

ELISA PELLEGRINO, VALERIA CARUSO, ANNA DE MEO

Discriminating between pathological and foreign-accented speech: an acoustic-perceptual study on Italian

This study investigates native Italian listeners' ability to discriminate pathological from foreign-accented speech, comparing the read performance in Italian of Italian and Ukrainian deaf talkers with those produced by hearing Japanese and Ukrainian L2 speakers. Seventeen native Italian listeners rated the speech samples for 'accentedness', 'comprehensibility' and 'speech type'. The comparison between perceptive and acoustic data shows the difficulties of listeners to distinguish between native and non-native deaf speech. The mixed prosodic profile of the Japanese speaker and the great number of segmental misarticulations lead listeners to rate her speech similarly to those of deaf speakers, namely as 'pathological' and 'scarcely comprehensible'. Conversely, the Vowel Space Area and the consonantal misarticulations type differentiate her from the deaf speakers, and are probably responsible for the judgement concerning her 'non-nativity' to the Italian language.

1. *Introduction*

This study focuses on the oral productions of prelingual deaf speakers, whose perceptive deficit has been recognized by many researchers as a type of impairment that prevents the development of a native-like competence in the vocal languages:

[...] one can report that all studies in the literature agree on the fact that young and adult deaf quite never reach a linguistic competence similar to that of a native speaker, neither in the oral nor in the written language" (Caselli, Maragna, Volterra, 2006: 235).

Starting from these assumptions, an analysis of deaf speech seems to be valuable for the double aim of enriching the existing literature on deaf people's skills, and of researching the features responsible for the perceptual discrimination of different kinds of oral productions. More precisely, this study will give insights into the ability of native Italians to distinguish non-native speech from the pathological oral productions of deaf people. As it will be clear in a moment, in fact, it is rather difficult for native listeners to discriminate between these two types of voices. In particular, the oral productions of deaf people prove to be undistinguishable with reference to the additional native/non-native parameter. Italian listeners are in fact unable to recognize if a deaf speaker has learned Italian as a first or second vocal language.

Moreover, the listeners' inability to detect any foreign accent in deaf speech is somehow paralleled by the analogous misinterpretation of some elementary learners' speech productions, which are mistaken for a kind of pathological speech.

Similar misinterpretations will be in focus in the following pages while discussing the results of a perceptive test along with data from the spectro-acoustical analysis of different types of voices.

Before going into the details of the experimental study, a brief description of deaf people's voices will be given by surveying the existing literature.

1.1 Features of deaf people's voice reported by the literature

The bulk of literature on deaf people's speech is made up of studies carried out by medical researchers interested in the quality estimation of their vocal productions. As a consequence, the analyses carried out so far have mainly evaluated the mispronunciations and the general comprehensibility of these atypical speakers on the segmental level.

Many of these studies, however, have also described the quality of deaf people's voices which, since the 1960s (Calvert, 1962), have been described as 'tense', 'hollow', 'breathy', 'flat', but also as 'harsh' and 'throaty', two characteristics that Shriberg, Kwiatkowski & Rasmussen (1990) refer to the 'laryngeal quality' of this pathologic speech. Deaf voices tend to be 'tense' or even strident since they receive most of the vocal linguistic input from tactile sensations (Lenden, Flipsen, 2007), while the reverberation of sounds in the pharynx is responsible for their hollow character (Boone, 1971; Finkelstein et al., 1993). On the other side, the excessive force put on plosives before a vowel (Lenden, Flipsen, 2007) determines the breathy quality of their speech productions and is also responsible for improper pauses and fragmented segmentation (Pellegrino, Caruso, 2012).

Even more distinguishable are, however, the prosodic traits of this atypical speech. The irregular phrasing is indeed the most perceptible prosodic feature, since deaf people pronounce utterances word by word, lacking co-articulations and sound blends. There is also a general tendency to prolong vowels (Boone, 1971) and to slow speech rate, two factors that weaken fluency. Additionally Nickerson (1975) pointed out how deaf speakers, being unable to vary pitch contours properly, produce a monotonous speech and use pauses to signal word and sentence boundaries (Stathopoulos et al., 1986). On the other hand, some other researchers have underlined how pitch, among other traits, is influenced by the degree of hearing loss in these subjects. Bonetti, Utović & Dulčić (2008) found that prelingual mild deaf adolescents, for example, were able to control specific vocal features, such as intonation, accent, pauses and speech rate. This proved that prosody was not uniformly affected in the category of deafness because some prosodic features remain preserved with lower degrees of deafness.

1.1 Other relevant features of deaf speech

The current study is more focused on deaf speakers' linguistic skills and competencies than it is in medical research.

In a previous study, Pellegrino and Caruso (2012) have highlighted that even the most disadvantaged deaf speakers, such as foreigners with a profound hearing

impairment, have communicative competences in the oral L2 that are independent of their segmental accuracy and the level of their hearing loss. Put to test in a semi-spontaneous communicative task, profound deaf foreigners proved in fact their ability to pronounce speech acts as single utterances, performing requests at a speech and articulation rate that are comparable to those of hearing native speakers. This suggests that the chief concern of whatever kind of speaker is that of preserving communicative units, even if the uttered sentences prove to be scarcely intelligible.

In the present study, however, the viewpoint adopted is perceptive, and the focus is on the ability of the native speakers to recognize pathological and non-native traits in the voices they are listening to. Considering, in fact, the results of a preliminary investigation (Pellegrino, Caruso & De Meo, 2012) showing that some L2 spoken productions can be mistaken for 'impaired' or even 'synthesized' speech, it seems interesting to investigate which acoustic features are responsible for mistaking the voices of non-native hearing speakers for those of impaired subjects.

