PEOPLE WITH PARKINSON'S DISEASE EXHIBIT PHONETIC FLEXIBILITY

Véronique Delvaux[1,2], Eva Goeseels[1], Kathy Huet[1], Myriam Piccaluga[1] Virginie Roland[1], Clémence Verhaegen[1], Bernard Harmegnies[1]

[1]Research Institute for Language Science and Technology, UMONS, Belgium
[2] Fund for Scientific Research (F.R.S.–FNRS), Belgium
veronique.delvaux@umons.ac.be

ABSTRACT

Phonetic flexibility is defined as the ability of the speaker-listener to adapt his speech behavior to the internal and external constraints weighing on the communication situation. This exploratory study deals with phonetic flexibility in elderly people with Parkinson's disease (PD). Twelve participants aged 57 to 79, half of whom had PD, produced the 25 sentences of the corpus under 6 successive experimental conditions: (i) reading; (ii) repetition of an oral model; (iii) interactive game; (iv)=(ii); (v) repetition with instruction to imitate the oral model; (vi) repetition with instruction to inhibit imitation. The results show that the participants, regardless of whether or not they belonged to the PD group, exhibited substantial phonetic flexibility. Indeed, the average pitch and sentence durations systematically deviated from those measured in reading to approach those of the target, and this according to the experimental condition: moderately in repetition, greatly in imitation, almost not in inhibition.

Keywords: Parkinson’s disease, phonetic flexibility, phonetic convergence, repetition, imitation.

1. INTRODUCTION

Phonetic flexibility is defined as the ability of the speaker-listener to adapt his/her speech behavior to the internal and external constraints of the communication situation. Speech is thus considered as a complex behavior supported by multiple, complex and interactive cognitive processes, under the control of the executive functions.

In the linguistic tradition, it is generally considered that the sound form of messages is determined by linguistic constraints and rules, but also depends on a large number of so-called “paralinguistic” factors (attitudes, emotions, dialectal or sociocultural origin of the speaker, etc.), or even “extra-linguistic” factors (e.g. sex, age, physical attributes of the speaker). The second half of the 20th century has documented many cases of phonetic variation related to the dialectal or sociolinguistic origin of the speakers. The last twenty years have seen many advances in our understanding of the effect on speech and voice of many other factors, which had been previously little explored, e.g., acoustic correlates of body size and mass [1], vocal expression of emotion [2], prosodic characterization of attitudinal factors such as irony [3], effects of stress and cognitive load [4], etc.

In traditional linguistics, the focus is on the form of the message, which is considered to be the resultant – unique? – of the combined action of the linguistic, paralinguistic and extralinguistic factors mentioned above. We propose to reverse the perspective and to consider speech from the point of view of cognitive sciences centered on the human subject, the speaker-listener. Each speech production is then the unique and irreducible product of an interaction between human subjects in a specific communication situation. The different factors determine a framework of constraints - the field of possibilities - but do not allow the prediction of a unique sound form. The individual shows flexibility, adaptability and phonetic control. According to her needs, expectations, preferred strategies, which can themselves evolve in relation to the situation, the speaker constantly adapts, in real time, the phonetic form of her speech productions.

Recent work has demonstrated the existence of important resources in speakers in terms of what we call "phonetic flexibility". This is evidenced by the numerous studies on imitation and vocal disguise abilities [5]. Moreover, it is now accepted that this flexibility potential is used daily by speakers during language interactions (cf. the phonetic convergence mechanism typically occurring between two interlocutors: [6]), as well as all the phenomena of adaptation to the interlocutor: 'child-directed speech', 'foreigner-directed' speech, speech addressed to the hearing impaired, etc.; for a review, see [7]). More generally, a large literature concerns adjustments to speech productions in response to constraints on the communication situation, both 'external' constraints (e.g., speech in noise [8]), and constraints internalized by the speaker (the traditional hypo- vs. hyper-articulation continuum: [9,10]). Finally, recent experiments using various speech perturbation paradigms have shown the extent to which speakers are able to adjust their motor programs in real time in order to compensate, totally or partially, for the effects of an articulatory (lip tube, bite block) or auditory (altered feedback) perturbation (e.g. [11]).

