



**HAL**  
open science

## **Working memory impairments in Multiple Sclerosis: effects on stress clash resolution.**

Caterina Petrone, Céline de Looze, Léa Gomis, Marjorie Wims, Alain Ghio,  
Noémie Moreau, Laurent Renié

### ► **To cite this version:**

Caterina Petrone, Céline de Looze, Léa Gomis, Marjorie Wims, Alain Ghio, et al.. Working memory impairments in Multiple Sclerosis: effects on stress clash resolution.. International Seminaire of Speech Production, May 2014, Cologne, Germany. pp.4. hal-01510226

**HAL Id: hal-01510226**

**<https://hal.science/hal-01510226>**

Submitted on 20 Apr 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Working memory and stress clash resolution in Multiple Sclerosis

Caterina Petrone<sup>1,2</sup>, Céline De Looze<sup>3</sup>, Léa Gomis<sup>2</sup>, Marjorie Wims<sup>2</sup>, Alain Ghio<sup>1,2</sup>, Noémie Moreau<sup>1,4</sup> & Laurent Renié<sup>1,4</sup>

<sup>1</sup>Laboratoire Parole et Langage, CNRS, <sup>2</sup>Aix-Marseille Université, France

<sup>3</sup>Phonetics and Speech Laboratory, CLCS, Trinity College Dublin, Ireland,

<sup>4</sup>Neurology Department, PAYS D'AIX HOSPITAL, Aix-en-Provence, France

caterina.petrone@lpl-aix.fr, deloozec@tcd.ie

## Abstract

In this paper, we investigate stress clash resolution, i.e., how speakers readjust the metrical patterns of an utterance to preserve its metrical regularity. We test the effect of working memory (WM) capacities in control (C) subjects and individuals with Multiple Sclerosis (MS) on their ability to look-ahead to the metrical structure of an utterance and compute changes in prominence structure in the clash condition. We predict that if WM capacities affect stress clash resolution, MS patients with low WM capacities will not be able to look ahead and will fail to apply stress resolution. Our results do not provide evidence of stress clash resolution for both C and MS subjects, suggesting neither stress shift nor stress deletion took place. Yet, some evidence is provided that MS speakers with low WM capacities narrow down the scope of the prosodic planning unit, thus suggesting that variation in planning strategies depend on cognitive skills. We discuss our results along different models of French prosody.

**Keywords:** stress clash, Multiple Sclerosis, working memory, prosodic structure, French.

## 1. Introduction

When speakers are to produce an utterance, they plan what they want to say and the way they want to say it, e.g. they choose the segments, the words, the prosody of that utterance. Although the incremental character of language production is uncontroversial (Levelt, 1989), it is still difficult to determine how far ahead speakers do plan and whether advance planning might differ at different representational levels (Wheeldon & Lahiri, 1997; Krivokapić, 2007). It has been suggested that the scope of planning does not coincide with a fixed linguistic unit, but might be flexibly adapted by the speakers (Swets et al., 2007). In particular, speaker-specific variations would reflect differences in working memory (henceforth WM) capacities. "Working memory" refers to the memory subsystem which is responsible for the active maintenance of mental representations along ongoing processing and/or distractions. Swets et al. (2007) argues that there is an effect of WM capacities on prosodic structure, such that readers with low WM are more likely to chunk a text into smaller prosodic phrases than readers with high WM. Petrone et al. (2011) found that German speakers with high WM capacities start their utterances with higher F0 peaks than speakers with low WM capacities in order to apply F0 declination (the gradual decrease of F0 over the course of an utterance) on larger portions of the utterance.

In this study, we investigate working memory constraints on underlying variability in the scope of prosodic planning by comparing healthy subjects with individuals affected by cognitive disorders due to multiple sclerosis (MS). MS is a

progressive inflammatory disease of the central nervous system which damages the myelin sheath. MS is characterized by cognitive deficits (such as WM impairments, cf. Fuso et al., 2010) which can occur independently of other disorders. A preliminary study on French (Petrone et al., 2013) found that MS patients affected by working memory impairments (but **not** by speech motor control disorders) were not capable of planning long prosodic phrases in French. Patients with low WM produced more prosodic breaks in unpredictable locations and their speech rate was slower than that of patients with high WM.