In order to achieve this aim, samples of different kinds of voices were collected: the oral productions of deaf speakers, both Italian and foreign citizens, as well as those produced by hearing non-native speakers. In the first research step, the speech samples were administered to a group of native Italian listeners, who rated for different parameters the voices they had been listening to. Secondly, the perceptive judgments were compared with the spectro-acoustical analysis of the oral corpus in order to identify:

- the existence of any transfer phenomena from the language of the first logopae-dic treatment to the L2 vocal Italian, and, additionally,
- the presence of any acoustic features that allow native listeners to recognize this type of pathological speech.

2. *The study*

The study thus consists of two parts: a perceptive test and a spectro-acoustical analysis. The perceptive test needed the recruitment of two groups of participants: one of readers, providing the necessary speech samples; one of native listeners, expressing judgments about the collected voices.

2.1 Participants

As for the readers, the group is composed as follows:

- 2 Italian sisters (henceforth D_L1_Fr; D_L1_Mo), with severe hearing impairment, aged 34 and 36 respectively, grown up in a hearing family, educated with an oral method, and having the vocal Italian as their L1 (see Pellegrino, De Meo & Caruso, 2014) and a low proficiency in the Italian Sign Language (LIS).
- 1 Ukrainian deaf woman (henceforth D_L2_Al), with severe hearing impairment, aged 42, living in Italy for 8 years, having a hearing son and employed as healthcare worker. She is proficient both in LIS and both in signed and vocal Ukrainian.

- 1 Japanese female learner of L2 Italian (henceforth H_L2_Ay), who had been recognized as 'impaired' or 'synthesized' in a previous, preliminary investigation (Pellegrino, Caruso & De Meo, 2012).

Additionally, two voices make up the control group in order to:

- a. identify the presence of any transfer phenomena in the vocal L2 productions from the language of the first logopaedic treatment of the deaf foreigner;
- b. assess the difference between the speech of a native Italian and that of a severe deaf.

The control group is thus composed by:

- 1 Ukrainian female, who is an L2 Italian learner (henceforth H_L2_Ha). The spectro-acoustic analysis of her voice has been compared to that of the Ukrainian deaf in order to verify point a) of the previous list.
- 1 native Italian speaker (henceforth H_L1_El), offering the necessary comparisons in order to estimate point b).

All foreign readers involved in the test have an elementary level of linguistic competence in Italian.

As for the listeners, the group is composed of 17 native Italians, aged between 20 and 28. Six of them declared to be unexperienced with any kind of impaired speech, six instead said to be familiar with deaf people's speech, and six others with the speech disorders caused by Down's Syndrome, Parkinson's Disease, stuttering, and schizophrenia.

2.2 Materials and Methods

In order to prevent different factors that might have biased the results, participants were involved in a read speech task. This made comparisons possible, since many voices were collected for the oral corpus and they were extremely different among each other. Moreover, the complexity of the text used as a stimulus was carefully controlled in terms of both word length, syntactic and informative structure. Participants were thus given as a reading material a simplified short excerpt from an article of fashion (see below):

G-Star è uno stilista famoso e fa jeans molto belli.

Oggi a Milano G-Star vende i suoi jeans a solo 20 euro, un prezzo molto basso.

L'attrice Laura Chiatti fa pubblicità ai jeans di G-Star nel negozio di Milano.

The speakers' voices recordings were taken in single sessions in an anechoic chamber, at 44.100 Hz sampling rate, at the Language Center of 'L'Orientale' University of Naples.

After that, the 17 native subjects listened, in a randomized order, to the single excerpts of the read speech uttered by the seven readers. Each speech sample was then rated for 'accentedness', 'comprehensibility' and 'speech'. Before taking the per-

ception test, the listeners were explained that 'accentedness' serves to assess whether the speakers' accent is judged as different from that of the L1 community (Derwing, Munro, 2005). 'Comprehensibility', on the other hand, expresses the effort required from the listener in order to understand the non-native speaker utterances (Warren, Elgort & Crabbe, 2009).

For the ratings regarding accentness, listeners had to judge if the speakers sounded native or non-native like; whereas 'comprehensibility' had to be rated on a five-point scale (0 = 'poor'; 5 = 'very good'). Although in studies on foreign accentedness, the term 'comprehensibility' often overlaps with that of 'intelligibility' and the two measures seem to be very well correlated (Munro, Derwing, 1999), in this study the main focus was on 'comprehensibility' rather than 'intelligibility'. This choice was made in accordance to the study aims and materials: 'comprehensibility' is usually assessed with human judgments and relies on prosodic cues, while 'intelligibility' is assessed with transcription tasks and mainly depends on segments (Munro, Derwing, 1995; 1999; 2001).

With respect to the evaluation of the 'kind of speech' parameter, instead, listeners could choose among four options: 'normal', 'pathological', 'synthesized' or 'other'.

2.3 Results

2.3.1 The perception test

Results of the perception test show for the 'accentedness' parameter that Italian listeners, regardless of their experience with impaired speech, succeed in distinguishing between native and non-native speech, when the utterances are produced by 'non-pathological', and 'non-synthesized' speakers (table 1), who in the following pages will be referred to as 'normal'.