These adjustment/adaptation mechanisms of the healthy speaker constitute in fact as many resources accessible to the individual confronted with a set of disturbances due to the development of a pathology affecting speech. Beyond the effects of the pathology itself, of the drug treatment, as well as their interactions, it is commonly admitted that patients set up specific compensatory strategies in order to maintain an adapted production of their speech sounds. These compensatory strategies, to be effective, must be based on a solid foundation of phonetic flexibility, as well as on other elements such as preserved access to cognitive resources (elderly and/or neurodegenerative patients) and motivation (personal, professional) to maintain rich and nuanced spoken communication with the environment.

Previous work in our laboratory has investigated phonetic flexibility abilities via a phonetic learning paradigm of aspirated plosives in French-speaking speakers over 65 years of age [12], as well as in individuals with speech aphasia [13]. In the study reported here, we explore phonetic flexibility ability in elderly individuals with and without Parkinson's disease via a repetition/imitation/inhibition paradigm [14]. Parkinson's disease (PD) is a neurodegenerative disease responsible for progressive disorders of the motor sphere. Among the associated disorders is dysarhtria, a generalized disorder of motor execution of speech, involving difficulty in initiation, lack of coordination or slowness in the execution of articulatory gestures, which typically results in dysphonia, articulatory imprecision and prosodic insufficiency (monopitch, monoloudness, timing abnormalities), associated or not with respiratory problems [15]. At an early stage of the disease, the atypicalities observed in the speech may be consecutive to the disorder itself (or even to the drug treatment), but also to compensatory strategies put in place by the patient to cope with it (e.g. maintaining speech rate at the expense of articulatory precision).

The principle of the experimental paradigm used here is to induce in the participants controlled variations in speech rate and pitch register, linked to instructions associated with a succession of experimental tasks or situations involving spoken communication. In addition to the clinical perspectives, this research aims at better understanding the cognitive and linguistic mechanisms that support the production of speech, conceived as a strategic behavior.

2. MATERIAL AND METHODS

2.1. Participants and speech material

Twelve participants, six men and six women between the ages of 59 and 79, took part in the study. Half of them had Parkinson's disease. Relevant biographical characteristics are collected in Table 1.

The corpus consisted of 25 sentences, 5 initial sentences (produced by a 22-year-old native French speaker) each having undergone 5 types of modifications. These modifications were performed via the "Manipulate" procedure in Praat and concern either the total duration (75%, 100%, 125% of the initial duration) of the sentence, or its average pitch (75%, 100% or 125% of the initial average pitch). The initial phrase is thus coded D100P100 and the four other versions: D75P100, D125P100, D100P75, D100P125. The sentences were: "Lalie vend du lilas (*Lalie sells lilacs*), "Le loup vit dans les bois" (*The wolf lives in the woods*), "C'est un papa ou un papy?" (*Is he a daddy or a grandpa*?), "Il a une barbe ou une moustache?" (*Does he have a beard or a moustache*?), "Il a une grande bouche ou une petite bouche?" (*Does he have a big mouth or a small mouth*?). The sentences in the corpus vary in length (6 to 11 syllables), syllabic complexity and phonological content (e.g., the first two sentences are totally voiced), as well as in prosodic structure.

	Healthy speakers					
Speaker	S01	S02	S03	S04	S05	S11
Sex	F	F	F	M	M	M
Age	64	64	75	75	67	79
Education	Ba	Ba	Ma	Ba	Ma	Ma
	PD patients					
Speaker	S06	S07	S08	S09	S10	S12
Sex	F	F	F	M	M	M
Age	59	71	66	64	57	75
Education	Ba	Ma	Ba	Ma	Ba	Ma
MMSE	30	29	30	30		29
Medication	Azi, Pro	Pro	Pro	Sta, Azi	Mo, Azi	Pro
Year of diagnosis	2013		2018	2013	2014	2016
Prior SLT	No	Sw	Sw	No	No	No
Complaint	N/A	Tr	N/A	A.R.	N/A	Tr

