INTERSPEECH 2021
30 August – 3 September, 2021, Brno, Czechia

# Phonetic complexity, speech accuracy and intelligibility assessment of Italian dysarthric speech

*Barbara Gili Fivela*[1], *Vincenzo Sallustio*[2], *Silvia Pede*[2], *Danilo Patrocinio*[3]

[1]University of Salento & CRIL-DReAM, Lecce, Italy
[2]Phoniatrics and Communicative Disorders Rehabilitation Center, Dept. Rehabilitation, ASL Lecce, Italy
[3]Università Cattolica del Sacro Cuore, Rome, Italy

barbara.gili@unisalento.it, v.sallustio@alice.it, pedesilvia@libero.it, danilopatrocinio@icloud.com

## Abstract

Intelligibility is the degree to which the speech of a person may be understood by a listener, and is related to functional limitation and disability. In protocols for the clinical assessment of dysarthria, intelligibility checks are included, as well as evaluations of speech accuracy, which is more directly related to the disease severity. However, both evaluations are usually based on subjective ratings.

Aim of this work is checking the correlation between intelligibility judgements, subjectively assigned as it may be the case in clinical procedures, and acoustic measures related to linguistically contrasting units. Two novelties characterize this work: a) acoustic measurements considered in the paper relate to both segments (vowel and consonants) and prosodic-intonational phonological events (e.g., pitch accents), that is linguistically relevant speech units; b) contexts of increasing phonetic-phonological complexity are considered, in order for the phonetic characteristics to challenge production accuracy, possibly affecting the realization of phonological features and intelligibility. Increasing complexity is expected to challenge intelligibility indeed and to have an impact on the correlation between intelligibility rates and acoustic measures. Results are preliminary, but confirm both 1) the correlation between acoustic measures of linguistically relevant events and speech intelligibility, as for both the segmental and the prosodic-intonational level, and 2) the role of increasing phonetic-phonological complexity in enhancing the above mentioned correlation.

**Index Terms**: Dysarthria, Parkinson's Disease, phonetic complexity, Italian, consonant clusters, prosody and intonation

## 1. Introduction

### 1.1. Intelligibility assessment, speech accuracy and clinical practice

Speech intelligibility may be defined as the degree to which the speech of a person may be understood by a listener [1, 2], and is only partially related to production accuracy. Various factors besides the actual realization of speech gestures play a role in affecting speech intelligibility, such as the contextual information and the knowledge of the speaker or the task [3, 4, 5, 6]. Nevertheless, speech intelligibility is related to functional limitation and disability, it is measured as an index of motor speech disorder severity and it is used "for documenting and monitoring change in functional speech performance" [7].

Both subjective and objective measures of intelligibility have been proposed in the literature. As for the former, various methods are described, such as calculating the percentage of words listeners think they understood, proposing visual analogue scales for listeners to mark the extent to which the speech is intelligible, or using Direct Magnitude Estimations, obtained by asking raters to rank the intelligibility of a recording against that of other recordings [8, 9]. As for objective measures, descriptions in the literature report on listener transcriptions of words heard and calculation of the percentage of words correctly understood, or listener selections of the word heard from a list [10, 8]. On the one hand, subjective measures are implicitly not completely reliable and, on the other hand, objective measures require procedures that may be difficult to adopt in clinical practice, besides that their efficacy varies depending on factors such as the speech task and the familiarity between the speaker and the listener [3, 4, 5, 6].

In protocols adopted for clinical assessments of dysarthria, subjective checks are usually included for practical reasons, but intelligibility judgements are only part of the ratings collected. For instance, clinical assessments by means of the Robertson's Dysarthria Profile (RDP's [11, 12]) also require an evaluation of aspects related to speech accuracy, that is to the precision in the realization and coordination of speech gestures (which, in comparison to intelligibility, may be more directly related to a motor speech disorder such as dysarthria). In this respect, various features are considered regarding diadocinesis (among which, rapid repetition of open syllables involving various consonants), the articulation of consonants, vowels, consonant clusters in real polisyllabic words and sentences, as well as features regarding the production of prosody (such as rhythm, speech rate, intonation, also by asking to imitate different patterns of sentence accentuation).