Table 1 - *Rates for accentedness*

Subjects	Native	Non Native
	Nr. listeners (%)	Nr. listeners (%)
H_L1_El	17 (100%)	
H_L2_Ay		17 (100%)
H_L2_Ha		17 (100%)
D_L1_Fr	16 (94%)	1 (6%)
D_L1_Mo	15 (88%)	2 (12%)
D_L2_Al	15 (88%)	2 (12%)

In fact, all listeners recognize which one is the Italian native speaker's voice, but only the hearing foreigners are identified as non-native. On the contrary, all deaf participants are considered as 'native', and no exception is made for the Ukrainian deaf women (D_L2_Al).

However, the fact that the speakers' provenience is unrecognizable when they are affected by a severe hearing loss is not surprising. This result is in fact compatible with the outcomes of another study by Pellegrino, Caruso, De Meo (in press) that demonstrates

how difficult it is to find strict correlations between prosodic profiles and deaf people's origin, degree of hearing loss and quantity of exposure to the vocal language. More precisely, the research has highlighted that fluency is the most unpredictable feature of deaf people's speech, since every deaf subject develops a personal pausing strategy in order to overcome articulatory difficulties and lack of self-control on his/her own utterances.

Additionally, the complexity of the 'non-native speech' category emerges also with respect to the assessment of another parameter, such as 'nature of speech'. When, in fact, the Italian listeners were asked to judge if the speaker they had been listening to was 'normal', 'pathological' or 'synthesized', they were able to recognize the hearing native as a 'normal' speaker and the three deaf women as 'pathological', while the hearing L2 speakers received contradictory rates (table 2). The Ukrainian woman (H_L2_Ha) was judged in fact as 'normal' by 11 listeners, and 'pathological' by six listeners. Even more ambiguous is the case of the hearing Japanese (H_L2_Ay), whose voice was considered 'pathological' by almost half of the listeners (8), 'normal' and 'synthesized' by four listeners each; while one listener considered her as both synthesized and pathological.

Table 2 - *Rates for kind of speech*

Subjects	Normal	Pathological	Synthesized	Synthesized/Pathological
H_L1_El	17 (100%)			
H_L2_Ay	4 (23.5%)	8 (47%)	4 (23.5%)	1 (6%)
H_L2_Ha	11 (64%)	6 (36%)		
D_L1_Fr	17 (100%)			
D_L1_Mo	17 (100%)			
D_L2_Al	17 (100%)			

Lastly, speech samples were rated on 'comprehensibility' (table 3).

Table 3 - *Rates for comprehensibility*

Subjects	Average rate (St. Dev.)
H_L1_El	5 (\pm 0)
H_L2_Ha	4.3 (\pm 0.46)
H_L2_Ay	2.5 (\pm 0.9)
D_L2_Al	2.3 (\pm 0.8)
D_L1_Fr	2.3 (\pm 0.7)
D_L1_Mo	1.4 (\pm 0.5)

As expected, the hearing native speaker (H_L1_El) was found very comprehensible (5), on the contrary, the three deaf subject were the least intelligible among the participants, with grades ranging from 2.3 of D_L2_Al to 1.4 of D_L2_Mo. Again, the more surprising results are those concerning the hearing L2 speakers. Although both of them were considered as non-native speakers, the Ukrainian hearing woman (H_L2_Ha) achieves a high level of comprehensibility (4.3), whereas the Japanese speaker (H_L2_Ay) gets scores (2.5) that are in line with those received by the deaf

subjects. One can thus infer that the scarce comprehensibility of this speaker is responsible for the fact that many native listeners (9 out of 17) have considered her as an impaired speaker.

2.4 Acoustic analysis

The next step of the research was to spectro-acoustically analyze the spoken performance of the six subjects in order to verify whether the listeners' rates on 'nature of speech', 'accentedness' and 'comprehensibility' could correlate with some shared prosodic profile.

To the purpose, using the software Praat (Boersma, 2001), the speech corpus was segmented into 2 different tiers: CV segments, syllables. CV boundaries were identified by visual inspection of speech spectrograms and waveforms. The segmentation of glides followed the rules adopted by Ramus et al. (1999): [w] and [j] were treated as consonants and the boundary was placed between the approximant and the vowel. Diphthongs were segmented in one or two vowels depending on the spectro-acoustic characteristics of the tract. If both vowels presented a specific steady-state formant pattern, the diphthong was divided into two vowels; otherwise, it was treated as a single interval. Voice Onset Time in stop consonants was included within the consonantal interval. In vowel+nasal consonant sequence, the nasalized portion of the vowel was assigned to the vowel. In case of initial voiced stop consonants, the first boundary corresponded to the onset of glottal pulses. Post-pausal voiceless plosives were assigned a duration equal to the mean value of single plosives in the same utterance. Respiratory and syntactic pauses were labeled as silent pauses (SP). Different kinds of disfluencies (laryngealizations, nasalizations, vocalizations, etc.) were treated as filled pauses (FP). The duration of all segments was extracted using a Praat script. Additionally, for every speech chain, defined as the portion of an utterance within two silences (Pettorino, Giannini, 2010), the lowest and highest f_0 values were manually measured.

On the basis of these measures, and following the Prosody–Voice Screening Profile for pathological speech by Schriberg et al. (1992), the following rhythmic-prosodic parameters were calculated:

- Articulation rate (AR), i.e. the ratio between the number of syllables really uttered and phonation time (syll/s).
- Speech rate (SR), i.e. the ratio between the number of syllables really uttered and the total time of the utterance time (syll/s).
- Fluency (F), i.e. the ratio between the number of syllables and the number of speech chains (syll/SC).
- Tonal range (TR), i.e. the interval between the f_0 minum and maximum, measured in semitones.