Since MS patients with low WM show difficulties in the temporal organization of their speech, we predict that they will also have difficulties in planning other aspects of prosody, such as the metrical structure. In the current paper, we focus on stress clash resolution, i.e., how speakers readjust the metrical patterns in a sequence [word1 word2] in order to preserve metrical regularity (Nespor & Vogel, 1986; cf. Tilsen, 2012 for a review). Phonologically, metrical regularity is the consequence of the rhythm rule, by which a stress is shifted at an earlier location in the word or deleted to avoid clash. A case in which the rhythm rule applies is when a word with late stress like JapaNESE creates a clash when followed by a word with early stress like INstitute (Keating & Shattuck-Hufnagel, 2002). The clash is resolved by reducing the prominence of the last syllable of word1 (e.g., by decreasing its duration and intensity and by deleting or shifting the pitch accent) and eventually by making the first syllable of word1 more prominent (e.g., by means of a pitch accent).

In French, the word-final (full) syllable is usually the most prominent and a primary accent can be assigned to this position. Stress clash has been traditionally defined as the occurrence of two consecutive primary accents (Post, 1999). Stress clash resolution is claimed to be obligatory when the two accents are within the same phonological phrase  $\phi$  (Post, 1999). For instance, in a sequence [word1 word2] $\phi$  such as "joLIS LYS" (*nice lilis*), accent resolution involves the deletion ("jolis liLAS") or the shift ("joLIS liLAS", *nice lilacs*) of the pitch accent of the word1 final-syllable (Post, 1999; Avanzi & Schwab, 2013). Moreover, the duration of the word1-final syllable is shorter when the stress clash resolution applies (Avanzi & Schwab, 2013).

Different accounts of stress clash in French have been proposed. In particular, Jun & Fougeron (2000) claimed the existence of an Accentual Phrase (AP), whose underlying pattern is /LHi LH\*/. In this account, clash resolution would be reinterpreted in terms of variation in the tonal pattern due to a specific constraint, by which, for instance, the AP final H\* would be realized as a L\* to avoid a sequence of three consecutive H tones (AVOID HHH). However, the syllable associated to L\* would still be lengthened.

Stress clash is a good candidate to test the effects of WM on prosodic planning, since clashes are not stored in lexical memory and thus their resolution require a more on-line building of the prosodic structure (Tilsen, 2012). When producing word1, the speaker must already take into account the metrical pattern of the upcoming word2 to apply the rhythm rule (Keating & Shattuck-Hufnagel, 2002). If WM capacities affect stress clash resolution, both healthy speakers and MS patients with high WM capacities will be able to look-ahead to the metrical structure of following words and compute changes in prominence structure in the clash condition. Following Post (1999), we predict that speakers will reduce the duration of the final syllable of word1 and they will either delete the accent or shift it from word1 to word2. On the contrary, the scope of planning will be narrower in MS patients with low WM spans and their look-ahead will be reduced. As a consequence, they will fail to apply stress clash resolution. The accent on word1 final syllable will be similar in clash and no-clash conditions or the prosodic structure of the sequence will be affected (e.g. prosodic break, hesitation inserted between word1 and word2 or within word1 and word2).

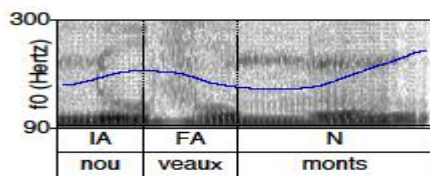


Figure 1: Pitch track for the clash sequence “nouveau monts” (*new mountains*). Straight lines indicate syllable boundaries.

## 2. Methods

### 2.1. Participants

Thirty MS patients (22 women and 8 men) and thirty-one healthy (22 women and 9 men) speakers participated in the experiment. The mean age is 43 y.o. for both the MS (ranging from 25 to 63 y.o.) and the control group (29-69 y.o.). The two groups were also matched in educational level. Results for twenty speakers analyzed are reported here, i.e., ten speakers for each group. The inclusion criteria are detailed below.

#### 2.1.1. Multiple sclerosis group (MS)

MS patients meeting with a clinically definite relapsing-remitting form of multiple sclerosis (i.e., periods of acute inflammation in which a deficit appear, followed by periods in which the patient eventually recovers from the deficit, cf. McDonald et al. 2001) took part in the study. Patients were included with (1) no clinical relapses and steroid treatment at least one month prior to the study; (2) no concomitant therapy with antidepressant or psychoactive drugs; (3) optimal visual acuity; (4) no dyslexia; (5) no motor disorders (as assessed by structural MRI scanning and clinical examination).

The experiments were run in the CNRS rooms at the Neurology Department in Pays d’Aix Hospital (Aix-en-Provence), in the anechoic room of the Laboratoire Parole et Langage and at the Hospital La Timone (Marseille). The participants agreed with the experimental procedure and signed an informed consent. Experiments were conducted under ethical conditions.

#### 2.1.2. Control group (C)

Speakers with no hearing, speech and motor disorders were selected.

