Speech planning in question-response interactions from a clinical perspective

C. Petrone[1], M. Zira[1], C. Zielinski[2], E. Sneed German[3] and N. Moreau[4]
[1]*Aix-Marseille Université, CNRS, LPL UMR 7039, Aix en Provence, France*
[2]*Brain and Language Research Institute, AMU, Aix en Provence, France*
[3]*Université Paul Valéry Montpellier 3, EMMA, Montpellier, France*
[4]*Centre Hospitalier du Pays d'Aix, Service de Neurologie, Aix-en-Provence, France*

In question-response interactions, short latencies (~200 ms) require that listeners grasp relevant information from their interlocutors' questions as soon as it becomes available in order to plan their responses before the ongoing turn is finished [1]. However, strategies for speech production planning also vary with cognitive abilities, such as speed of processing [2].

This study investigates the effects of cognitive abilities in question-response interactions from a clinical perspective. We compared (Hexagonal) French healthy controls (HC) and individuals with cognitive disorders related to Multiple Sclerosis (MS) at an early stage of their disease. MS is an autoimmune disease characterized by the production of demyelinating lesions in the brain and spinal cord. Cognitive impairment affects 40-65% of MS patients and includes deficits on specific capacities involved in speech planning [3].

In a question-response game, we monitored eye movements to lexical competitors (**ca**nard/**ca**non, "duck/cannon") during question comprehension as well as latencies of speech responses. Participants orally replied to 24 trials consisting of a sequence of two pre-recorded questions (Q1 and Q2). Q1 asked whether the location of one object was above or below a geometric shape (e.g., *Est-ce que le canard est au dessus du rond?*, "Is the duck above the circle?"), and the following Q2 (e.g., *Et est-ce que le canard/canon est en dessous du carré?* "And is the duck/cannon below the square?") asked the location either of the previously mentioned object (anaphoric condition) or of a new object (non-anaphoric condition). After listening to Q2, participants provided a full response, whose form was kept constant (e.g., "No, the duck is above the square"). Target questions were interspersed among filler questions (which could also require a positive response). Questions were controlled for prosodic patterns for all trials. They were parsed into 3 Accentual Phrases (the basic prosodic unit in French). The object name could either carry a focal accent on the ambiguous syllable or not [4]. If intonation is used for reference resolution [5], the focal accent on the ambiguous syllable should facilitate a non-anaphoric interpretation, while the lack of the accent should induce higher fixations to the just-mentioned object. However, if the presence/absence of an accent cannot be reliably used to infer the contrastive status of a word (as for Canadian French [6]), listeners should just use segmental information to disambiguate lexical competitors. Given that planning a response is cognitively more demanding than understanding a question, differences across the two populations should more strongly emerge in response production. MS patients were predicted to have increased speech latencies and higher number of pauses in their responses.

35 patients of early stages of relapsing-remitting MS (mean disease duration = 4.5 year) and 35 HC (matched in age, sex, and education) participated in the experiment. (Among the inclusion criteria: no relapses at the time of the study; no history of optic nevritis; optimal vision; absence of dysarthria, assessed through tests from the BECD). Participants underwent standard neuropsychological tests, with the two populations differing in semantic and phonemic fluency, and speed of information processing (WAIS IV) ($p<0.05$). We found a bias for fixating "old" pictures, i.e., that were already referred to in Q1, in both HC and MS patients. For HC, cluster-based permutation analyses showed no effects of prosody on eye movements ($p> .05$). For MS patients, the likelihood of fixations to the already mentioned object demonstrated a bias to interpret deaccented words as anaphoric and accented words as non-anaphoric, for both anaphoric ($p < 0.001$; see Fig. 1) and non-anaphoric conditions

(p=0.004). An exploration of the onset contingent plots suggest that, in the non-anaphoric condition, MS patients were slower than HC in switching away from the already mentioned object when the word was deaccented. If confirmed, this would indicate that the use of intonational cues is strategic. HC might choose a less costly strategy by relying only on the stronger cue, i.e., segmental information, for lexical disambiguation. MS patients spread attention to both prosodic and segmental information but this is more source demanding. As for response planning, mixed models showed that, while the pause duration *before* "Non" was similar across the two populations, the pause duration *after* "No" was longer for MS (456 ms) than for HC (273 ms) (p<.01). The number of pauses was higher in MS patients (p=.02). Hence, MS patients used a within-turn pause to plan their response, possibly as a strategy to preserve turn-taking. Troubles in response planning, as evidenced by pauses, are in line with expectations. Our results support the hypothesis that speech production planning is flexible [7]. We further aim at investigating the correlations between neuropsychological scores and individual eye movements and verbal performances.

