PHONOLOGICAL AND SOCIOPHONETIC INFORMATION IN DYSARTHRIC SPEECH: A FIRST ARTICULATORY INVESTIGATION ON ITALIAN

Gili Fivela Barbara, Sonia d'Apolito, Anna Chiara Pagliaro

University of Salento – CRIL-DReAM, Lecce, Italy
{barbara.gili,sonia.dapolito,annachiara.pagliaro}@unisalento.it

ABSTRACT

In Italian, alveolar plosives and fricatives are different phonemes, and in some varieties the formers may be realized as aspirated consonants, which represent sociophonetic markers. Dysarthric speakers may experience different types of difficulty in producing the segments mentioned above, as fricatives are known to require finer supralaryngeal control than plosives, and a specific synchronization of laryngeal and supralaryngeal gestures is required in producing aspirated plosives.

In this paper, plosives, fricatives, and plosive aspiration phases are acoustically and articulatorily investigated in Parkinsonian dysarthric Italian speakers and matching controls. Research questions concern both aspects of motor control strategies in consonant production, and possible differences in the realization of phonemic distinctions and sociophonetic markers. Results suggest that, even if fricatives and plosives are kept different, the disease seems to have a stronger impact on the production of the latter. Changes may regard both the realization of phonemic distinctions and the production of sociophonetic markers.

Key Words: Parkinson's Disease, plosives, aspiration, fricatives

1. INTRODUCTION

Plosives and fricatives are among the most frequent sounds in the world languages. Almost 92% of languages have voiceless plosives series [1:27], and moreover it "is natural for a language to have at least one sibilant, namely, a voiceless alveolar [s]" [2:108]. Italian, for instance, shows both plosives and fricatives alveolar phonemes. In motor control terms, the production of plosives is considered simpler than that of fricatives, in line with the hypothesis that a collision with another articulator requires less precision than the realization of a narrowing between articulators [3, in line with 4]. This may hold for Parkinson's Disease (PD) dysarthric speech, where [5] found that /t/ is distorted by realizing incomplete closures or voicing (word medially more than word initially), while /s/ may be distorted by producing stopping, voicing, unintelligible segments or even omissions; consistently, [6] found a reduction of the tongue blade elevation in PD's alveolar fricatives (/s/, /z/) in comparison to controls'. However, previous work on Italian [7, 8] suggests a different picture, where PDs realize fricatives similarly to controls and differ in plosive production (where the physical obstacle ends the closing gesture but requires, e.g., a longer release [7]).

Besides the distinction between plosives and fricatives, in some languages plosives may also be aspirated, with aspiration playing a phonological role (e.g., Korean) or representing an allophonic variant (e.g., English). In some varieties of Italian (e.g., Salentinian Italian), aspirated plosives are allophonic variants and aspiration is affected by prosodic conditions, in that it is stronger in word internal poststress position or before pauses [9:§150, 10]. Acoustically, aspiration may be measured as the Voice Onset Time (VOT, the time interval from consonant release to start of vocal folds vibration), since aspiration represents the air outflow through the glottis after the closure release in the vocal track [11]. From an articulatory point of view, in aspirated plosives the glottal opening starts after the oral closure and reaches its peak at the closure release, rather than earlier than it as it happens in non-aspirated plosives [12, 13]. As for pathological speech, the review in [14] recalls that VOT abnormalities in dysarthria (and apraxia) mainly reflect loss of motor control, with dysarthric speakers showing in plosives longer closure duration and VOT than controls (see also the review in [15] and, for English, [16]). As for English speakers with PD hypokinetic dysarthria, however, results are inconsistent. They point to decreased VOTs in PDs and reduced contrasts in voiced and voiceless plosive VOT [17, 18], as also reported in Korean bilabials [19] and alveolars [20]; however, results also point to the lack of difference in VOT length between PDs and controls ([21]; see also [22]). Crucially, though, Korean dysarthric speakers seem to control the synchronization between laryngeal and supra-laryngeal articulators to distinctively produce phonologically different (aspirated vs. non-aspirated) plosives [19]. What about aspiration when it represents a sociophonetic trait?