However, while attention is payed to the accurate production and the intelligibility of linguistically contrasting units at the segmental level, a more superficial evaluation of linguistic prosody is performed (see the evaluation of "Ability to maintain appropriate intonation" in RDP's, with no clear definition of what appropriate is). Nonetheless, we know that prosody plays relevant linguistic functions such as accentuation, phrasing, focus, and in some languages, such as Italian, it may even convey the information related to the modality of the sentence.

Copyright © 2021 ISCA  http://dx.doi.org/10.21437/Interspeech.2021-1862

### 1.2. Acoustic predictors of intelligibility

A number of works in the literature focus on the correlation between acoustic measures and speech intelligibility, and the identification of the best acoustic predictors of intelligibility. The assumption is that a strong correlation exists between intelligibility and the phonetic characteristics carried by the acoustic signal. Despite the fact that the phonetic analysis of the acoustic signal offers a number of information on speech articulation, and therefore on production accuracy, such studies mainly aim at showing the relation between objective acoustic measurements and intelligibility or at offering detailed phonetic analyses in order to identify the features that account for intelligibility [1].

Various measures have been found to be good predictors of intelligibility. For instance, measures of global speech timing (such as speech rate and articulation rate), measures of sentence-level fundamental frequency (F0), or sound pressure level, as well as measures directly related to segment (such as vowel space area, second formant slope or second formant interquartile range) have been reported to predict intelligibility (as for other measures, see [13]). Nevertheless, the relation between acoustic measures and intelligibility appears to be complex, due to heterogeneous results (e.g. on global speech timing) and to the high amount of variation observed [14, 15]. Concerning prosody, the attention is usually payed to global aspects (e.g. F0 slope, peak intensity [16, 17]), with no specific reference of those changes that may be directly related to phonological units and, therefore, may be more tightly related to linguistic messages and speech intelligibility. On the contrary, it is well known that, e.g., the timing of tonal events with respect to syllables composing the phonic chain is crucial to distinguish phonologically relevant intonational events [18, 19].

The point of view adopted here is that there is a need to keep in mind the linguistic contrasts and their functional load, as for both segments and prosody-intonation. Working on the correlation between acoustic measures and intelligibility may then be promising because of the double outcome that it may offer: getting objective information on accuracy *per se*, through acoustics, and offering objective measures to correlate to intelligibility ratings (and possibly, at a later stage, to predict them).

### 1.3. Phonetic and phonological complexity

Challenging speech accuracy may be a way to enrich the acoustic signal of information on possible difficulties in the realization of speech gestures, possibly increasing the correlation between values of acoustic measures and speech intelligibility ratings. A way to challenge accuracy may involve the production of increasingly complex items. However, defining what complexity is may be difficult, and also trying to tease apart what phonological and phonetic complexity are may be not easy, or even not fruitful [20].

On the one hand, it is possible to interpret the higher complexity as higher phonetic difficulty (e.g., calculating the number of articulators involved, the characteristics of their phasing). On the other hand, the need to preserve distinctiveness should be taken into account, bringing into play more phonological issues. In this respect, for instance, higher complexity may be related to phonological specifications, that is marks, or markedness in Trubetzkoy's terms, or to a higher number of linguistic features (e.g., [21], and see also the complexity scores assigned to consonants in [22]). Further, the balancing between the expected economy of articulatory effort and the need to preserve distinctiveness may be complex *per se*.

In any case, complexity may regard different structures, having an impact on both segments and prosody. For instance, a syllable composed by a Consonant and a Vowel (CV) is widely recognized as a non-complex structure, the acquisition of a two-member cluster is marked, and the tendency to simplify clusters, restoring simpler CV syllables via deletion or vowel epenthesis is attested [23, 24] (in fact, CV syllables are preferred in both typical and atypical speech acquisition, in aphasia [25] and adult degenerative speech [26]). Along similar lines, among tonal structures, rises are reported to be more complex than falls, and events that include both a rising and a falling component are the most complex (e.g. [27]).

