Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

Speech markers of Cancer-Related Cognitive Impairment: A pilot study

Amélie B. Richard[1,2], Alexandre Foncelle[2], Fabrice Hirsch[1], Sophie Jacquin-Courtois[2,3], Karen T. Reilly[2], Manon Lelandais[4]

[1]Université Paul-Valéry Montpellier 3, UMR5267 Praxiling, France
[2]Université de Lyon, U1028; UMR5292 Lyon Neuroscience Research Centre, France
[3]Henry-Gabrielle Hospital, Hospices Civils de Lyon, France
[4]Université Paris Cité, UR3967 CLILLAC-ARP, France
amelie.richard@inserm.fr, manon.lelandais@u-paris.fr

Abstract

Speech is sensitive to mild cognitive changes due to age-related diseases, and prosodic features can identify patients with early-stage dementia from controls. Few studies have investigated speech markers of subtle cognitive impairment in non-neurodegenerative pathologies in younger populations, such as Cancer-Related Cognitive Impairment (CRCI). Little is known about the cognitive mechanisms underlying CRCI, but it is frequently encountered by cancer patients who mainly report memory-related concerns (i.e., forgetting words). Despite its substantial impact on patient quality of life, CRCI is difficult to detect with neuropsychological tools and often remains underdiagnosed. Our aim is to test whether previously documented speech markers are likely to detect CRCI in patients with breast cancer. We compared speech rate, F0 variability and pause duration in 11 breast cancer survivors with a cognitive complaint, 11 breast cancer survivors without any cognitive complaint and 10 controls in two narrative tasks (memory-based; picture-based). A Bayesian analysis showed no significant effects of group or task, but a qualitative analysis of pauses allowed us to generate hypotheses about the cognitive mechanisms underlying the patients' reported memory concerns. Even though speech markers specific to CRCI have yet to be defined, prosodic analysis is a promising approach for detecting subtle cognitive impairment.

Index Terms: Speech markers, cognitive impairment, pauses, breast cancer.

1. Introduction

Prosody is known to be sensitive to changes in cognition. Language is a complex cognitive function which relies on phonological and semantical mechanisms, as well as memory, executive and attentional processes. Speech production, as its output, reflects everyday cognitive ability.

Some prosodic features reach satisfactory discrimination power distinguishing patients with cognitive decline from controls, and identifying variants of a same pathology with either manual or machine learning methods [1].Speech rate, F0 variability and pause duration in particular have been shown to be good candidates for assessing cognition. For instance, individuals with amnestic Mild Cognitive Impairment have more F0 variability than individuals with non-amnestic Mild Cognitive Impairment [2]. Research in psycholinguistics suggests that these speech markers might be associated with different brain processes. For instance, a lower speech rate might reveal deficits in episodic memory, executive functions, and lexical retrieval difficulties in individuals with early-stage dementia [3] while a longer pause duration might reflect lexical-semantic difficulties [4]. These speech markers are even more manifest when the cognitive load is increased with a demanding task [5].

However, the large majority of studies using prosodic analysis to detect cognitive impairment focus on geriatric populations with dementia. Given that dementia is caused by clear degenerative physiological alterations of brain structures, its speech characteristics may differ from those of cognitive change not related to dementia. Little is known about speech markers in younger individuals with acquired cognitive impairment in non-neurodegenerative contexts. A new challenge in clinical prosody analysis is to test the efficiency of speech markers in differentiating between individuals with subtle cognitive impairment and healthy controls.

Subtle non-neurodegenerative cognitive impairment is increasingly frequent, and covers a wide range of pathologies (e.g., mild traumatic brain injury, long-Covid-19, minor stroke sequalae) in a population that is often young. It refers to self-reported cognitive difficulties that are hard to assess using current diagnostic tools. One of these impairments is Cancer-Related Cognitive Impairment (CRCI). CRCI is a functional change in cognition encountered by up to 70% of patients with cancer in the years following the end of curative treatments. In addition to decreasing patients' quality of life, CRCI hinders their return to work and limits their participation in social activities.