In this paper, the distinction between plosives and fricatives in Italian, as well as the plosive aspiration phase, are investigated from both the acoustic and articulatory point of view. Crucially, besides the phonological contrast between plosives and fricatives, in the variety of Italian considered here, namely the Salentinian variety (Southern Apulia, Italy), the aspiration of plosives is an allophonic variant that is sociolinguistically (but not phonologically) relevant.

2. GOALS AND HYPOTHESES

We aim at observing if PD speakers, who are still able to preserve the plosive vs. fricative contrast: 1) exhibit more difficulties in maintaining the narrowing gesture than the closure; 2) master as precisely as controls the coordination between laryngeal and supra-laryngeal gestures in order to maintain the sociophonetic feature of plosive aspiration; 3) show possible compensation strategies in realizing both phonological contrasts and sociolinguistically relevant features.

We hypothesize that PD speakers who maintain the plosive vs. fricative phonological difference may show: a) imprecise realization of the frication narrowing (see [3]); b) longer closure and VOT duration than controls in aspirated plosives (see [15, 16]), though a stronger aspiration in word internal poststress position may be expected (see [9, 10]); finally, we hypothesize that c) dynamical adjustment related to compensation strategies may be stronger in order to keep fricatives and plosives different than to preserve sociolinguistic traits.

3. METHODS

Tongue dynamics was observed in unvoiced alveolar plosives and fricatives, inserted as singletons (/t/, /s/) in a stable vowel context (/a/-to-/a/, e.g., /la'tata/, /la'sasa/). They were onset of both word initial, stress syllables, as well as word medial, poststress syllables, and were inserted within the carrier sentence *La X blu* 'the blue X'. Plosives could be realized as aspirated ([t^h]).

Nine Salentinian subjects, that are 5 mild-to-severe PD subjects (males; average age 71.6 y.o.) and 4 age/gender-matched healthy controls (HC: average age 69.75), have been acoustically and articulatorily recorded in ON phase through Electromagnetic Articulography (AG501). Articulatory data were acquired by means of 7 sensors, glued on subjects: 2 on tongue mid-sagittal plane (dorsum and tip), 2 on lips vermillion border (upper and lower), 1 on the nose and 2 behind the ears for normalization (one HC subject was excluded from the articulatory analysis because of errors in tongue tip tracking, which is pivotal for the present paper). Subjects read aloud the corpus of written sentences for a minimum of 5 times. After an auditory check, aimed at selecting plosives and fricatives realization of target segments, acoustic and articulatory signals were segmented and labelled in PRAAT and MAYDAY respectively [23, 24]. The acoustic duration of each segment was measured and, as for plosives, both the closure (stop closure to spike) and the VOT duration (spike to beginning of following vowel) were collected, differently from previous investigations [7, 8]. Regarding kinematics, the following landmarks were semi-automatically inserted along the vertical axis of tongue dorsum (TD) and tongue tip (TT) tracks: gesture target, located at the zero velocity, and, for each tracked segment, maximum velocity, labelled at the velocity peak of the relevant coil. Further, on the position track we labelled the beginning/end of the closure phase of the consonantal gesture (e.g., beginning of release, at the 20% of velocity increase after the release; measure missing in previous investigations [7, 8]). Concerning the duration of tongue tip/dorsum gestures, the following measures were calculated: (a) the interval including both the closing and the opening consonantal gestures (Clo-Op); with regards to the release phase, the interval from the consonantal target both to (b) the vowel target (Ctarget-Vtarget latency) and (c) the onset of the vowel (20% of peak velocity - Ctarget-Vonset). Further, as for the consonantal gesture to the following vowel, we measured the peak velocity and stiffness (peak velocity/gesture amplitude). Time measurements were analyzed as both absolute and normalized (over the acoustic/articulatory word duration) measures.