Education: Ba/Ma = bachelor's/master's degree
Medication: Azi: Azilect; Pro: Prolopa; Sta: Stalevo; Mo: Modapar
Prior Speech-Language Therapy: Sw: swallowing
Complaint: Tr: Tremor; A.R.: Arm rigidity

Table 1: Participants' characteristics

2.2. Experimental paradigm and measurements

The experimental paradigm consisted of six successive steps, completed in a single experimental session. At each stage of the paradigm (except (iii)), the participant produced the 25 sentences of the corpus according to a specific task: (i) reading; (ii) repetition of an oral model (the "target" voice); (iii) interactive game; (iv) repetition of an oral model (same as (ii)); (v) repetition with explicit instruction to imitate the oral model; (vi) repetition with explicit instruction to inhibit as much as possible the imitation of the oral model. The 25 sentences were presented in the same pseudo-random order at each stage. The interactive game was a slightly adapted version of the board game "Who am I?", so only certain sentences from the corpus were likely to be produced, without guarantee. The games (at least four games per participant) were all played between the participant and one and the same interlocutor, the one to whom the target voice belonged. The speech productions during the game are not analyzed here.

Acoustic data were recorded using a cell phone (iPhone 8) placed 30 cm from the participant's mouth on a separate stand. The recordings were converted to .wav format (44100Hz, mono) and processed in Praat. A dedicated script allowed the alignment between the transcriptions of the produced sentences and the speech signal, as well as the taking of measurements under the supervision of the user. The measures are: the total duration of the sentence, then converted into speech rate (number of syllables per second), as well as the fundamental frequency (Hz) calculated every 5 ms, from which the median value was extracted and expressed in z-score. For each utterance of each participant, we thus have two indicators, which form the main dependent variables of the statistical analysis: the median pitch (z-score) and the speech rate (syll/s).

Moreover, for tasks (ii), (iv), (v), (vi), these two indicators were transformed into a "convergence rate". The principle is to express (in %) the proportion of the path covered in relation to the distance to be covered to reach the target. The absolute value (e.g., the average measured pitch for a given repetition of a given sentence) is expressed relatively to the target:

$$\frac{\text{Measured\_Pitch} - \text{Baseline\_Pitch}}{\text{Target\_Pitch} - \text{Baseline\_Pitch}} * 100$$

The baseline selected was the average of the values measured for the respective participant during the reading task. The convergence rate could indicate either a partial convergence (e.g. 25%) or almost total convergence (90-100%), or even an overshoot of the target (above 100%) or a divergence (negative value).

3. RESULTS

Two repeated measures analyses of variance (ANOVA) were first performed, one with speech rate (syll/s) as the dependent variable, the other with pitch (z-score). The within-subjects factor was the task (pregame repetition; postgame repetition; imitation; inhibition), the between-subjects factor concerned the characteristics of the stimuli (respectively, for rate: D75, D100, D125; for pitch: P75, P100 and P125). In these analyses, the participant was considered as a random factor and the group (control vs. PM) was not taken into account. Thus, this was an initial exploration of the data, all subjects combined.

Analyses revealed a significant effect of task and stimulus characteristics on rate and pitch, as well as a significant interaction of these two factors illustrated in Fig.1 (Rate: $F(4,40.003)=30.255$, $p<.001$; Pitch: $F(4,40.081)=10.272$, $p<.001$). In repetition tasks, speakers tended to follow the oral model, producing sentences with higher/slower speech rates in response to higher/slower speech rates (similarly for pitch). In the imitation task, the observed differences in production as a function of stimulus characteristics were exacerbated (slopes are steeper in Fig.1). In the inhibition task, the line is almost horizontal, which means that speakers were globally successful in maintaining a given level of rate or pitch, regardless of those characterizing the voice that pronounced the sentences to be repeated.

