ARTICULATORY BEHAVIOUR DURING DISFLUENCIES IN STUTTERED SPEECH

Ivana Didirkova, Sébastien Le Maguer, Fabrice Hirsch, Dodji Gbedahou

ARTICULATORY BEHAVIOUR DURING DISFLUENCIES IN STUTTERED SPEECH

Ivana Didirková[1,2], Sébastien Le Maguer[3], Fabrice Hirsch[4], Dodji Gbedahou[4]

[1] EA 1569 TransCrit, Université Paris 8, France,
[2] CNRS & UMR 7018 Laboratoire de phonétique et phonologie, Université Paris 3, France
[3] ADAPT Centre, Sigmedia Lab, EE Engineering, Trinity College Dublin, Dublin, Ireland
[4] CNRS & UMR 5267 Praxiling, Université Paul-Valéry Montpellier 3, France
ivana.didirkova@univ-paris8.fr, lemagues@tcd.ie, fabrice.hirsch@univ-montp3.fr, hubertgbedahou@yahoo.fr

ABSTRACT

The aim of this study is to analyse articulatory movements that occur during Stuttering-Like Disfluencies (SLD) and to propose a new classification of SLD based on supraglottic articulatory gestures. To carry out this study, ElectroMagnetic Articulography (EMA) data were collected within two Persons Who Stutter (PWS) reading two texts. All pathological disfluencies were identified in the production of PWS categorized as blocks, repetitions and prolongations. Results show four articulatory patterns occurring during the SLD: Reiterations of series of movements leading to sound(s) or syllable repetitions, global maintain of the articulatory posture, anarchical movements and a combination of above. While the first category only concerns repetitions, the three others can concern SLD categorized as repetitions, prolongations or blocks.

Keywords: stuttering; stuttering-like disfluencies; speech production; articulatory description.

1. INTRODUCTION

1.1. Stuttering

Stuttering can be defined as an alteration of speech fluency having negative implications on communication ([7]). More precisely, this disorder is considered as a motor trouble that momentarily stops speech flow. Several types of stuttering are mentioned in literature: developmental stuttering starting between age 3 and 7 and disappearing spontaneously, persistent stuttering beginning at the same period but remaining present in adolescence and adulthood; as well as acquired stuttering, generally due to a neurological accident ([7]). According to [15], 5% of the worldwide population have been concerned by this disorder but its prevalence is at 1% since the rate of 'spontaneous' remissions in children is evaluated at 80%. If the origins of stuttering remain a challenge for researchers, recent works allow formulating several hypotheses about the aetiology of developmental and persistent stuttering. Indeed, the origins of this trouble should be multi-factorial since different studies point out genetic and neurological specificities in Persons Who Stutter (PWS).

1.2. Phonetics of stuttering

As mentioned above, stuttered speech is characterized by the presence of disfluencies that are more frequent than in non-stuttered speech. These Stuttering-Like Disfluencies (SLD) can be classified mainly as blocks, prolongations and repetitions (e.g. [11]) but other types of speech flow alterations can be found in PWS (see [3] for a literature review).

Moreover, SLD present several specificities compared to non-pathological disfluencies. For example, stuttering is one of the scarce disorders where disfluencies can frequently split a syllable ([17]).

In another study, [4] show that alterations of speech flow by PWS are generally accompanied by audible tensions. The same research shows that consonants can be prolonged in stuttering-like disfluencies, while this is not the case in normal alterations of speech flow in French. Finally, they observe that the duration of SLD is generally more important and more variable.

1.3. Physiological description of SLD

However, classification of speech disfluencies as non-pathological or pathological is not an easy task given that several types of disfluencies are present both in non-stuttered and stuttered speech. For instance, sound prolongations, repetitions, as well as silences, can also be observed in people who do not stutter. This is the reason why physiological descriptions are necessary to determine what distinguish stuttering-like disfluencies and non pathological disfluencies.