As for statistical analysis, mixed models were implemented (in R, lme4 [25, 26]). The fixed effects were *Population* (PD vs. HC), *Constriction* (plosives vs. fricative) *Position* (initial-stress vs. medial-poststress, henceforth stress and poststress), and *Repetition* (number of productions). Interactions were also investigated. By subject random slopes were set in order to account for inter-subjects variability in the realization of *Population*, *Constriction* and *Position*. Analyses of aspiration were run on plosives only, discarding the *Constriction* factor. Residual plots did not reveal obvious deviations from homoscedasticity or normality. Significance ($p<0.05$) of fixed effects was checked via a *Likelihood Ratio Test*.

4. RESULTS

4.1. Auditory check and acoustic measures

Two PD subjects produced plosives and fricatives as barely indistinguishable segments, in that they realized plosives without a neat closure phase (slightly noisy closure and no spike in spectrogram), or as voiced rather than unvoiced segments. Thus, they have no longer been considered, as results of measurements concerning their productions were not useful in investigating the difficulties in maintaining the narrowing gesture vs. the closure nor in analyzing the plosive aspiration phase.

Consonant acoustic duration is affected by *Position*, (absolute measures: χ2(1)=23.77 p=,000; normalized: χ2(1)=28.55 p=,000) and duration increases in stress in comparison to poststress position (absolute: 11.15ms ± 2.21 S.E.; normalized: 2.57% ±0.46 S.E.) for HC speakers and fricative condition. Moreover, consonant normalized duration is affected by *Population* too (χ2(1)=4.29 p=,03), while absolute duration changes depending on *Constriction* (χ2(1)=11.14 p=,000). Values are smaller in PDs (-2.07ms ±0.85 S.E.) and in plosives in comparison to fricatives (-7.50ms ±2.21 S.E.). No interactions are found (Figure 1).

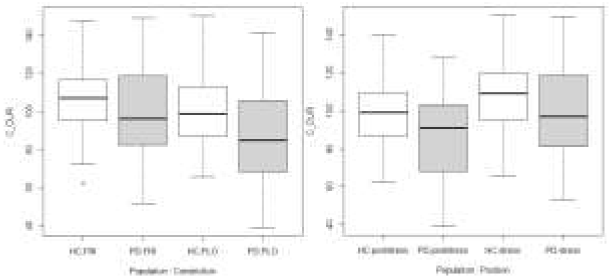

Figure 1: Consonant absolute acoustic duration for Constriction (left panel) and Position (right panel) in HCs and PDs (white and grey columns, respectively).

When considering plosives as for their closure, the duration differs significantly only with respect to *Position* (absolute: χ2(1)=11.50 p=,000; normalized: χ2(1)=13.57 p=,000), being longer in stress position (absolute:9.62 ±2.74 S.E; normalized:2.19% ±0.57 S.E.).

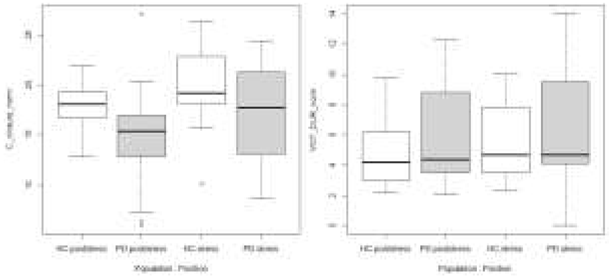

Figure 2: Consonant normalized closure (left panel) and VOT acoustic duration (right panel) for Position in HCs and PDs (white and grey columns, respectively).

The interaction between *Population* and *Position* does not reach significance, though the closure absolute duration is shorter in PDs, especially in poststress position (χ2(1)=9.54 p=,08). As for the VOT phases, the duration is only affected by the *Position* factor (absolute: χ2(1)=7.06 p=,007; normalized: χ2(1)=3.87 p=,004) with no significant interactions. Measures show that VOT duration increases in stress in comparison to poststress position (absolute: 3.96ms ±1.45 S.E.; normalized: 0.89%- ±0.44 S.E.), though its duration appears to be highly variable in PDs - Figure 2.