Cancer patients mostly report difficulties with memory and language (i.e., forgetting words and names, forgetting what they wanted to say) [6], [7]. Subjective difficulties in CRCI are confirmed by validated self-report questionnaires which signal the presence of a general cognitive complaint. Despite this, most studies note no or little association between the intensity of cognitive complaints and scores on neuropsychological tests [8]. This discrepancy might be explained by the fact that such tests were designed for more severe disorders and evaluate a single cognitive function in optimal conditions. The lack of appropriate objective tools hinders our understanding of CRCI, and prevents cancer patients from accessing appropriate care. There is thus a need for new approaches to explore CRCI and its underlying mechanisms.

Based on the results of previous studies using speech markers for detecting subtle cognitive impairment [9], we assume that prosodic analysis is sensitive and ecological enough to discriminate cancer patients with CRCI from controls.

10.21437/SpeechProsody.2024-16

1.1. Aim & hypotheses

Our aim is to test whether previously documented speech markers are likely to be specific to breast cancer patients with a cognitive complaint. We hypothesized that breast cancer survivors with a cognitive complaint would produce slower speech, with a longer pause duration and more F0 variability than those without a cognitive complaint and healthy controls. Based on the patients' memory complaints we think that they will perform worse in a memory-based narrative task than in a picture-based narrative task.

2. Method

The present study is part of two larger projects. The first one aims at assessing the effectiveness of a rehabilitation program offered to women with breast cancer who have finished the main phase of their treatment (ActiCog), and has been approved by the ethical committee of the Hospices Civils de Lyon (n°22-5009). The second project focuses on Cancer-Related Cognitive Impairment in women with breast cancer (DisCCo), and was approved by the Ethics Committee Est-III in August 2023.

2.1. Population

Thirty-seven survivors enrolled in a physical rehabilitation program, Alizés, had a 30-minute one-to-one interview with a speech therapist. The Alizés program is offered by the Physical Medicine and Rehabilitation department at the Henry Gabrielle hospital (Hospices Civils de Lyon), and is open to women who are less than one year after completion of an adjuvant treatment for breast cancer (surgery, chemotherapy and radiation therapy). Women taking hormone therapy were included.

The women interviewed for the current study were required to be enrolled in the Alizés program and to have a good level of French (i.e., native speaker or fluent). Exclusion criteria included: neurological (i.e., stroke) or psychiatric (i.e., schizophrenia) history, currently taking medication known to modify cognitive ability (e.g., antidepressants), and the presence of a speech or language disorder. Four survivors were not eligible because of epilepsy, depression, or stroke history, and one was not fluent in French. The FACT-Cog questionnaire is a validated measure of cognitive complaint in patients with cancer. A score below 55/72 on the Perceived Cognitive Impairment subscale signals the presence of a cognitive complaint [10]. Among the thirty-three eligible participants, we randomly chose eleven who had a significant complaint revealed by the FACT-Cog, and selected eleven others with scores above this cutoff. In total, twenty-two participants were included in the current study. In addition, we selected ten healthy control participants from the DisCCo study without any cancer history or cognitive complaint. Inclusion and exclusion criteria for controls were the same as described above.

To sum up, participants were divided into three groups: 11 cancer survivors with a cognitive complaint (BCcog), 11 cancer survivors without any cognitive complaint (BC), and 10 healthy controls (HC).

2.2. Procedure

2.2.1. Tasks and speech variables

Participants were asked to complete two narrative tasks: a **picture-based** task for which they were instructed to tell a story from a 5-picture sequence [11], and a **memory-based task** for which they were instructed to listen to and memorize a 1-minute short story [12], then to immediately recall it with a maximum of detail. Healthy controls were instructed to read and memorize a short text displayed on a screen for forty seconds, then to immediately recall it with a maximum of detail.