At segmental level, contrastive analyses on vowel formant patterns were carried out. From the CV tier, the steady state formant pattern of stressed and unstressed vocalic segments, without transitions to and from adjacent sounds, was identified. Then,

the frequency of F1 and F2 were measured manually. The raw vowel measurements were normalized through the web-based software package NORM, using Neary 2 (formant extrinsic) as normalization method.

Consonantal pronunciation mistakes were also taken into consideration. Given deaf speakers' inability to rely on acoustic cues in the process of speech perception, the first objective was to verify whether deaf speakers were able to reproduce correctly the information mainly provided by the visual system (place of articulation), and the cues derived primarily from the acoustic signal (manner of articulation and voicing) (McGurk, MacDonald, 1976). A further goal was to investigate whether deaf talkers would produce errors that are of a different type from those made by the hearing speakers.

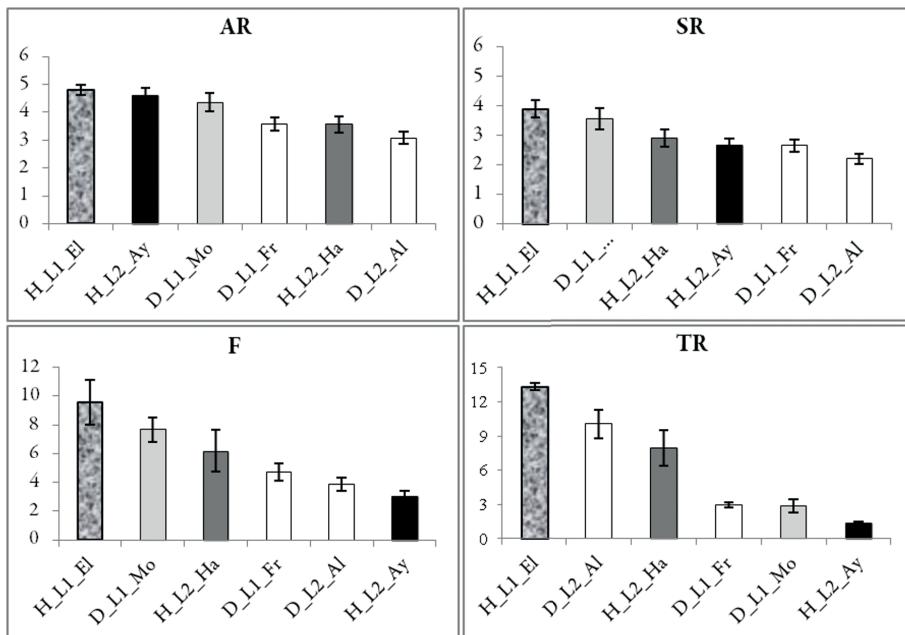
2.5 Acoustic correlates of accentedness, perceived nature of speech, and comprehensibility

2.5.1 Suprasegmental level

The results of spectro-acoustic analysis carried out on the corpus shows no strict relationship between the native Italian listeners' judgements and the prosodic profiles of the different speakers' groups. This means that the suprasegmental clues alone are not robust enough to explain the assessment results of the perceptive test. However, this is still in line with the results of the study by Pellegrino, Caruso, De Meo (in press) showing that the spoken productions of deaf people can have the same prosodic features, regardless of the speakers' geographical origin, level of hearing impairment and the amount of exposure to the vocal Italian. Moreover, the present study provides evidence of an imperfect correspondence between prosodic and perceptual data that seems to indicate the stronger relevance of some other segmental clues for specific categories of judgments, like 'native/non-native' and 'pathological' (figs. 1-4).

Among the clearest mismatches between acoustic and perceptual data one can find the case of the Italian deaf D_L1_Mo (light grey column in graphs above). The Italian listeners judged her spoken productions as an example of scarcely comprehensible, pathological speech even though her prosodic profile is similar to the one of the native hearing speaker (H_L1_El, granite column) in terms of Articulation rate, Speech Rate and Fluency. In addition, the Tonal range is the only parameter that distinguishes this same Italian deaf (D_L1_Mo) from the native speaker and equates her to the other speakers who receive the same grades for comprehensibility and speech type (D_L1_Fr and H_L2_Ay).

Figures 1-4 - Cross-comparison of the oral corpus per AR, SR, F and TR



The hearing Ukrainian speaker represents, instead, a different case (H_L2_Ha, dark grey column). Her lower values of Articulation Rate, Speech Rate, Fluency and Tonal Range than the hearing native Italian (H_L1_El) seem not to increase the listeners' effort to understand her speech, but they are probably responsible for the judgement of 'non-nativeness'. The majority of the listeners, in fact, considers her speech as 'highly comprehensible' and 'non-native'. On the other hand, her Fluency values are similar to those of one deaf Italian, D_L1_Mo, and her Tonal range is in line with the Ukrainian deaf woman (D_L2_Al), two clues that have probably biased the correct recognition of her voice by some listeners. Six out of 17 of them have in fact considered her voice as 'pathological'.