4.2. Articulatory measures

The *Population* factor affects the interval including the whole consonant, i.e. the closing plus the opening gesture, as for the normalized measure (Clo-Op: χ2(1)=4.50, p=0.03). A slight increase is found in PDs (e.g., by about 2% ± 0.7 S.E. in poststress position and fricative condition). The duration of the opening gesture alone does not significantly differ (Ctarget-Vtarget latency: χ2(1)=3.05, p=0.08), though it slightly increases in PDs. No interaction with *Position* and *Constriction* is found.In line with the literature, absolute Clo-Op duration is affected by *Constriction* (χ2(1)=4.51, p=0.03), with plosives showing shorter duration than fricatives (16.8 ms ± 6.5 S.E. in the case of plosives for HC speakers and poststress condition) - Figure 3, left panel. Interaction between *Population* and *Position* is found for both the absolute Clo-Op and the Ctarget-Vtarget measure (χ2(1)=4.90, p=0.03; χ2(1)=12.016, p=0.01), as PDs show shorter duration especially in poststress position.

As for the release phase corresponding to the Ctarget-Vonset, the absolute interval does not reach significance as for the *Position* factor (χ2(1)=2.89, p=0.08),though values increase (by about 8 ms in HC ± 4 S.E. in the case of stress position and fricative) and increase even more in PDs realization.

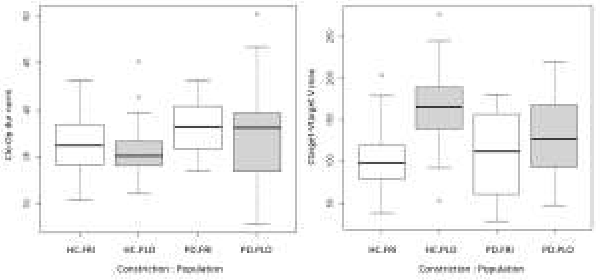

Figure 3: Consonant absolute articulatory duration (Clo-Op, left) and peak velocity in consonant opening (Ctarget-Vtarget, right) for Populationin Fricatives and Plosives (white and grey columns, respectively).

Peak velocity and stiffness measures showed the significant impact of the *Position* factor (respectively, χ2(1)=5.48, p=0.02; χ2(1)=3.7334, p=0.05), with expected higher values in stress than in poststress position. The opening gesture peak velocity is the only measure that is affected by the *Constriction* degree

($\chi 2(1)=5.04$, $p=0.02$), as it increases in plosives (by about 45 ± 15 S.E. for HC speakers and poststress condition). However, an interaction between *Constriction* and *Population* is found, as the peak velocity is lower in PD's plosives, not differing much from that found in fricatives – Figure 3, right panel. Further, *Population* also interacts with *Position* ($\chi 2(1)=3.7289$, $p=0.05$) with PDs showing increased peak velocity in poststress than in stress position.

5. DISCUSSION AND CONCLUSIONS

The analysis of productions by HC and PD speakers who are still able to preserve the plosive vs. fricative contrast shows that acoustic measurements regarding the difference between plosives and fricatives (*Constriction* factor) did not significantly differ depending on the PD vs. HC *Population*. The overall normalized duration of consonants is smaller in PDs than in HCs, but results are in line with the expected longer duration in fricatives than in plosives, as well as in stress in comparison to poststress positions. In PDs, fricatives do not seem to vary more than plosives. Rather, PDs tend to shorten the plosive closure acoustic duration in comparison to HCs, especially in poststress position, which is in line with the overall shorter duration of acoustic consonants in PDs. Finally, the plosives' VOT, which is longer in stress than in poststress position, is not significantly different in PDs and HCs. Thus, within the inconsistent results recalled in section 1, our findings support, e.g., [21, 22] rather than [15, 16], even though the VOT duration seems to be more variable in the PDs than in HCs. In fact, the VOT by PDs and HCs is longer in stress than in poststress position, in line with prosodically induced VOT modulation (overall VOT lengthening in stress position), with no evident or consistent sociophonetic differences ([9, 10]).