In a second step, the effect of group (participants with and without Parkinson's disease) was studied. A multivariate analysis of variance was performed with convergence rates calculated for flow and pitch as dependent variables and task (repetition: includes pre- and post-game repetition tasks; imitation; inhibition) and group (control vs. PD) as independent variables.

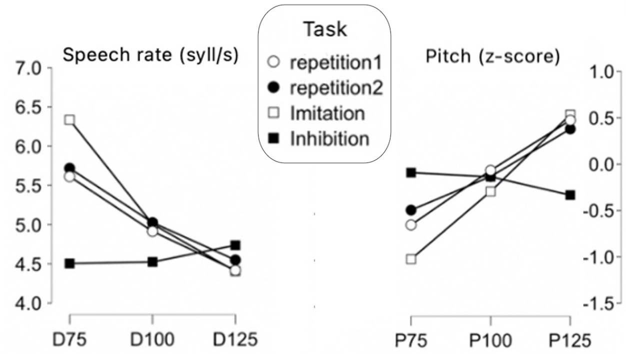

Figure 1: Speech rate (syll/s) (left) and pitch (z-score) (right) as a function of task and stimulus properties. All participants combined.

Regarding the convergence rate on speech rate, the analysis revealed only a significant effect of task ($F(2,1045)=6.713$, $p= .001$). There was no group effect, i.e. subjects with PD behaved like control subjects for speech rate (Fig.2, left): there was low convergence in repetition, high convergence (about 50% on average) in the imitation task and high divergence rates in inhibition.

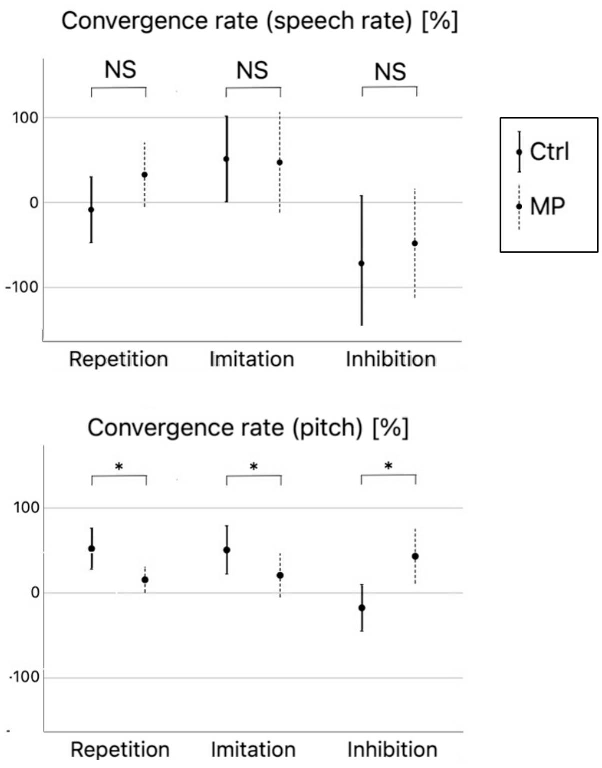

Figure 2. Convergence rate (%) as a function of task and participant group. Left: Speech rate, right: Pitch.

Regarding the convergence rate on pitch, only the Group*Task interaction was significant (F(2,1045)=7.939, p< .001). Regarding the convergence rate on pitch, only the Group*Task interaction was significant (F(2,1045)=7.939, p< .001). Fig.2 (right) illustrates this interaction: speakers in the control group converge to the pitch of the target voice, in both repetition and imitation, and manage to counteract this tendency in inhibition. In contrast, speakers with Parkinson's disease are unable to follow the given instructions, especially in the inhibition task: they approach (slightly) the pitch of the target voice in all tasks.

Note that when we added Sex to the independent variables in the statistical analysis, we observed a significant triple interaction Sex*Group*Task for the convergence rate on pitch (F(2,1039)=7.915, p< .001). This interaction reflects the fact that the trend shown in Fig.2 (right) concerned exclusively female subjects. Men, with or without PD, did not converge on pitch, regardless of the experimental task.