Completely different is the case of the hearing Japanese (H_L2_Ay, black column). Her mixed prosodic profile might have confused the listeners, who assigned very low comprehensibility ratings and chose extreme options when asked to rate the kind of speech they had been listening to. She has in fact some vocal traits in common with all the groups of participants. For the Articulation Rate, she is similar to the native speaker and to one Italian deaf, D_L1_Mo. In terms of Speech Rate, instead, she resembles the Ukrainian speaker and the other Italian deaf (D_L1_Fr). In terms of Fluency, her speech flow is as fragmented as the one produced by the Ukrainian deaf (D_L2_Al). What on the contrary differentiates her from the rest of the group is the narrow interval between f_0 maximum and minimum. This type of non-native intonation, therefore, might have been perceived as unnatural as the flat and artificial sounding intonation produced by text-to-speech systems (Clark,

1999). This can explain the 'synthesized speech' judgments expressed by four listeners.

2.5.2 The segmental level

At segmental level, contrastive analyses on vowel formant patterns were conducted. Before starting to discuss the results, it is important to underline that the text used for the read speech task was not originally created to allow for a punctual analysis of the Italian vowel system. The objective, instead, was to look at the effect of the hearing impairment and non-nativeness on the size of vowel areas used by the six speakers reading the same text. The passage includes all seven phonetic vowels, though they occur with a different frequency (table 4).

Table 4 - *Vowel frequency*

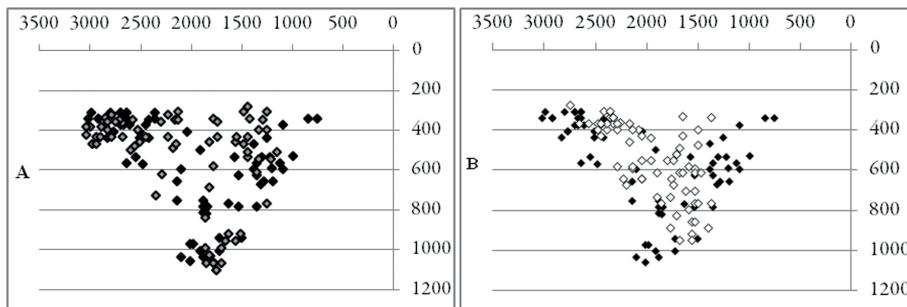
[i]	[ɛ]	[e]	[a]	[ɔ]	[o]	[u]
29%	8%	11%	26%	3%	20%	3%

The diagrams in figs 5A-7 plot the normalized F1/F2 values of the vowels produced by the Italian hearing speaker in comparison to the ones of:

1. Italian deaf readers (figs. 5A-5B).
2. Ukrainian deaf and hearing readers (figs. 6A-6B).
3. Japanese and Ukrainian hearing readers (fig. 7).

As regards the first comparison (figs. 5A-5B), it is interesting to point out that even in the case of vowel articulation, the Italian deaf speaker D_L1_Mo is very close to the native model, especially for the area of front-close and open vowels (fig. 5A). The same does not apply to her sister.

Figures 5 A-B - *Vowel areas A of the hearing native speaker (black rhombi) in comparison with D_L1_Mo (grey rhombi in fig. A, on the left) and D_L2_Fr (white rhombi in fig. B, on the right)*



In line with the results of acoustic and kinematic studies on vowel articulation by deaf and normal speakers (see Pratt, Tye-Murray, 2011 for a review), D_L1_Fr presents vowel space reduction with a general centralization pattern in the degrees of

fronting and opening. This means that D_L1_Fr differs from the native speaker as well as from her sister not only on the suprasegmental level, but also on the degree of variability in tongue and shape position across vowels. Tye-Murray (1991) had already shown that deaf subjects are not capable of producing vowel distinct movement trajectories, unlike normal subjects whose tongue movements vary in view of the vowel to be uttered.

Figures 6 A-B - *Vowel areas of the hearing native speaker (black rhombi) in comparison with D_L2_Al (white squares, in fig. A, on the left) and H_L2_HA (grey x marks, in fig. B, on the right)*

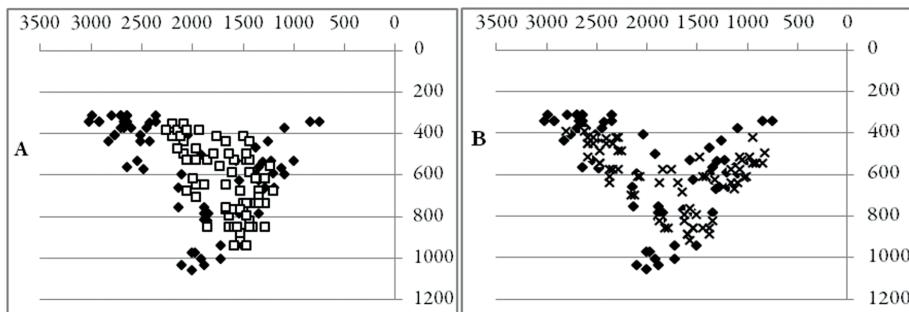
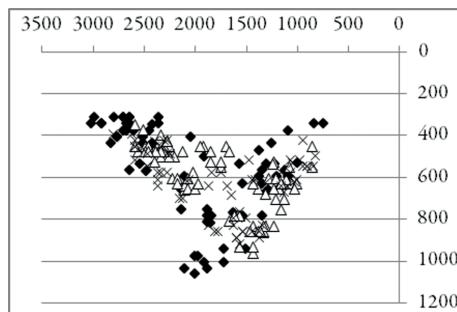


Figure 7 - *Vowel areas of the hearing native speaker (black rhombi) in comparison with H_L2_HA (grey x marks) and H_L2_Ay (white triangles)*