As for the articulatory analysis, the normalized measures regarding both the whole closing plus opening consonantal gesture and, to a certain extent, the opening gesture (Clo-Op and Ctarget-Vtarget) increase in PDs. Thus, articulatory data point to longer gestures in PDs, despite the shorter duration of acoustic segments. Only absolute measures confirm expectations regarding the shorter duration of gestures in plosives than in fricatives together with a shorter duration for PDs, especially in poststress position. Peak velocity and stiffness measures regarding the opening gesture showed higher values in stress than in poststress position, while peak velocity is the only measure that is affected by the *Constriction* degree, as it increases in plosives than in fricatives. Interestingly, peak velocity is lower in PD's plosive opening gesture, where it does not differ much from that found in fricatives. The auditory check preliminary to the analysis allows us to discard the hypothesis that the lower peak velocity is due to an incomplete realization of the closure. Nevertheless, from the articulatory data it may be hard to tell if the lower peak velocity is related to laryngeal-supralaryngeal gesture phasing difficulties [12, 13]. What we can observe, though, is that the aspiration phase seems to acquire characteristics that are similar to those found in fricatives, and this could be a stage prior to the incomplete closure realization observed in other PD speakers (e.g., those excluded in the present analysis). On the other hand, the increased peak velocity value in poststress than in stress position by PDs seems to be in line with the observed shorter aperture phase in poststress position, and with possible differences related to aspiration and consonant prosodic position.

The abovementioned results do not seem to support expectations related to major differences between PD and HC in fricatives more than in plosives (see goal 1). Rather they point in the direction of some possible inaccuracies in the realization of plosives by PDs, as their opening gesture in plosives show similarities with that found in fricatives. Further, the plosive aspiration phase in PDs (see goal 2) seems to be affected by the stress/poststress consonant position as in the case of HC speech, in line with prosodic modulation and with no evident or consistent difference due to sociolinguistic factors. In fact, no observation of longer aspiration in poststress rather than stress position [9, 10] has been found in either PDs or HCs. However, PDs show higher peak velocity in poststress than in stress position, which could be a cue of aspiration differences in comparison to what observed in HCs. Besides, difficulties in laryngeal-supralaryngeal gesture phasing in PDs could have appeared in terms of variability in VOT acoustic duration or with reference to the observed overall lowering in peak velocity in plosive opening gestures. Therefore, speakers who are still able to differentiate plosives and fricatives seem to have already modified some characteristics of their opening gesture in plosives, and changes may be detected both in the case phonemic distinctions are realized and when the production of sociophonetic markers is concerned (see goal 3).

Of course, more material is needed to clarify the matter, together with further acoustic (e.g., noise intensity measures during VOT) and articulatory measures (e.g., opening gesture amplitude).

ACKNOWLEDGEMENTS

This work was funded by the project PRIN 2017JNKCYZ. We thank F. Sigona for technical support, M. Iraci for help in the corpus collection, both M. Iraci and G. Pucci for a preliminary analysis of part of the material, and all experimental subjects for participation.

6. REFERENCES

[1] Maddieson, J. 1984. *Patterns of Sounds*. CUP.

[2] Hockett, C.F. 1955.*A Manual of Phonology(IJAL Monographs 11)*. Indiana University.

[3] Fuchs, S., Perrier, P., Geng, C., Mooshammer, C. 2006. What role does the palate play in speech motorcontrol? Insights from tongue kinematics for German alveolar obstruents, in Harrington, J., Tabain, M., (eds.), *Speech Production: Models, Phonetic Processes, and Techniques,* Psychology Press, 149-164.

[4] Löfqvistand, A., Gracco, V.L. 1997. Lip and jawkinematics in bilabial stop consonant production, *JSLHR*40, 4, 877-893.

[5] Antolík, T., Fougeron, C. 2013. Consonant distortions in dysarthria due to Parkinson's disease, Amyotrophic Lateral Sclerosis and Cerebellar Ataxia, *Proc. Interspeech 2013*, August, Lyon, France, 2152-2156.