In a first attempt to investigate the correlation between intelligibility judgements and acoustic measurements in Italian dysarthric speech, the work described in this paper regards speech material of increasing complexity in relation to the production of both segments and prosody, without trying to tease phonetic and phonological complexity apart.

## 2. Goal and hypotheses

The main aim of the work described here is to check the correlation between intelligibility judgements and acoustic measurements on both segments and prosodic-intonational events that play a linguistic role. A subgoal is investigating such correlation in contexts of increasing phonetic-phonological complexity, that is taking into account changes in phonetic characteristics that challenge the identity of relevant phonological units, and that therefore may more probably challenge intelligibility.

The hypotheses are that: 1) acoustic measures of linguistically relevant events are correlated with speech intelligibility at both segmental and prosodic-intonational level; 2) increasing complexity represents a challenge for accuracy and intelligibility, and enhances the correlation of acoustic measures with intelligibility judgements.

## 3. Methods

The corpus used for the analyses was composed by two blocks, one including single words and one including sentences. Block I was composed by 30 initially stressed Italian disyllabic words proposed in isolation, whose initial syllable showed a minimal CV or a maximal sC(C)V(C) structure, that is an increasingly complex structure [28]. Specifically, the minimal structure was, e.g., /ba/ and the maximal was /sbran/ (expected as [zbran]). Block II included 4 statements in Italian, proposed in isolation for eliciting broad focus, i.e. and out-of-the-blue statement, or proposed within a context in order to induce a contrastive-corrective focus interpretation. Items were randomized within blocks, and blocks were randomized across subjects. Crucially, syllables with simple vs. complex structures, also involve different segmental compositions and a different number of consonants, and pitch accents produced in different focus conditions are characterized by different tonal compositions (H+L* vs. H*+L in broad and contrastive-corrective focus, respectively, in the variety of Italian under investigation – see below) and require different F0 modulations (falling vs. rising-falling patterns).

Participants were 12 volunteers: 8 patients with Parkinson's Desease and mild-to-moderate hypokinetic dysarthria (PDs, 4F/4M, from the province of Lecce – Italy, mean-age: 72.8, range: 64-81;

Activities of Daily Living (ADL): 4-6/6, Instrumental Activities of Daily Living (IADL): 4-7/8; Dysarthria severity: 1-2), and 4 controls with no known neurological disease (CTRs, 2F/2M, from the province of Lecce – Italy, mean-age: 67.75, range: 60-85). Patients were audio recorded in ON-phase during a reading task, in a quiet room at the Phoniatrics and Communicative Disorders Rehabilitation Center (Lecce, Italy), by means of an Edirol Sound-Card and a high-quality Sennheiser microphone (96000Hz s.r.).

Intelligibility ratings were collected by asking 3 raters to judge the intelligibility of each item on a 5-point Lickert scale, where 1=min, 5=max (along the lines of the Robertson's Dysarthria Profile, where, though, a 1-4 scale is adopted by a therapist, and expected also by a family member and unknown person; raters engaged here were a therapist, a phonetician and a naïve rater). Individual raters listened to stimuli via headphones in a quiet room, and judged the stimuli with no knowledge of speaker identity or neurological diagnosis. The rates given to each collected item were used to compute the mean intelligibility rate for each item and for each subject. The inter-rater reliability was calculated (SPSS - Cronbach's Alpha). Mean and Standard Deviations of ratings in the Parkinson's group were used to group subjects in "intelligibility" groups; the group consistency was also checked with reference to IADL/ADL and RDP scores.