### 2.2. Working memory tasks

Prior to the acoustic recordings, MS and healthy speakers underwent a battery of cognitive assessments including a series of working memory span tasks. The span tasks are based on the forward and backward recall of digits as well as on the recall of letters and number sequences (Brissart et al., 2008). The administration time of this battery was 15 minutes. The letter-number sequencing task better discriminated the two groups [ $\beta = -0.33$ ,  $SE = 0.13$ ,  $z = -2.57$ ,  $p = 0.01$ ]. In this task, random sequences of letters and numbers are presented by the experimenter, from 2 to 8 letter-number combinations. Participants have to remember and repeat them by, first, repeating the numbers in ascending order, then the letters in alphabetical order (e.g., T-9-A-3; correct response is 3-9-A-T).

Only correlations between the WM score obtained with this task and the acoustic parameters are described below. Raw scores on this WM span task were converted to standard scores (between 1 and 19). Participants were then split into high vs. low WM capacity sub-groups depending on whether their WM score was above vs. below the mean WM score of the group to which they belong to (MS = 9; C = 11). As expected, the mean WM score also differed for high (MS = 6.8; C = 8) and low (MS = 13.2; C = 16) WM speakers across the two groups.

### 2.3. Acoustic task

A reading corpus was collected, which consists in three sets of sentences composed of Subject-Verb-Object phrases (S-V-O) such as in “Mon frère regarde les jolis lys dans le parc” (*My brother is looking at the beautiful lilies in the park*). The O constituent contains a sequence [word1 word2] $\phi$  in which word1 is a bisyllabic adjective and word2 is the head noun. In the clash-condition (CL), word2 is monosyllabic and the two primary accents are close to each other (“joLIS LYS”, *nice lillis*); in the no-clash condition (no-CL), word2 is bisyllabic and one unaccented syllable intervene between the two accented ones (“joLIS liLAS”, *nice lilacs*). The two target words were always placed within the same phonological phrase in order to induce speakers to solve clash in the clash condition. The target sentences were interspersed among fillers and presented three times randomly. The acoustic experiment lasted 15 minutes. To sum up, a total of 960 observations were analyzed (20 speakers x 2 groups x 2 clash conditions x 4 sets x 3 times). Both MS and C groups were recorded by means of the EVA2 workstation (Ghio et al., 2012). The acoustic task lasted 15 minutes for each speaker.

### 2.4. Acoustic annotation

The syllables in word1 and word2 were manually aligned and annotated. The two syllables in the adjective (A) are labeled IA and FA, respectively for their Initial and Final position in the Adjective.

The sequence [word1 word2] $\phi$  containing dysfluencies were annotated R and were excluded from the statistical analyses. In total, 6R are found for the control group C (among 4 speakers) and 23R for the MS group (among 7 speakers). R stands for omissions (RO; *jolis -las*), hesitations (RH; *jolis heu lilas*), lengthening (RA; *de nouveaux mmmmmonts*), repetitions (RR;

jo-jolis lilas), prosodic breaks (RB; [jolis][pause][lilas]) and errors (RE; beaux lilas).

## 2.5. Acoustic measurements

Acoustic measurements were automatically carried out using the Praat software (Boersma & Weenink, 2014). For each syllable, the syllable duration (ms) and the f0 maximum (maxf0, Hz) were measured. The ratios between IA and FA maxf0 and syllable duration (ratioPitch\_IA/FA, ratioDur\_IA/FA) were also computed. A ratio equals to 1 indicates similar duration and maxf0 while a ratio > 1 a shorter duration and lower maxf0 FA and a ratio < 1 a longer duration and higher maxf0 for FA. In no-CL condition, FA is expected to be longer and higher than IA. In CL condition, FA is expected to be shorter and lower than IA. As a consequence, the ratio is expected to be < 1 in no-CL and >1 in CL.

## 2.6 Statistical analyses

A series of linear mixed models were carried out to test the effect of WM capacities (high vs. low), Group (C vs. MS) and Condition (CL vs. no-CL) on the duration and f0 of the word1 last syllable (FA) and on the ratios IA/FA. For f0 analysis, Gender (male vs. female) was also included as the fourth fixed factor. The f0 data were log-transformed to meet normality assumption. Speakers and sentence sets were the random intercepts. Forward selection of random slopes and intercepts based on likelihood-ratio tests was used to decide which random terms should be retained in the models. Non-significant interactions among fixed factors were also excluded based on likelihood-ratio tests. The standard mixed models do not provide p-values when random slopes and intercepts are correlated. We assumed that a fixed factor is significant if its *t* value is greater than 2 (Baayen et al., 2008).