Concerning the respiratory level, [13] observe that the respiratory movements during pre-phonatory phases are different in PWS. Other studies dealing with this topic have been carried out (e.g. [10], [18]). The laryngeal level also presents some specificities. Indeed, [2] observe an abnormal activity of the vocal folds during stuttering-like disfluencies. Therefore, [8] prefer to talk about myoclonic movements (spasms) to describe the glottis functioning in PWS.

Curiously, The literature concerning the supraglottic level in subjects who stutter remains scarse and often deal with fluent speech produced by PWS [5]. Among studies dealing with SLD, [6] reveal a deficiency in the jaw-phonatory connection. The speech motor behaviour in PWS would tend to be less efficient or even immature in the management of the coordination of different articulators. Thus, [14] supposes alterations of speech flow are due to a coarticulation disruption. More precisely, [14] estimates that, in a [CV] sequence, transition between the two sounds should be the consequence of a disrupted antagonist muscles activity. This fault line would correspond to the moment where stuttering-like disfluencies emerge.

1.4. Objective and hypothesis

As mentioned above, few studies have been carried out on the way disfluencies are produced. Furthermore, most of these studies are based on extrapolations made from acoustic data. However, while it is possible to obtain many informations thanks to the acoustic signal, EMA data allow a more direct observation.

Consequently, the aim of this study is to provide a description of articulatory behaviour during SLD. More precisely, our objective is to analyse articulatory movements that occur during SLD and to propose a classification based on supraglottic articulatory gestures. Our hypothesis is that the nomenclature generally used to describe disfluencies does not reflect the articulatory behaviour. If a same percept can be a result of different articulatory gestures ([12]), depending on speaker, phonetic environment, etc., we postulate that the same articulatory patterns could be at the origin of several perceptual types of disfluencies ([3]).

2. METHODS

2.1. Data acquisition & participants

EMA data were collected by means of an electromagnetic articulograph Carstens AG501 3D at the Lorraine Research Laboratory in Computer Science and its Applications (LORIA, Nancy, France) with a sampling rate of 250 Hz and an accuracy of 0.3 mm. All data were stocked in a .pos file and synchronized with a sound recording (44.1 kHz, 16 bits, .wav). 10 sensors (2x3 mm) per subject were used: two were fixed on the lips of each subject (1 in the middle of the upper lip and another one in the middle of the lower lip). 3 coils were situated on the tongue of each subject; one on the tongue tip, one on the tongue body and one on the tongue back. To track the mandible's movements, another sensor was placed on the subjects' jaw. The palate's form was indicated by means of a seventh coil. Other sensors were used to control head's movements.

Two PWS, one female and one male, aged respectively 23 and 26, both native speakers of French and Wolof, were recruited for this study. Participants were recorded while reading the text of an Alphonse Daudet's novel, *La chèvre de Monsieur Seguin* (Mister Seguin's goat), in French and an Aesop's fable, *Le lion et le rat* (The lion and the rat). These recordings took place in a soundproof room.

2.2. Data analysis

2.2.1. Acoustic and perceptual analysis

Data analysis rely on perceptual and acoustic identification of stuttering-like disfluencies. First, three persons (two of the authors and a speech therapist specialized in stuttering) identified all SLD in the production of PWS, based on perception and on the speech signal, without classifying these disfluencies. They then discussed cases where they did not reach agreement. In order to confirm their annotations and identify the perceptual class of every disfluency, a perception test has been carried out within the freeware Perceval ([1]), based on .wav files extracted for each SLD. Five naïve listeners were then asked to categorize SLD as blocks, repetitions, prolongations or combined disfluencies (Fleiss' kappa: 0.752). Authors discussed cases where naïve listeners did not reach agreement. Speech alterations identified as combined disfluencies were eliminated from further study. Moreover, repetitions of diphones, syllables, words and other sequences containing more than one phone were excluded from our research in order to minimize influence of coarticulation on our observations.