As for acoustics, the speech signal was manually labelled as for the relevant landmarks (PRAAT), which were segment boundaries, position of the plosive burst, and tonal targets position. Regarding segments, measurements concerned consonants, and specifically, duration (absolute. and normalized over the word duration), intensity and Center of Gravity (CoG) in fricatives and plosive bursts, Voice Onset Time (VOT), and Δintensity for adjacent Cs and Vs; as for vowels, duration, first and second formants (F1, F2), and intensity were measured. With respect to intonation, time and fundamental frequency values (F0) at tonal landmarks were measured, together with the tonal target-to-segment boundary and tonal target-to-tonal target latencies (absolute and normalized). Such measures were used to calculate the rise and fall F0 delta, the F0-slope and the pitch range (in semitones, ST). Levene's test was used to check for homogeneity of variance among samples for each acoustic measure. As in many cases the assumption was not met, non-parametric Kruskal-Wallis tests were used to compare acoustic measures ($p<0.05$).

Acoustic measures were then correlated with perceptual ratings by means of Spearman rank order correlations, considering the strength of the correlation for both items and intelligibility groups.

# 4. Results

### 4.1. Intelligibility ratings

The inter-rater reliability on the whole set of data was high (Alpha=0.871, $p<.001$), and slightly better on the sentence block than on the word block (word: Alpha=0.860,$p<.001$; utterances subset: Alpha=0.896,$p<.001$). The average rating for items ranged from 3 to 5 (mean 4.54; sd 0.69), and that for subjects ranged from 4.1 to 4.9 (mean 4.47; sd 0.29). Mean and standard deviation of intelligibility levels for subjects were used to identify three PD "intelligibility" groups (group 1: ratings 4.1-4.3; group 2: 4.4-4.6; group 3: 4.7-4.9).

Consistency was found in terms of group internal IADL, ADL and RDP's scores (group 1=ADL:4-5/6,I ADL:4-5/8, RDP's=2/3; group 2=ADL:5-6/6, IADL:6-7/8, RDP's:2; group 3 =ADL:6/6, IADL:8/8, RDP's:1). Thus, intelligibility groups could then be expected to correspond to different dysarthria severity levels.

### 4.2. Acoustic measurements

#### *4.2.1. Segmental measures*

Due to space limits, it is not possible to report statistic measurements for all the acoustic measures considered. However, differences in mean and standard deviation values were found even though statistical significance was not always reached and, in any case, not all PD speakers differed from CTR's. For instance, group 1 showed longer /t/ normalized durations than other groups (Figure 1), the Kruskal-Wallis test indicated that acoustic measures were indeed statistically different ($p<0.05$, that is 0.0027), with the Mann–Whitney U tests showing that it was indeed group 1 to be significantly different from others. For other measures, less neat situations were found (e.g. group 3 showed significantly greater VC delta intensity, in comparison to other groups).

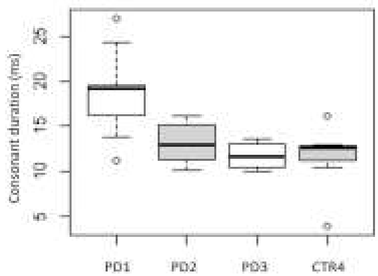


Figure 1: */t/ consonant normalized duration (ms) by intelligibility/subject group*

#### *4.2.2. Prosodic measures*

Differences in mean and standard deviation values were observed across groups in relation to prosodic measurements too, even though statistical significance was never reached. For instance, pitch range mean values changed especially for group 1 in both Broad Focus and Corrective Focus conditions (Figure 2, left BF vs. right CF boxes), but the Kruskal-Wallis test indicated the lack of statistical differences ($p>0.05$).

