane Kent, Hourii Vorperian, and Joseph Duffy (1999). Acoustic studies of dysarthric speech: methods, progress and potential. *Journal of Communication Disorders* 32(3): 141–186
- Koopmans-van Beinum, Florian J. (1983). Systematics in vowel systems. In von den Broecke, Marcel, Vincent van Heuven, and Win Zonneveld, eds. *Sound Structures. Studies for Antonie Cohen*, Dordrecht, De Gruyter Mouton: 159–171
- Landau, Barbara (2010). Path in language and cognition: Universal asymmetries and their cause. In Marotta Giovanna, Alessandro Lenzi, Linda Meini, and Francesco Rovai, eds. *Space in language. Proceedings of the Pisa International Conference*, Pisa: ETS, 73–94
- Lansford, Kaitlin L., and Julie M. Liss (2014). Vowel acoustics in dysarthria: speech disorder diagnosis and classification. *Journal of Speech Language and Hearing Research* 57(1): 57–67

- Löfqvist, Anders, Birgitta Sahlén, and Tina Ibertsson (2010). Vowel space in Swedish adolescent with cochlear implants. *Journal of Acoustical Society of America* 128(5): 2064–3069
- Liu, Hiu Mei, Feng Ming Tsao, and Patricia K. Kuhl (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy, *Journal of Acoustical Society of America* 117(6): 3879–3889
- Monsen, Randall B. (1976). Normal and reduced phonological space: the production of English vowels by deaf adolescents. *Journal of Phonetics* 4(3): 189–198
- Mou, Zhiwei, Zhuoming Chen, Jing Yang, and Li Xu (2018). Acoustic properties of vowel production in Mandarin-speaking patients with post-stroke dysarthria, *Scientific Reports* 18, 14188, <https://doi.org/10.1038/s41598-018-32429-8>
- Moura, Carla P., Luís M. Cunha, Helena Vilarinho, Maria J. Cunha, Diamantino Freitas, Miguel Palha, Siegfried M. Poeschel, and Manuel Pais-Clemente (2008). Voice parameters in children with Down Syndrome. *Journal of Voice* 22: 34–42
- Nadel, Lynn (ed.), (1988). *The psychobiology of Down syndrome: issues in the biology of language and cognition*, Cambridge (Mass.): MIT Press
- Naderifar, Ehsan, Ali Ghorbani, Negin Moradi, Hossein Ansari, Ozra Arghadoost, Faezeh Asadollahpour and Martina Ozbic (2017). Evaluation of formant frequencies in Persian speaking children with different degrees of hearing loss. *Shiraz E-Medical Journal* 18(7):e13094
- Neumeyer, Veronika, Jonathan Harrington, and Christoph Draxler (2010). An acoustic analysis of the vowel space in young and old cochlear-implant speakers. *Clinical Linguistics and Phonetics* 24(9): 734–741
- Neville, Helen J., Daphne Bavelier, David Corina, Josef Rauschecker, Avi Karni, Anil Lalwani, Allen Braun, Vince Clark, Peter Jezzard, and Robert Turner (1998). Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. *Proceedings of the National Academy of Sciences* 95, 922–929
- Nicolaidis, Katerina, and Anna Sfakianaki (2016). Acoustic characteristics of vowels produced by Greek intelligible speakers with profound hearing impairment I: Examination of vowel space. *Journal of Speech Language Pathology* 18(4): 1–10
- Ozbič, Martina, and Damiana Kogovšek (2008). An acoustic comparison of formant frequencies in individuals with normal hearing, profound and severe hearing impairment. *Investigationes Linguisticae: VXi*, 1–13
- Penz, Arthur, and Michael J. Moran (1988). Voice disorders in Down Syndrome. *Journal of Childhood Communication Disorders* 12(1): 109–117
- Rahilly, Joan (2013). Vowel disorders in hearing impairment. Ball Martin J., and Fiona E. Gibbon, eds. *Handbook of vowels and vowel disorders*. New York & London: Psychology Press, 364–383
- Sandoval, Steven, Visan Berisha, Rene L. Utianski, and Julie M. Liss (2013). Automatic assessment of vowel space area. *Journal of Acoustical Society of America* 134(5): 477–483
- Sapir, Shimon, Lorraine O. Ramig, Jennifer L. Spielman, and Cynthia Fox (2010). Formant centralization ratio (FRC): A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech Language and Hearing Research* 53(1): 114–125