After exclusion of combined disfluencies and disfluencies concerning more than one phone, 250 SLD were obtained. Their distribution according to the perceptual type of disfluency and according to the subject can be found in the Table 1. Although both subjects have a severe stuttering, the distribution of disfluencies is not the same: whereas speaker F produces 89 disfluencies, mostly blocks and prolongations, 161 of disfluencies analysed in this paper were produced by M. For the speaker M, repetitions are the most present perceptual disfluency type. Due to these idiosyncratic characteristics of SLD in our speakers, only 55 disfluencies (22%) were blocks. Other 39.2% of disfluencies were prolongations. Repetitions represented 38.8% of all analysed SLD. All of SLD were spontaneous, e.g. no factors were manipulated to elicit these disfluencies.

Table 1: Distribution of stuttering-like disfluencies according to their perceptual type and according to the speaker

	Repet.	Prolong.	Block	**Total**
Female	17	34	38	89
Male	80	64	17	161
Total	97	98	55	250

2.2.2. Automatic articulatory analysis

We can assess if there was a movement during the production by inspecting the articulatory dynamics. To do so, we have defined the following methodology. First, we consider $t \in [0..T]$ the index of the frame and C_t the set of coils at frame t. From each coil $c_t \in C_t$, we compute the local velocity based the central finite difference as defined in [16] using Equation (1).

$$(1) \quad \Delta(c_t) = \begin{cases} 0 & t = 0 \\ 0 & t = T \\ \sqrt{\sum\limits_{i \in \{x,y,z\}} \frac{c_{t+1}(i) - c_{t-1}(i)}{2}} & \text{else} \end{cases}$$

Then, a movement at frame t, is detected if the following criterion is validated:

$$(2) \quad \sum_{c_t \in C_t} f(\Delta(c_t), \theta_c) > |C_t|/2$$

with $f(v, \theta) = 1$ if v is beyond θ and 0 else.

In this study, we define θ_c at 30% of the average dynamic of the whole corpus for each coil c. This large threshold allows us to be less sensitive to a movement and therefore enhance non-activity detection. Finally, we focus our analysis to the segments annotated as a disfluent production. For each segment, we ignore the 10% first and the 10% last frames in order to avoid transition effect. From the remaining ones, we compute the percentage of frames considered in movement.

2.2.3. Manual articulatory analysis

After this classification, the Visartico software ([9]) was used to visualize and analyse the vertical movements (the z axis) of the upper and lower lip, the tongue tip, the tongue body, the tongue back and the mandible in segments that included the stuttered phone and its preceding and subsequent phones.

3. RESULTS

3.1. Percentage of frames in movement by type of SLD

As we can see in Table 2, even though a threshold has been defined to capture a maximum of non-movement frames, there are still around 40% of them considered as moving in average. Furthermore, the standard deviation shows an important variability across the segments. Some segments are even reaching 80% of movement.

Table 2: Average movement percentage per type and per disfluency type. The standard deviation is indicated in parenthesis.

	Female	Male
Repetition	46.28 (13.15)	48.05 (12.91)
Prolong.	31.18 (13.43)	40.80 (9.60)
Block	42.26 (11.69)	47.18 (11.08)

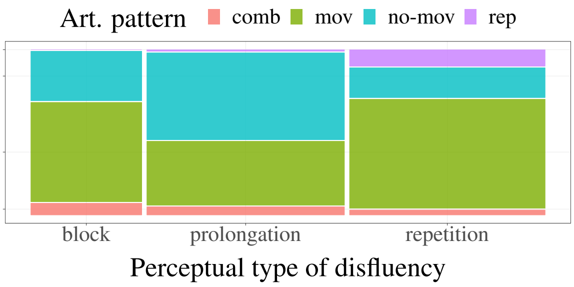
If we observe what happens for each type of SLD, it is possible to notice, in average, that repetitions are the disfluencies produced with the greatest amount of frames in movement in speakers F (average: 46.28%, SD: 13.15%) and M (average: 48.05, SD: 12.91%). Blocks constitute the second type of SLD where frames in movement are the most present in F (average: 42.26%, SD: 11.69%) and M (average: 47.18%, SD: 11.08%). Finally, less frames are in movement in prolongations (average: 31.18%, SD: 13.43 in F; average: 40.8%, SD: 9.6 in M).