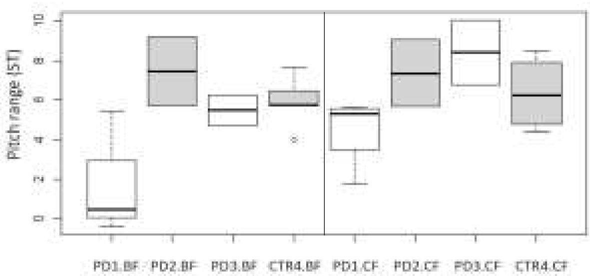


Figure 2: *Pitch range (ST) mean values by intelligibility/subject group (PD group 1, 2, 3, CTR 4), in Broad (left, BF boxes) and Corrective Focus conditions (right, CF boxes)*

### 4.3. Correlation between acoustic measurements and intelligibility ratings

#### *4.3.1. Segmental measures*

Correlation results are reported with reference to plosives and the following vowel in Table 1, as plosive are always present in the sequences considered. Considering correlations involving both items and intelligibility groups, we found significant correlations and, even more importantly, moderately-to-strong correlations, corresponding to r values equal or higher than 0.5 (the sign of the correlation, which indicates the direction of the association is not discussed here).

In the case of syllables including unvoiced plosives (Table 1, left) and especially when considering intelligibility groups, we found a strong correlation between ratings and consonant duration (Cdur, even more when the normalized measure, Cdur%, is taken into account). Concerning vowel measurements, Vowel intensity (Vint) is also strongly correlated with intelligibility, and a slight correlation is found for the vowel-to-consonant delta in intensity (Δint CV). Results for syllables including voiced plosives (Table 1, right), confirm the strength of the correlation between intelligibility judgements and both vowel intensity (Vint), and vowel-to-consonant delta in intensity (Δint CV).

Table 1: *Spearman rank order correlation results for Item and Intelligibility groups – syllable including unvoiced (left), voiced (right) stops* (**0.01; *0.05)

| Measure | Item | Intelligibility group | Item | Intelligibility group |
|---|---|---|---|---|
| C dur | **r=-.639*** | r= -.596* | r=-.137 | r=-.372 |
| C dur% | r=-.594* | **r= -.775**** | r=-.098 | r=-.454* |
| C int | r= -.420 | r= -.302 | r= .057 | r= -.107 |
| C CoG | r= -.425 | r= -.265 | r= -.286 | r= -.427* |
| VOT | r= -.264 | r= -.381 | r= -.412 | r= -.428 |
| Δint CV | **r= .620*** | **r= .605*** | **r=.606**** | **r= .770**** |
| V dur | r= -.007 | r= .180 | r= .417* | r= .307 |
| V int | r= .324 | **r= .775**** | r= .388 | **r= .715**** |
| F1 | r= -.386 | r= -.302 | r= .369 | r= .346 |
| F2 | r= -.372 | r= -.426 | r= .205 | r= -.430* |

Concerning complexity, the analysis regarded measures that turned out to be strongly correlated in Table 1, and aimed at checking if there were differences related to syllable structure/composition. Taking complexity into account, even stronger correlations could be found. For instance, in the case of CV delta intensity for voiced plosives (Table 2, bottom), the significant positive correlation was stronger for complex structures/compositions (e.g., C r=.731* vs. sCr r=.882**).

Table 2: *Spearman rank order correlation results for Intelligibility groups – syllable of increasing complexity including unvoiced (top) and voiced (bottom) stops* (**0.01; *0.05)

| Measure | C | Cr | sC | sCr |
|---|---|---|---|---|
| Cdur% | r= -.375 | **r= -.445*** | **r= -.466*** | **r=-.575*** |
| Δint CV | **r= .605*** | **r= .678**** | **r= .775**** | **r=.805**** |
| Vint | r= .475 | **r= .575*** | **r= .665**** | **r=.775**** |

| Measure | C | Cr | sC | sCr |
|---|---|---|---|---|
| Δint CV | **r= .731*** | **r= .850*** | **r= .806*** | **r=.882**** |
| Vint | r=.415 | **r=.521*** | **r=.535*** | **r=.715**** |

### *4.3.2. Prosodic measures*

As you can see in Table 3, we found significant correlations concerning prosodic measures too. We first looked at global measures, which are given much attention in the literature. A moderately-to-strong significant correlation was found for various measures in either the Broad focus or the Corrective focus condition, or both. In the case of measures related to linguistically relevant intonational patterns, the impact of complexity is checked by looking at local measures (Table 4), involving specific tonal event (by comparing the pitch accent that is considered to be more complex with the less complex one (the Corrective, rising-falling pattern vs. the Broad, falling one). Correlations are indeed more and stronger in the Corrective focus than in the Broad focus, especially for latency and slope measurements (e.g., HLlatency, HLF0slope).