The centralization pattern is even more evident in the case of the Ukrainian deaf speaker (D_L2_Al) (fig. 6A). Probably due to reduced tongue mobility, she succeeds in uttering only close mid and open mid vowels, but is not able to distinguish between anterior and posterior vowels. The drastic centralization trend observed in D_L2_Al cannot be explained only as the effect of the language of the logopedic treatment she received, namely the vocal Ukrainian. The VSA of the hearing Ukrainian reader (H_L2_Ha) (fig. 6B), similarly to the one of the Japanese (H_L2_Ay) (fig. 7), is more reduced than those of the native speakers, but she was at least able to produce front and back vowels. Behavioral data similar to those of the Ukrainian deaf speaker were already found by Stein (1980), who demonstrated that hearing impaired children were able to produce tongue height distinctions between

high and mid vowels, but they were unable to differentiate front and back vowels. All this things considered, one can affirm that the trend to vowel centralization of the deaf speakers' voices analyzed so far is partially imputable to the sensorial impairment, and partially to the fact that they are L2 speakers.

2.5.3 The segmental level: consonants

In addition to the vowel areas, attention was also paid on consonantal articulation, since a sizeable body of research has demonstrated that deaf subjects produce a "myriad of consonantal errors" affecting speech intelligibility (Pratt, Tye-Murray, 2011: 208). This further level of analysis had a twofold objective: on the one hand, it was intended to verify whether Italian and foreign deaf speakers commit the same kind of mistakes while reading aloud in Italian. On the other hand, it was specifically aimed at ascertaining further reasons for the judgment expressed by the group of listeners for the speech produced by the Japanese, who was considered as an 'impaired' speaker.

The pronunciation mistakes were analyzed both quantitatively and qualitatively. For each speaker, the number of misarticulations, with respect to the native model, was calculated. We also considered the kind of segmental errors committed and the sound category mostly affected by mispronunciations.

The results of quantitative analysis have shown that the speakers that score limited comprehensibility (the 3 deaf participants and the Japanese learner) were also the ones producing consonantal errors (table 5). Out of the 83 consonantal sounds uttered by the native speaker, the deaf Ukrainian (D_L2_Al) gets 10 sounds wrong, D_L1_Mo and the Japanese speaker err 15 consonants, D_L1_Fr 23. No segmental errors were instead produced by the hearing Ukrainian speaker, whose spoken productions were judged as highly comprehensible.

Table 5 - *Number and Percentage of consonantal mistakes per speaker and comprehensibility level*

Subjects	Target number of consonants	N. of mispronunciations per speaker (%)	Comprehensibility level
H_L2_Ha	83	-	4.3 (± 0.46)
D_L2_Al	83	10 (12%)	2.3 (± 0.8)
D_L1_Mo	83	15 (18%)	1.4 (± 0.5)
H_L2_Ay	83	15 (18%)	2.5 (± 0.9)
D_L1_Fr	83	23 (28%)	2.3 (± 0.7)

The errors produced by the least comprehensible speakers were then classified according to the commonest taxonomy available in the literature for the errors produced by deaf and non-native speakers. Therefore, accent shift, insertions, omissions, substitutions, errors in consonantal clusters, voiced-voiceless confusion were considered (Major, 2001; Pratt, Tye-Murray, 2011; Tye-Murray et al., 1996). As table 6 shows, the most recurrent type of mistake is phone substitution rather than

omission, as Pratt and Tye-Murray (2011) on the contrary report. In descending order, there are voiced-voiceless confusions, omissions and lastly insertions, which have considerably lower errors percentage.

Table 6 - *Percentage of consonantal mistake type*

<i>insertions</i>	<i>omissions</i>	<i>substitutions</i>	<i>voiced-voiceless confusion</i>
11%	14%	51%	24%

With respect to the distribution of the error type across speakers, fig. 8 shows that all the speakers produce substitution errors. To this typology, the Japanese speaker adds also insertions and deletions; the Italian deaf subjects add only omissions, while the Ukrainian only insertions.

Figure 8 - *Number of errors for speech type and speaker*

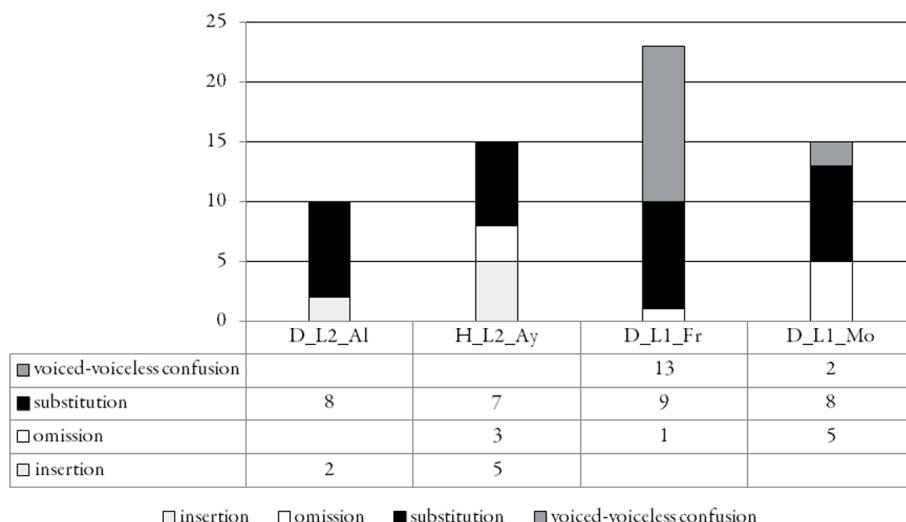
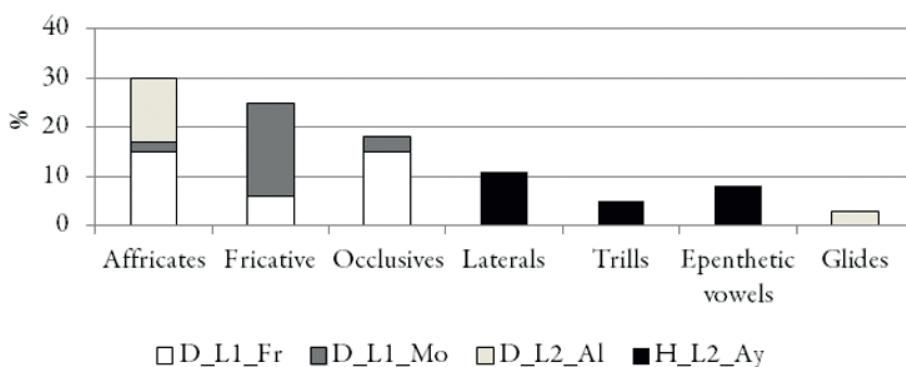