3.2. SLD and articulatory patterns

Four main categories of disfluencies have been revealed by EMA data (a chi-squared goodness-of-fit test: χ_2=187.28, df=3, p=.000): *a*) Reiterations of series of movements leading to sound repetitions (*rep*); *b*) Combination of a global maintain of an articulatory posture and articulatory movements (*comb*); *c*) Global maintain of the articulatory posture with or without an acoustic output and with or without anticipation of the subsequent phone (*no–mov*); *d*) Presence of articulatory movements with or without inter-articulatory coupling (*mov*).

As shown in Figure 1, while the first category mostly concerns repetitions, the three others can concern SLD categorized as repetitions, prolongations or blocks, showing that a same articulatory pattern can be observed for the 3 types of disfluencies (chi-squared test for independence: χ_2=39.302, df=6, p=.000, effect size: 0.560).

Figure 1: Proportions of disfluency types. The area of each rectangle gives the proportion of the perceptual type (width) and the articulatory pattern (height).



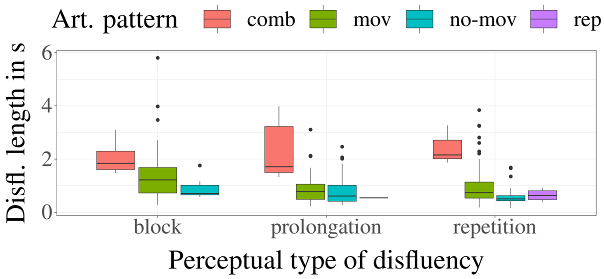
3.3. Duration and SLD type

A linear model was fit with articulatory pattern and perceived disfluency as the independent variables and length as the dependent variable. The model was significant: F=11.68 on 5 and 244 df, p=.000. Our data indicate a clear preference of the combined articulatory pattern to occur within the longest disfluencies. The duration decreases when the disfluency is characterized by the presence of a movement during whole disfluency. The shortest disfluencies are those where we observe a global maintain of articulatory posture and a repetition of an articulatory movement. These effects are mostly prominent in blocks as shown in Figure 2.

4. DISCUSSION AND CONCLUSION

To sum up, most of SLD are carried out with frames in movement. If movements are generally observed, it is possible to note that their 'efficiency' is variable: indeed, if some of them are audible during prolongations and repetitions for instance, others are ineffective since they are inaudible, as in blocks. It

Figure 2: Types of disfluencies and their length.



is important to highlight that movements' efficiency during SLD can be due *a*) to the degree of constriction between the different articulators and *b*) to the respiratory and/or the laryngeal level. Indeed, if air pressure and/or vocal folds configuration are not adapted, the acoustic output will be absent.

Moreover, several types of articulatory patterns have been observed during disfluencies. These patterns can be divided in two categories: those which are carried out with slight vertical movements or an immobilization of most articulators, and those produced with movements. Among the last category cited, there are SLD presenting inter-articulators coupling and SLD where articulators move independently of each other. These different patterns are present for blocks, prolongations and repetitions. In other terms, it means that a same type of disfluencies can be produced in different ways. This allows to draw a parallel between disfluencies and the Quantal theory [12]; This theory supposes that a same percept can be the results of several different articulators' positions. As for disfluencies, a same disfluency can be the result of different configurations.

Concerning articulatory patterns, it has been shown that the longest disfluencies are carried out with a combination of several articulatory configurations. This result suggests that more there are different articulatory patterns during a disfluency and longer the disfluency will be. Consequently, making an effort seems to be ineffective in PWS when they have to overcome a disfluency.

Our study reports articulatory patterns seen in their totality (all articulators taken together). Thus, it seems necessary to investigate the contribution of each articulator to observed patterns. Our result should finally be compared to an articulatory description of disfluencies of non-stuttering speakers.

Finally, we think that this research should be carried out in a longitudinal approach in order to verify how articulatory patterns progress during a speech therapy, showing that articulatory data should be used more frequently during a stuttering reeducation to note patient's evolution.