Table 3: *Spearman rank order correlation results for Intelligibility groups in Broad Focus (left) and Corrective Focus conditions (right)* (**0.01; *0.05)

| Measure | Broad Focus Intelligibility group | Corrective Focus Intelligibility group |
|---|---|---|
| F0 range | r= .556 | **r= .869**** |
| H F0 | **r= -.729*** | r= -.510 |
| L F0 | **r= -.815**** | r= -.435 |
| F0 fin | **r= -.815**** | **r= -.945**** |
| F0 max | **r= -.692*** | r= -.510 |
| L1F0 | -- | **r= -.680*** |

Table 4: *Spearman rank order correlation results for Intelligibility groups in Broad Focus (left) and Corrective Focus conditions (right)* (**0.01; *0.05)

| Measure | Broad Focus Intelligibility group | Corrective Focus Intelligibility group |
|---|---|---|
| Vbnd-to-L lat. | r= **.605** | **r=-.837**** |
| Vbnd-to-L lat.% | r= .556 | **r=-.956**** |
| HL latency | r= .544 | **r= -.869**** |
| HL latency% | r= .321 | **r= -.850**** |
| HL F0 slope | r= .272 | **r= .945**** |
| HL F0 delta | **r= .667*** | **r= .680*** |

## 5. Discussion and conclusions

The work aimed at checking the correlation between intelligibility judgements and acoustic measurements on both segments and prosodic-intonational events that play a linguistic role, and at investigating such correlation in contexts of increasing complexity in phonetic-phonological terms.

Acoustic results did not always turn out significantly different across groups, thus they do not seem to offer clear hints on accuracy. However, our hypotheses on the correlation with intelligibility have been confirmed, as 1) the correlation was found between acoustic measures of linguistically relevant events and speech intelligibility at both segmental and prosodic level, and 2) increasing phonetic complexity, always taking into account language/variety specific features (e.g., intonation patterns) appears to enhance the correlation. However, results are preliminar and more data are needed to confirm them.

## 6. Acknowledgements

This work was partially funded by the PRIN 2017 project 2017JNKCYZ. We thank all the pathological and control subjects who participated in the experiment.

## 7. References


[1] R.D. Kent, G. Weismer, J.F. Kent and J.C. Rosenbek. Toward phonetic intelligibility testing in dysarthria. *Journal of Speech and Hearing Disorders,* 54, 482-99, 1989.

[2] Duffy, J. R. (2005). Motor speech disorders (2nd ed.). St. Louis, MO: Mosby.

[3] Hustad , K. C . (2007) . Effects of speech stimuli and dysarthria severity on intelligibility scores and listener confi dence ratings for speakers with cerebral palsy . Folia Phoniatrica et Logopaedica , 59 , 306 – 317 .

[4] Pennington , L. , & Miller , N . (2007) . Infl uence of listening conditions and listener characteristics on intelligibility of dysarthric speech . Clinical Linguistics and Phonetics , 21 , 393 – 403 .

[5] Pennington , L. , Miller , N. , Robson , S. , & Steen , N . (2010) . Speech and language therapy for older children with cerebral palsy: A systems approach . Developmental Medicine and Child Neurology , 52 , 337 – 344 .

[6] Pennington , L. , Roelant , E. , Thompson , V. , Robson , S. , Steen , N. , & Miller , N . (2013) . Intensive dysarthria therapy for younger children with cerebral palsy . Developmental Medicine and Child Neurology , 55 , 464 – 471 .

[7] Yorkston KM, Beukelman DR, Strand EA, Bell KR. Management of motor speech disorders in children and adults. 2nd Ed Pro Ed; Austin, TX: 1999.