## 3. Results

### 3.1 Duration

Fig. 2 shows the FA duration across group, WM and clash conditions. No significant differences on FA duration between CL and no-CL conditions were found in C group, independent of WM capacities. Similarly, MS speakers with high WM showed no significant differences across the clash conditions. However, for MS speakers with low WM, FA duration is significantly *longer* in CL (246 ms) than in no-CL conditions (188 ms). [ $\beta = 0.014$ ,  $SE = 0.006$ ,  $t = 2.325$ ].

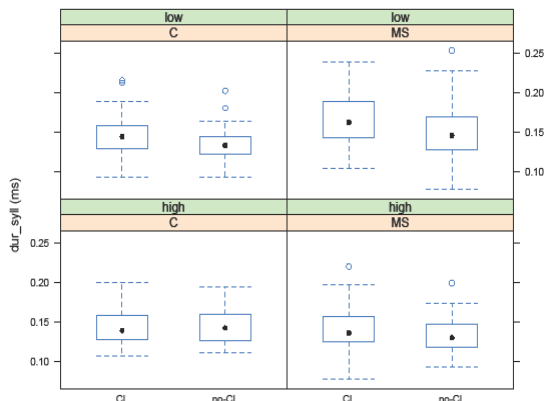


Figure 2: FA duration in regards to group (C vs. MS) and WM capacities (high vs. low).

Concerning the ratioDur\_IA/FA, no significant differences in duration between IA and FA in regards to conditions, groups and WM.

### 3.2 F0

Our analyses on the FA maxf0 do not show any effect of Group, WM and Condition. The mean F0 maximum was in the upper F0 range for both female (C = 217 Hz; MS = 234 Hz) and male (C = 147 Hz; MS = 100 Hz) speakers, suggesting the presence of a pitch peak on FA. Fig. 3 shows that ratioPitch\_IA/FA is around 1 in the CL condition, meaning that the maxf0 had similar value in IA and FA. However, the ratio was significantly lower in no-CL than in CL for both MS and C subjects [ $\beta = -0.07$ ,  $SE = 0.01$ ,  $t = -5.33$ ]

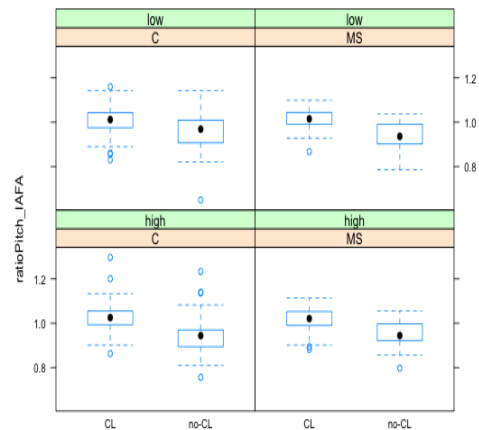


Figure 3: RatioPitch\_IA/FA in regards to group (C vs. MS) and WM capacities (high vs. low).

## 4. Discussion and conclusion

Our results do not provide evidence of stress clash resolution in the way predicted by Post (1999), both for duration and f0 measurements. First, the duration of FA was similar in both clash and no-clash condition for the control group, suggesting no changes in the degrees of stress. Moreover, contrary to our expectations, FA duration was even longer in the clash than in the no clash condition for the MS group with low WM, indicating that other prosodic strategies were adopted by these speakers. The ratioDur\_IA/FA also showed no difference across the two clash conditions between the initial and final syllable of the adjective. This further supports the idea that neither stress-shift nor stress deletion took place. Moreover, if clash resolution would have applied at tonal level, the accent on the final syllable of word1 should have been deleted or shifted to the initial syllable. However, no significant differences in maxf0 within FA were found in the clash and no-clash condition. This was true for both the C and MS groups, independent of WM capacities. The ratioPitch\_IA/FA showed that, for both groups, the maxf0 has the same value for both IA and FA in the clash condition, whereas in the first syllable carried a higher F0 value in the no-clash condition.

The negative results for clash resolution can be explained by taking into account other models of French prosody than Post (1999). Jun & Fougeron (2000) claimed that the lower constituent of French prosody is the Accentual Phrase (AP). The AP is the domain of primary stress and is characterized by an obligatory final rise (LH\*), preboundary lengthening and an optional initial rise (LHi). Phonetically, the H\* of the final rise

is consistently realized at the end of the last full syllable of the AP. The Hi of the early rise is not anchored to a specific landmark but it can be realized on the first or second content word syllable (Welby, 2006). The /LHi LH\*/ can be fully realized or realized in different phonetic variants, depending on several factors such as the speech rate or the speaker.