- Sapir, Shimon, Lorraine O. Ramig, Jennifer L. Spielman, and Cynthia Fox (2011). Acoustic metrics of vowel articulation in Parkinson's disease: vowel space area (VSA) vs. vowel articulation index (VAI). Manfredi Claudia, ed. *7th International workshop on Models and analysis of vocal emissions for biomedical applications*. Firenze: Firenze University Press:173–175
- Skodda, Sabine, Wenke Visser, and Uwe Schlegel (2011). Vowel articulation in Parkinson's disease. *Journal of Voice* 25(4): 467–472
- Sorianoello, Patrizia (2012). *Linguaggio e Sindrome di Down*. Milano: FrancoAngeli
- Sorianoello, Patrizia (2015). Some effects of Vowel Space Area (VSA) reduction on speech intelligibility. Botinis, Antonis, ed. *Proceedings of the 6th International Conference of Experimental Linguistics (ExLing2015, ISCA)*, 26–27 June 2015, Athens, Greece: 78–81
- Verhoeven Jo, Oydis Hide, Sven De Maeyer, San Gillis, and Steven Gillis (2016). Hearing impairment and vowel production. A comparison between normally hearing, hearing-aided and cochlear implanted Dutch children. *Journal of Communication Disorders* 59: 24–39
- Watt, Dominic, and Anne H. Fabricius (2003). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the f1-f2 plane. *Leeds Working Papers In Linguistics* 9: 159–173
- Weismer, Gary (2007). The acoustic vowel space in speech disorders: Data and interpretation. *Journal of the Acoustical Society of America* 122(5): 3015
- Weismer, Gary, Jing Yi Jeng, Jacqueline S. Laures, Ray D. Kent, and Jane Kent (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica* 53(1): 1–18
- Yasuyoshi, Kato, and Yoshino Tomoyoshi (1989). Acoustic analysis of vowels in adolescents with profound hearing impairments. *Bull. Spec. Educ*, 14(1): 17–24

Doprinos određenih akustičkih mjera vokala cjelokupnoj govornoj razumljivosti

Primarni je cilj ovog istraživanja da provjerimo je li vokalski prostor Talijana s poremećajima izgovora različit od osoba tipičnoga govornog razvoja. Nekoliko akustičkih mjera ispitanih na vokalima, kao što su VSA-3, FCR-3, VAI-3 F1-Ratio i F2-Ratio, primijenjeno je na trima različitim eksperimentalnim skupinama: 1) osobama s Downovim sindromom, 2) govornicima s teškim oštećenjem sluha i 3) čujućim osobama koje nemaju nikakve kromosomske poremećaje. Rezultati su pokazali da ispitanici s oštećenjima imaju sužen vokalski prostor, koji se u usporedbi s govornicima tipičnoga govornoga razvoja ogleda u suženome rasponu F1 i F2, velikom preklapanju vokala te niskoj razini razumljivosti govora. Ispitanici s izgovornim poremećajima razlikovali su se u svim mjerama od kontrolne skupine ispitanika. Nasuprot tome, osobe s Downovim sindromom nisu se u FCR-3 i F1-Ratio mjerama razlikovale od gluhih osoba.

Keywords: speech disorders, vowels, acoustic analysis, Italian

Ključne riječi: govorni poremećaji, samoglasnici, akustička analiza, talijanski jezik