5. ACKNOWLEDGMENT

This research was partly supported by « MoSpeeDi. Motor Speech Disorders : characterizing phonetic speech planning and motor speech programming/execution and their impairments », subside CRSII5_173711/1Sinergia du Fond National Suisse de la Recherche Scientifique

This research was partly supported by the French *Agence Nationale de la Recherche* and by the *Caisse nationale de solidarité pour l'autonomie* under Grant No. ANR-18-CE36-0008 (Project BENEPHIDIRE, PI: Fabrice Hirsch).

This research was also partly supported by the Irish Research Council (IRC) and by the ADAPT Centre. The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and isco-funded under the European Regional Development Fund.

6. REFERENCES

[1] André, C., Ghio, A., Cavé, C., Teston, B. 2003. PERCEVAL: a Computer-Driven System for Experimentation on Auditory and Visual Perception. *International Congress of Phonetic Sciences (ICPhS)* Barcelona, Spain. UAB 1421–1424.

[2] Conture, E. G., Schwartz, H. D., Brewer, D. W. 1985. Laryngeal behavior during stuttering: A further study. *Journal of Speech, Language, and Hearing Research* 28(2), 233–240.

[3] Didirkova, I. 2016. *Parole, langues et disfluences: une étude linguistique et phonétique du bégaiement*. PhD thesis Université Paul Valéry-Montpellier III.

[4] Didirkova, I., Fauth, C., Hirsch, F., Luxardo, G., Diwersy, S. 2016. Disfluences normales vs. disfluences sévères: une étude acoustique. *JEP-Journées d'Etudes sur la Parole* volume 1 191–199.

[5] Heyde, C. J., Scobbie, J. M., Lickley, R., Drake, E. K. 2016. How fluent is the fluent speech of people who stutter? a new approach to measuring kinematics with ultrasound. *Clinical linguistics & phonetics* 30(3-5), 292–312.

[6] Loucks, T. M., Luc, F., Sasisekaran, J. 2007. Jaw-phonatory coordination in chronic developmental stuttering. *Journal of Communication Disorders* 40(3), 257–272.

[7] Monfrais-Pfauwadel, M.-C. 2014. *Bégaiement, bégaiements. Un manuel clinique et thérapeutique*. De Boeck-Solal.

[8] Monfrais-Pfauwadel, M.-C., Tromelin, O., Mougin, A.-L., Ormezzano, Y. 2005. Utilisation des explorations multimédia synchrones dans l'objectivation des événements laryngés lors des bégayages. *Revue de laryngologie, d'otologie et de rhinologie* 126(5), 341–345.

[9] Ouni, S., Mangeonjean, L., Steiner, I. Sept. 2012. VisArtico: a visualization tool for articulatory data. *13th Annual Conference of the International Speech Communication Association - InterSpeech 2012* Portland, OR, United States.

[10] Peters, H. F., Boves, L. 1988. Coordination of aerodynamic and phonatory processes in fluent speech utterances of stutterers. *Journal of Speech, Language, and Hearing Research* 31(3), 352–361.

[11] Piérart, B. 2011. *Les bégaiements de l'adulte* volume 5. Editions Mardaga.

[12] Stevens, K. N., Hanson, H. M. 2012. Chapter 12: Articulatory–acoustic relations as the basis of distinctive contrasts. *The Handbook of Phonetic Sciences* 116, 424.

[13] Watson, B. C., Alfonso, P. J. 1987. Coordination of prephonatory events in mild and severe stutterers. In: *Speech Motor Dynamics in Stuttering*. Springer 197–207.

[14] Wingate, M. E. 2012. *The structure of stuttering: A psycholinguistic analysis*. Springer Science & Business Media.

[15] Yairi, E., Ambrose, N. 2013. Epidemiology of stuttering: 21st century advances. *Journal of fluency disorders* 38(2), 66–87.

[16] Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D., others, 2006. *The HTK book (v3. 4)*.

[17] Zellner, B. 1992. Le bé bégayage et euh... l'hésitation en français spontané. *JEP* volume 19 481–487.

[18] Zocchi, L., Estenne, M., Johnston, S., Del Ferro, L., E. Ward, M., T. Macklem, P. 07 1990. Respiratory muscle incoordination in stuttering speech. *The American review of respiratory disease* 141, 1510–5.