4. DISCUSSION

The aim of this study was to explore phonetic flexibility in French-speaking participants over 60 years of age, half of whom have Parkinson's disease. The paradigm used allowed us to study the fluctuations of the phonetic parameters of speech productions in relation to the production conditions. On the one hand, the instructions varied according to the task: simple repetition, explicit imitation, repetition without imitation, of the model voice. On the other hand, the sentences to be produced were manipulated in such a way as to obtain 5 versions of the same sentence, differing only in the speech rate or the pitch register.

First, the results showed that the subjects in the study were overall capable of phonetic flexibility. Their productions varied in the expected direction. When asked to simply repeat the sentences produced by a target voice, participants tended to imitate the speech rate as well as the fundamental frequency of the target voice, in a moderate way. This result, typical of the shadowing task, can be interpreted as a latent phenomenon of phonetic convergence, which occurs even in non-interactive communication situations [16]. When the instruction was explicitly to imitate the target voice, the convergence was more marked (the average convergence rate was close to 50% in imitation for the control subjects, for both speech rate and pitch). In contrast, convergence was close to zero, or even turned into divergence (speech rate), during the voluntary inhibition of imitation task, confirming the ability of the participants to actively adapt their speech productions in order to follow a given instruction. Comparative analysis of convergence rates as a function of conditions (shown in Fig.2) suggests that speech rate was "easier" for participants to manipulate at will than the mean register of f0.

Overall, this study confirms previous results obtained on a population of elderly French-speaking subjects via a phonetic learning paradigm [12]. The two experimental paradigms are different, as well as the subjects who participated in the two studies, but the finding at the end is similar: even elderly speaker-listeners are able to adapt in real time their speech productions in order to accomodate the internal and external constraints that frame the communication situation. Theoretical implications are numerous, especially for the study of the effects of aging on speech motor control (see e.g. [17]).

The study presented here also provides an exploratory approach to the question of the potential preservation of phonetic flexibility resources in people with PD. The exploratory nature is due to the small number of participants studied to date, with additional subjects currently being recruited to complete the numbers. Initial results suggest that PD participants exhibit phonetic flexibility, as they do not behave significantly differently from controls, at least for speech rate. On the other hand, PD subjects were found to be unable to converge to (or diverge from) the fundamental frequency of the target voice when instructed. This result could be interpreted in light of the prosodic insufficiency typically observed in these patients [18]. Recall, however, the interaction found with Sex: men, with or without PD, did not converge on the pitch, regardless of the experimental task. It is indeed worth noting that the target voice was a female voice (mean pitch at P100: 209 Hz), which makes the imitation/inhibition task radically different for male participants compared to female participants. The small number of participants does not allow at this stage to generalize the effects highlighted by the statistical analysis on our sample: is the lesser convergence on the pitch an effect of gender, an effect of the disease, or more globally an expression of individual variation? Future analyses on a larger sample should answer this question.

In any case, the phonetic flexibility demonstrated by PD patients with respect to speech rate (whereas deficits related to the temporal properties of speech are typical of Parkinsonian dysarthria [15,18]) opens interesting clinical perspectives. The speech therapy management of these patients – unfortunately not very systematic in the typical care pathway – could be based on these resources in order to develop compensatory strategies best adapted to their motor speech disorders.

In conclusion, the exploratory study presented here confirms previous results which suggest that phonetic flexibility capacities are maintained after 60 years of age. These abilities would also be preserved, at least in part, in people with Parkinson's disease. Thus, the production-perception link would remain effective, mobilizable in order to respond to the constraints of the communication situation, and this despite the presence of a motor execution disorder more or less extensive in these patients.

5. ACKNOWLEDGEMENTS

This study was partially funded by the Action de Recherche Concertée ARC "EvalDy" Grant (Wallonia-Brussels Federation, UMONS), Belgium. Thanks to Yousra El Omari for her help in data collection.