Figure 9 - *Mistaken sound category*



It is interesting to underline that, despite the across-speakers similarity in the type of mistakes made, what differentiates the hearing Japanese from all deaf speakers is the mistaken sound category (fig. 9)¹.

The Japanese, in fact, errs particularly with liquids. For example she tends to delete the alveolar trill in coda position (i.e. 'G-star' uttered like [dZ"ist" a]; |dZ| i| s| t| a|- r|); she substitutes the long alveolar lateral in intervocalic position with the voiced alveolar plosive (i. e. 'belli' uttered like [b"eddi]; |b| e| dd| i|) and single alveolar lateral with alveolar trill (i.e. 'solo' uttered like [s"oro] |s| o| r|o|). She usually adds epenthetic central vowels in loan words ending with consonant (i.e. 'jeans' uttered like [dZ"ins@]; |dZ| i| n|s| +@|) and in sequences involving complex consonantal clusters (i.e. 'pubblicità' uttered like [pub@ditSit"a]; |p| u| b| +@|d|i| tS|i|t|a|) in order to simplify the syllable structure.

As regards deaf speakers, instead, regardless of their provenience, the sounds mostly affected by errors are the plosives, the fricatives and the affricates both in initial and in intervocalic positions. In the examined corpus, there is evidence of:

- shift in manner of articulation ts > s (es. 'prettso' uttered like [pr"esso]; |p| r| e| ss| o|; tʃ > f ('attrice' uttered like [attr"iSe]; |a| tt| r| i| S| e|);
- shift in place of articulation: st>sk ('stilista' uttered like [stil "iska]; |s| t| i| l| i| s| t| a|);
- devoicing: b > p ('basso' uttered like [b_0"asso]; |b_0| a| s| s| o|); dZ > tS ('oggi' uttered like ["oddZ_0i]; |o| ddZ_0i|); v > f ('venti' uttered like [v_0"enti], [|v_0| e|n| t| i|);
- omission of fricative sounds ('stilista' uttered like [til "itta]; | - s| t| i| l| i| tt| a|).

Considering the latter category of error, this was produced only by one Italian deaf speaker, D_L1_Mo, whose hypernasal speech reflects the inability to close the soft palate completely, and thus to produce fricatives. A more detailed analysis of the kinds of errors listed above seems to suggest that deaf people do not find it hard to produce complete closure. The plosives indeed are primarily object of a shift in the place of articulation and devoicing. On the contrary, all the errors related to a shift in the manner of articulation involve sounds with a mixed articulatory mechanics. The affricates were substituted by the corresponding fricatives, probably because the subjects find it difficult to coordinate two consecutive movements in uttering a single sound. Rather than fully stopping the air from leaving the vocal tract, then releasing it through a constricted opening, they produce only the second part of the movement, resulting in a fricative.

3. Conclusions

Summing up the results achieved so far, it is important to underline, in the first place, the difficulties encountered by the native listeners to distinguish the spoken productions of deaf speakers with reference to the native/non-native parameter.

¹ The consonantal mistakes were transcribed and annotated according to the rules adopted for the corpus CLIPS (see Savy, 2005).

Despite the spectro-acoustical evidence of some phonetic transfers from the language of the logopaedic treatment, these features are overheard by native listeners, who fail to distinguish between native and non-native Italian deaf speakers.

Similarly, the mixed prosodic profile of the Japanese speaker and the high number of segmental mistakes she makes might have confused the native listeners, who judged her oral productions similarly to those of deaf speakers, namely as 'pathological' and 'scarcely comprehensible'. On the contrary, her VSA and consonantal misarticulations differentiate her from the deaf speakers and thus may have caused the different judgment concerning her 'non-nativity' to the Italian language. In fact, when the Japanese speaker reads in Italian, she errs mainly with liquids and with complex consonantal clusters, while the sounds that are more affected by errors in deaf speech are the plosives, fricatives and affricates, both in initial and in intervocalic positions. These features derive from the difficulties to control voicing and the degree of constriction that deaf speakers commonly have.

All this things considered, one can underline the relevance of the data collected so far for better outlining the types of deaf speakers' linguistic inabilities and for offering further evidence of the wide-ranging variations admitted by the category of 'non-native Italian speech'.

Bibliografia

BOERSMA, P. (2001). Praat, a system for doing phonetics by computer. In *Glot International*, 5, 341-345.