It is possible that the sequence [word1 word2] has been mostly uttered as a single AP with the default /LHi LH\*/ pattern in both clash and no-clash conditions (Fig. 4). Specifically, speakers of the C group and speakers of the MS group with high working memory could have consistently placed an initial Hi on word1 since in both conditions there is no violation of AVOID HHH. Given the ‘loose’ association of Hi with the left edge of the AP, its phonetic could have variably realized on IA and FA. This could explain the high values of maxf0 in FA in both clash and no-clash conditions, as well as the lack of differences for ratioPitch\_IA/FA in the CL-condition. Moreover, the lack of durational results can be explained by the fact that the syllable carrying the initial rise is not marked by lengthening.

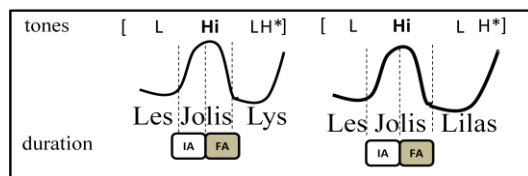


Figure 4: Schema of the possible prosodic pattern of stimuli with and without clash, produced by a healthy speaker.

However, a different prosodic pattern might have been chosen by MS speakers with low WM. For these speakers the final syllable was lengthened in the CL-condition. A possible hypothesis is that the adjective and the noun have been uttered in two separate APs: [word1]<sub>AP</sub> [word2]<sub>AP</sub>. As a consequence, FA may become the phrase-final syllable and, as such, it may be marked by a final rise and preboundary lengthening. In other words, MS speakers with low WM could have narrowed down the scope of prosodic planning compared to the other speakers because of their reduced cognitive resources. Note also that planning difficulties also resulted in a higher number of dysfluencies in the MS than in the C group. This suggests that the planning of the prosodic unit is flexibly adapted by speakers depending on their cognitive skills (cf. Swets et al. 2007).

In conclusion, the prosodic patterns found in our data are tentatively reanalyzed following Jun & Fougeron (2000)’s model of French intonation. While speakers might have not planned clash resolution, the comparison between healthy and MS speakers provide evidence of individual differences in the scope of the planning unit. A more detailed phonetic study is needed to quantify the amount of speaker-specific behavior in the realization of the prosodic structure and its link with cognitive capacities.

## 5. References

Avanzi, M., & Schwab, S. (2013). Stress clash resolution in French. A corpus-based account. Présenté à Phonetics & Phonology in Iberia, Lisbon.

Baayen, R., Davidson, D., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, p. 390-412.

Boersma, Paul & Weenink, David (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.70, retrieved 2 April 2014 from <http://www.praat.org/>

Brissart, M., Leiniger, M., Le Perf, L., Taillemite, E., Morele, M., & Debouverie, H. (2008). La mémoire de travail dans la sclérose en plaques: revue de la littérature. *Revue Neurologique*, p. 15-27.

Fuso, S. F., Callegaro, D., Pompéia, S., & Buono, O. F. A. (2010). Working memory impairment in multiple sclerosis relapsing-remitting patients with episodic memory deficits. *Arquivos de Neuropsiquiatria*, p. 205-211.

Ghio, A., Pouchoulin, G., Teston, B., Pinto, S., Fredouille, C., De Looze, C., ... Giovanni, A. (2012). How to manage sound, physiological and clinical data of 2500 dysphonic and dysarthric speakers? *Speech Communication*, p. 664-679.

Keating, P., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics*, p. 112-156.

Krivokapić, J. (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, p. 162-179.

Levelt, W. J. M. (1989). *Speaking. From intention to articulation* (MIT). Cambridge, Massachusetts.

McDonald, W., Compston, A., & Edan, G. et al. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol*, p. 121 - 127.

Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.

Petrone, C., Fuchs, S., & Krivokapić, J. (2011). Consequences of working memory differences and phrasal length on pause duration and fundamental frequency. *Proceedings of the ISSP*, p. 393-400. Montréal.

Post, B. (1999). Restructured phonological phrases in French. Evidence from clash resolution. *Linguistics*, p. 41-63.

Swets, B., Desmet, T., Hambrick, D., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, p. 64-81.

Tilsen, S. (2012). Utterance preparation and stress clash: planning prosodic alternations. In *Speech production and perception: Planning and dynamics* (ed. Fuchs, S., Perrier, P., Weirich, M. & Pape, D.) (p. 119-156). Frankfurt am Main: Peter Lang Verlag.

Welby, P. (2006). French intonational structure: Evidence from tonal alignment. *Journal of Phonetics* 34(3), p. 343-371.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. *Journal of Memory and Language*, p. 356-381.