[8] Kent , R. D. , Miolo , G. , & Bloedel , S . (1994) . The intelligibility of children's speech: A review of evaluation procedures . American Journal of Speech-Language Pathology, May , 81 – 95 .

[9] Zraick , R. , & Liss , J. M . (2000) . A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality . Journal of Speech, Language, and Hearing Research , 42 , 979 – 988 .

[10] Hustad , K. C . (2006) . Estimating the intelligibility of speakers with dysarthria . Folia Phoniatrica et Logopaedica , 58 , 217 – 228 .

[11] Robertson S.J. (1982) Dysarthria profile. Backgound and development. College Speech Therapist Bullettin, 359.3.82.

[12] Fussi F, Cantagallo A. (1999): *Profilo di valutazione della disartria. A*dattamento italiano del test di Robertson, raccolta di dati normativi e linee di trattamento (1982) Ed. Omega

[13] Liss JM, Utianski R, Lansford K. Crosslinguistic application of English-centric rhythm descriptors in motor speech disorders. Folia Phoniatrica et Logopaedica. 2013; 65:3–19.

[14] Yunusova Y, Weismer G, Kent RD, Rusche NM. Breath-group intelligibility in dysarthria: Characteristics and underlying correlates. Journal of Speech, Language, and Hearing Research. 2005; 48:1294–1310.

[15] LYNDA FEENAUGHTY, KRIS TJADEN, and JOAN SUSSMAN (2014) Relationship between acoustic measures and judgments of intelligibility in Parkinson's disease: A within-speaker approach Clin Linguist Phon. 2014 November ; 28(11): 857–878.

[16] Patel, R. (2003). Acoustic characteristics of the question-statement contrast in severe dysarthria due to cerebral palsy. Journal of Speech, Language, and Hearing Research, 46(6), 1401–1415.

[17] Patel, R., Connaghan, K. P., & Campellone, P. J. (2013). The effect of rate reduction on signaling prosodic contrasts in dysarthria. Folia Phoniatrica et Logopaedica, 65(3), 109–116.

[18] PIERREHUMBERT J. (1980). *The phonology and phonetics of English intonation*. PhD thesis, MIT – published in 1988 by IULC.

[19] LADD D. R. (1996). *Intonational Phonology*. Cambridge Studies in Linguistics, Cambridge University Press, Cambridge.

[20] *Ioana Chitoran and Abigail C. Cohn. (2009)* Complexity in phonetics and phonology: gradience, categoriality, and naturalness *In* Pellegrino, F, Marsico, E, Chitoran, I, and Coupe, C eds. 2009. *Approaches toPhonological Complexity*. Berlin - New-York: Mouton De Gruyter, 21-46.

[21] Chomsky, N. and M. Halle 1968 *The Sound Pattern of English*. New York, NY: Harper and Row.

[22] Lindblom, Björn and Ian Maddieson. 1988. Phonetic universals in consonant systems. In *Language, Speech and Mind* (ed C. Li and L.M. Hyman). Routledge, London. 62-78.

[23] Clements, G. N. & J. Keyser. (1983). CV phonology. Cambridge, MA: MIT Press

[24] Reimers, P. M. (2015). Markedness in first language acquisition. In M. Yavaş (Ed.), Unusual productions in phonology: Universals and language-specific considerations (pp. 140–158). New York: Psychology Press

[25] Romani, C., Galuzzi, C., Bureca, I. & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. Cognitive Psychology, 62, 151–192.

[26] MJ. Ball, C Code, J. Tree, K. Dawe, J. Kay (2004). Phonetic and phonological analysis of progressive speech degeneration: a case study. CLINICAL LINGUISTICS & PHONETICS, 2004, VOL. 18, NO. 6–8, 447–462

[27] Gandour, Jackson, T., Richard A. Harshman. 1978. Crosslanguage differences in tone perception: A multi-dimensional scaling investigation. *Language and Speech* 21:1-33.

[28] Kenstowicz, M. (1994). Phonology in generative grammar. Cambridge, MA: Blackwell.