6. REFERENCES

[1] Pisanski K., Fraccaro P., Tigue C., O'Connor J., Röder S., Andrews P., Fink B., Debruine L., Jones B., Feinberg D. 2014. Vocal indicators of body size in men and women: A meta-analysis. *Animal Behaviour* 95, 89–99.

[2] Grandjean D., Baenziger T. 2019. Chapitre 4. Expression vocale des émotions. In : Sander D. (ed.), *Traité de psychologie des émotions* (pp. 109-155). Paris:Dunod.

[3] Rouas J.-L., Shochi, T., Guerry M., Rilliard A. 2019. Categorisation of spoken social affects in Japanese: human vs. machine. *Actes de l'International Congress of Phonetic Sciences ICPhS*, Aug 2019, Melbourne, Australia. ffhal-02317743

[4] Huttunen K., Keränen H., Väyrynen E., Pääkkönen R., Leino T. 2011. Effect of cognitive load on speech prosody in aviation: Evidence from military simulator flights. *Applied ergonomics* 42, 348-57.

[5] Revis J., De Looze C., Giovanni A. 2013. Vocal flexibility and prosodic strategies in a professional impersonator, *Journal of Voice* 27, 524.e23- 524.e31.

[6] Pardo J. S., Urmanche A., Wilman S., Wiener J., Mason N., Francis K. & Ward M. 2018. A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics* 69, 1-11.

[7] Cooke M., King S., Gargnier M., & Aubanel V. 2014. The listening talker: A review of human and algorithmic context-induced modifications of speech. *Computer Speech & Language* 28, 2, 543-571.

[8] Garnier M., Henrich Bernardoni N. 2014. Speaking in noise: How does the Lombard effect improve acoustic contrasts between speech and ambient noise? *Computer Speech & Language*. Doi: 10.1016/j.csl.2013.07.005.

[9] Moon S.J., Lindblom B. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *J. Acoust. Soc. Am.* 96, 40–55.

[10] Smiljanic R., Bradlow A.R. 2008. Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics* 36, 91-113.

[11] Ménard L., Perrier P., Aubin J. 2016. Compensation for a lip-tube perturbation in 4-year-olds: Articulatory, acoustic, and perceptual data analyzed in comparison with adults. *J. Acoust. Soc. Am.* 139(5), 2514.

[12] Delvaux V., Cano-Chervel J., Huet K., Leclef C., Verhaegen C., Harmegnies B. 2015. Capacités d'apprentissage phonétique et vieillissement, in : A.-M. Etienne (ed). *Evolutions sociales, Innovations et Politiques: nouvelles questions pour la psychologie de la santé*, Mardaga

[13] Verhaegen C., Delvaux V., Huet K., Fagniart, S., Piccaluga M. Et Al. 2020. Capacités d'apprentissage phonétique chez des patients aphasiques francophones: étude de cas. *Actes des 33e Journées d'Études sur la Parole (JEP),* 2020, Nancy, France, 606-616.

[14] Delvaux, V., Harmegnies, B., Huet, K., Piccaluga, M., Roland, V., Verhaegen, C. (2022) Les personnes atteintes de la maladie de Parkinson font preuve de flexibilité phonétique. *Proc. XXXIVe Journées d'Études sur la Parole*, JEP 2022, Noirmoutier, 289-297.

[15] Rusz J., Cmejla R., Ruzickova H., Ruzicka E. (2011). Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J. Acoust. Soc. Am*. 129, 350–367.

[16] Delvaux V., Soquet A. 2007. The Influence of Ambient Speech on Adult Speech Productions through Unintentional Imitation. *Phonetica* 64, 145-173.

[17] Fougeron C., Guitard-Ivent F., Delvaux V. 2021. Multi-Dimensional Variation in Adult Speech as a Function of Age. *Languages* 6(4), 176.

[18] Skodda S., Visser W., Schlegel U. 2011. Gender-Related Patterns of Dysprosody in Parkinson Disease and Correlation Between Speech Variables and Motor Symptoms. *Journal of Voice* 25(1), 76–82.