[6] Logemann, J.A., Fisher, H.B. 1981. Vocal tract control in Parkinson's Disease: Phonetic feature analysis of misarticulations, *JSHD*, 46, 348-352.

[7] Iraci, M., Grimaldi, M., Gili Fivela, B. 2017. Dynamic Aspects of the Alveolar Fricative vs. Stop Contrast in Parkinson's Disease, PAPE Conference (poster presentation), *Phonetics and Phonology in Europe*, Cologne, 11th-15th June, 2017.

[8] Iraci, M. 2017. Vowels, consonants and co-articulation in Parkinson's Disease. Unpublished PhD Dissertation. University of Salento, Lecce, Italy.

[9] Rohlfs, G. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti. Fonetica*. Einaudi, v. 1.

[10] Sobrero, A., Romanello, M.T. 1981.*L'italiano come si parla in Salento*. Milella.

[11] Kim, H.S., Kim, H.H. 2019. Acoustic analysis of Korean stop sounds in patients with dysarthrias, *Clinical Arch. of Communic. Disord*, 4, 3, 201-213.

[12] Best, C.T.1995. A direct realist view of cross-language speech perception, in Strange, W. (ed.), *Speech perception and linguistic experience: issues in cross-language research*. York Press, 171–204.

[13] Browman, C.P., Goldstein, L. 1986. Articulatory gestures as phonological units, *Phonology*, 6, 151-206.

[14] Auzou, P., Ozsancak, C., Morris, R. J., Jan, M., Eustache, F., Hannequin, D. 2000. Voice onset time in aphasia, apraxia of speech and dysarthria: a review, *Clinical Ling.& Phonetics*, 14, 2, 131-150.

[15] Kent, R.D., Weismer, G., Kent, J.F., Rosenbek, J.C. 1989. Toward phonetic intelligibility testing in dysarthria, *JSHD*, 54, 482-499.

[16] Forrest, K., Weismer, G., Turner, G. 1989. Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric males, *J. of the Acoustical Society of America*, 85, 2608-2622.

[17] Weismer, G. 1984. Articulatory characteristics of Parkinsonian dysarthria: Segmental and phrase-level timing, spirantization, and glottal-supraglottal coordination, in McNeil, M.R., Rosenbek, J.A., Aronson, A.E. (eds.), *The dysarthrias: Physiology, acoustics, perception, management*, College-Hill Press, 101-130.

[18] Morris, R.J. 1989. VOT and dysarthria: A descriptive study,*J. of Communication Disorders*, 22, 23-33.

[19] Park, S., Sim, H., Baik, J.S. 2005. Production ability for Korean bilabial stops in Parkinson's disease, *Journal of Multilingual Communication Disorders*, 3, 90-102.

[20] Kang, Y., Kim, Y. D., Ban, J. C., Seong, C. J. 2009. A comparison of the voice differences of patients with idiopathic Parkinson's disease and a normal-aging group, *Phonetics and Speech Sciences*, 1, 99-107.

[21] Fischer, E.,Goberman, A. M. 2010. Voice onset time in Parkinson disease",*J. of Communic. Dis.*, 43, 21-34.

[22] Bunton, K., Weismer, G. 2002. Segmental level analysis of laryngeal function in persons with motor speech disorders. *Folia Phoniatrica etLogopaedica*, 54, 223-239.

[23] Boersma, P. 2002. Praat, a system for doing phonetics by computer, *Glot International*,5 no. 9/10, 341-345.

[24] Sigona F., Stella A., Grimaldi M.,Gili Fivela B. 2015. MAYDAY: A Software for Multimodal Articulatory Data Analysisin Romano, A., Rivoira, M., Meandri, I. (eds.), *Aspetti prosodici e testuali del raccontare, Atti del X Conv. AISV*, Edizioni dell'Orso, 173-184.

[25] R Core Team, 2019. R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria.

[26] Bates, D.S., Maechler, M., Bolker, B., Walker, S. 2015.FittingLinear Mixed-Effects Models Using lme4, *Journal of Statistical Software*, 67, 1, 1-48.