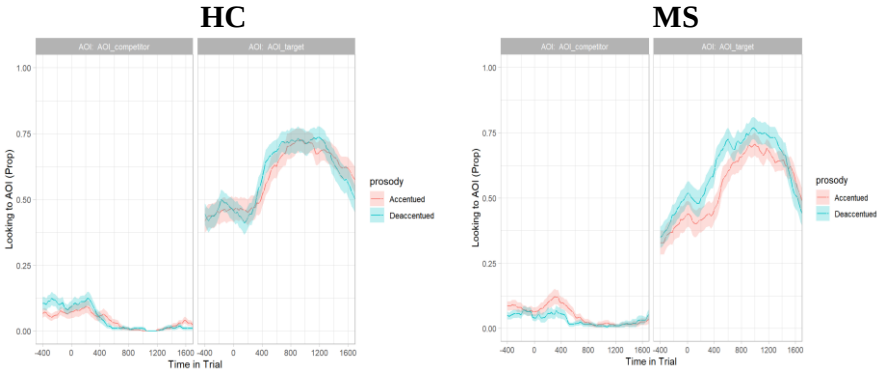

Figure 1. Fixation logits for Q2 in the anaphoric condition (canon -> canon) for HC (left) and MS patients (right). For each graph, data are separately plot for fixations to the old object, i.e., to the object already mentioned in Q1 (right) and for fixations to the new object (left). Trials with vs without focal accent are signaled by different colors (pink/blue). The zero value in the x axis is the beginning of the ambiguous syllable ("ca-").

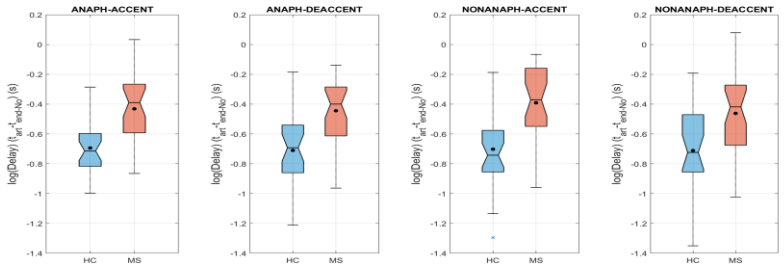

Figure 2. Latency of speech reponses (log) after the response "No" across discourse contexts and prosody in HC (blue) and MS patients (pink).

References. [1] Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Sci.Rep.*, 5: 12881. **[2]** Swets, B., Fuchs, S., Krivokapić, J. & Petrone, C. (subm.). A cross-linguistic study of individual differences in speech planning. **[3]** De Looze, C., Moreau, N.,... & Petrone, C. (2017). Effects of cognitive impairment on prosodic parameters of speech production planning in multiple sclerosis. *J. Neuropsychol*. **[4]** Jun S.-A. & Fougeron C. (2000). A Phonological model of French intonation. In A. Botinis (ed.) *Intonation: Analysis, Modeling and Technology*. Dordrecht : Kluwer AP. 209-242. **[5]** Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *J. Mem. Lang.*, 47(2), 292-314. **[6]** Swerts, M., & Vroomen, J.(2015). Accentshifts in spoken noun phrases affect verification latencies of listeners in Dutch but not Canadian French. *J. Phon.*, 52, 170-182. **[7]** Swets, B., Jacovina, M. E., & Gerrig, R. J. (2014). Individual differences in the scope of speech planning: Evidence from eye movements. *Lang.Cogn.*,6, 12–44.