BONETTI, L., UTOVIĆ, V. & DULČIĆ, A. (2008). Prosody in Speech Produced by Deaf Persons. In *Croatian Review of Rehabilitation Research*, 44, 1-13.

BOONE, D.R. (1971). *The voice and voice therapy*. Englewood Cliffs, NJ: Prentice-Hall.

CALVERT, D.R. (1962). Deaf voice quality: A preliminary investigation. In *The Volta Review*, 64, 401-403.

CLARK, R.A.J. (1999). Using Prosodic Structure to Improve Pitch Range Variation in Text to Speech Synthesis. In OHALA, J.J. (Ed.), *Proceedings of the XIVth International Congress of Phonetic Sciences*, vol. 1. San Francisco: University of California Press, 69-72.

DERWING, T.M., MUNRO, M.J. (2005). Second language accent and pronunciation teaching: A research-based approach. In *TESOL Quarterly*, 39, 379-397.

FINKELSTEIN, Y., BAR-ZIV, J., NACHMANI, A., BERGER, G. & OPHIR, D. (1993). Peritonsillar abscess as a cause of transient velopharyngeal insufficiency. In *Cleft Palate-Craniofacial Journal*, 30, 421-428.

LENDEN, J.M., FLIPSEN, P.J. (2007). Prosody and voice characteristics of children with cochlear implants. In *Journal of Communication Disorders*, 40, 66-81.

MAJOR, R.C. (2001). *Foreign Accent: The Ontogeny and Philogeny of Second Language Phonology*. Mahwah, NJ: Erlbaum.

McGURK, H., MACDONALD, J. (1976). Hearing lips and seeing voices. In *Nature*, 264, 746-748.

MUNRO, M.J., DERWING, T.M. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. In *Language and Speech*, 38, 289-306.

MUNRO, M.J., DERWING, T.M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. In *Language Learning*, 49 (supp. 1), 285-310.

MUNRO, M.J., DERWING, T.M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech. In *Studies in Second Language Acquisition*, 23, 451-568.

NICKERSON, R. (1975). Characteristics of the speech of deaf persons. In *The Volta Review*, 77, 342-362.

PELLEGRINO, E., CARUSO, V. (2012). "And Nevertheless they do speak Italian". Prosodic explorations in speech of deaf immigrants. In DE MEO, A. & PETTORINO, M. (Eds.), *Prosodic and Rhythmic Aspects of L2 Acquisition. The case of Italian*. New Castle upon Tyne: Cambridge Scholar Publishing, 83-102.

PELLEGRINO, E., CARUSO, V. & DE MEO, A (2012). Other voices: prosodic discriminative features of foreign and impaired speech, oral presentation to *Prosody in Typical and Atypical Populations Congress*, University of Reading, UK, 3-4 September.

PELLEGRINO, E., CARUSO, V. & DE MEO, A (in press). La competenza prosodica nell'italiano parlato di sordi nativi e stranieri. In *Il linguaggio disturbato. Modelli, strumenti, dati empirici*, Bari, Università degli studi di Bari.

PELLEGRINO, E., DE MEO, A. & CARUSO, V. (2015). "Chi compie l'azione?" L'applicazione del Competition Model su sordi italiani. In FAVILLA, M.E., NUZZO, E. (Eds.), *Grammatica applicata: apprendimento, insegnamento, patologie*. Milano: studi AItLA, 165-178.

PETTORINO, M., GIANNINI, A. (2010). Il parlato dei mass media. In PETTORINO, M., GIANNINI, A. & DOVETTO, F. (Eds.), *La Comunicazione Parlata 3*, Napoli, OPAR, Università degli studi di Napoli 'L'Orientale', 71-83.

PRATT, S.R., TYE-MURRAY, N. (2011). Speech Impairment Secondary to Hearing Loss. In MC NEIL, M.R. (Ed.), *Clinical Management of Sensorimotor Speech Disorders*. New York: Thieme, 204-234.

RAMUS, F., NESPOR, M. & MEHLER, J. (1999). Correlates of linguistic rhythm in the speech signal. In *Cognition*, 72, 265-292.

SAVY, R. (2005). Specifiche per l'etichettatura dei livelli segmentali. In ALBANO LEONI, F., GIORDANO, R. (Eds.), *Italiano Parlato. Analisi di un dialogo*. Napoli: Liguori.

SHRIBERG, L.D., KWIATKOWSKI, J., RASMUSSEN, C., GREGORY, L.L. & MILLER, J.F. (1992). The prosody-voice screening profile (pvsp): psychometric data and reference information for children. Phonology Project Technical Report n. 1, Waisman Center on Mental Retardation and Human Development, University of Wisconsin-Madison, 1-54.

STATHOPOULOS, E., DUCHAN, J., SONNENMEIER, R.M. & BRUCE, N.V. (1986). Intonation and pausing in deaf speech. In *Folia Phoniatrica*, 38, 1-12.

STEIN, D. (1980). A study of articulatory characteristics of deaf talkers, Doctoral dissertation, University of Iowa.

TYE-MURRAY, N., SPENCER, L., BEDIA, E.G., WOODWORTH, G. (1996). Differences in children's sound production when speaking with a cochlear implant turned on and turned off. In *Journal of speech and hearing research*, 39, 604-10.

TYE-MURRAY, N. (1991). The establishment of open articulatory postures by deaf and hearing talkers. In *Journal of Speech and Hearing Research*, 34, 453-459.

WARREN, P., ELGORT, I. & CRABBE, D. (2009). Comprehensibility and prosody ratings for pronunciation software development. In *Language Learning & Technology*, 13, 87-102.