The samples for survivors were recorded with a Rode Lavalier Go in a quiet room at the hospital (Saint-Genis-Laval, France). The samples for healthy controls were recorded with a Shure lavalier in a soundproof cabin at the Lyon Neuroscience Research Centre (Bron, France). Three speech variables were selected according to the following characteristics:

- **Pause duration (ratio):** was measured by dividing total pause time by total speech time. Pauses included silent pauses (speechless segments with a duration above 250 ms), filled pauses (non-lexical fillers perceived as hesitation markers such as "euh" and "hum" (/ʌ/, /ø/ and /m/) with a flat F0 contour and a duration above 250 ms), and lengthenings (voiced phonemes with flat F0 contour and a duration above 250 ms). The literature on pause length agrees on a threshold equal to or above 200 ms for pauses related to cognitive processes such as hesitation or demarcation [13]. However, we recorded survivors' samples in a hospital resulting in medium audio quality. We thus chose a threshold of 250 ms to avoid annotating unrelated acoustic phenomena such as echoes or background noise.
- **Speech rate (wpm):** was measured by dividing the number of words per minute. This included lexical and grammatical items and word fragments, and excluded filled pauses and noises.
- **F0 variability (Hz):** was automatically measured with Praat [14] as the standard deviation of the total speech sample.

2.2.2. Subjective evaluation

In addition to the PCI subscale of the FACT-Cog questionnaire, participants completed the Perceived Cognitive Ability subscale. Despite having no standardized cutoff, the subscale is used in clinical contexts as a useful tool to measure self-reported cognitive ability. Participants also filled the standardized Hospital Anxiety and Depression (HAD) scale for anxiety and depression.

2.3. Data processing

We recorded 45 minutes of picture-based task and 31 minutes of memory-based task. All samples were automatically transcribed with Whisper [15] and manually checked by the first author. Pause annotation was performed with the SPPAS software [16] and errors were manually corrected by the first author in Praat. In total, we annotated 874 pauses for the picture task and 1355 pauses for the memory task.

2.4. Statistical analysis

2.4.1. Linear mixed-effects model

The null hypothesis (H0) postulates that there is no difference in F0 variability, pause ratio or speech rate, depending on groups and conditions (i.e. tasks). To test H0 against H1 (i.e., the presence of a significant difference), linear mixed models were designed and compared for each of these three dependent variables. A random intercept was attributed to the *subject* variable, as well as fixed effects to *group* (HC, BC and BCcog

levels), *condition* (picture and memory levels), *age*, *anxiety*, *depression*, and the various interaction terms were combined with all these fixed effects. Models were conducted in R [14] using the lme4 package [17].

2.4.2. Bayesian statistics

To confirm the results of the mixed models, we conducted a Bayesian analysis. There was no strong *a priori* on the parameter values of the models, so non-informative priors were set for all parameters in all models with a Normal prior distribution N(0,3). The Bayes factor was calculated between each model comparison for hypothesis testing and, for the most likely model, the summary of the parameter estimate with a 95% credibility interval was considered. The Bayesian analysis was conducted in R using the brms package [18].

3. Results

3.1. Population description

Table 1 shows the population's mean age and scores in questionnaires on cognitive complaint and anxiety/depression, along with standard deviations. Controls were slightly younger than cancer outpatients, although a Kruskal-Wallis test showed no significant difference between groups (p = .31). A lower score on the PCI subscale signals a higher cognitive complaint. As mentioned in the Method section, only BCcog had a significant cognitive complaint. A lower score on the PCA subscale signals lower perceived cognitive abilities. Cognitive abilities were significantly lower in the BCcog group. A score above 11/21 on the HAD subscales signals the presence of anxiety or depression. None of the participants reported anxiety or depression.

Table 1: *Population age and scores in questionnaires on cognitive complaint, and anxiety/depression; mean (sd)*

Characteristics	**BC**	**BCcog**	**HC**
Age	47.18 (7.15)	49.36 (9.11)	44.10 (9.96)
FACT-Cog PCI score /72	62.73 (5.10)	40.55 (9.32)	66.70 (5.12)
FACT-Cog PCA score /28	20.27 (6.02)	13.27 (4.55)	23.70 (2.87)
HAD – Anxiety score /21	6.36 (2.99)	5.11 (2.77)	8.20 (3.34)
HAD – Depression score /21	4.18 (3.59)	4.44 (3.74)	6.30 (3.38)

3.2. Prosodic variables

Table 2 shows median pause duration ratio, speech rate and F0 variability for each group in the memory and picture conditions. Pause duration ratio is higher for the BCcog group in the picture condition (0.42) than for BC (0.23) and controls (0.28).

Table 2: *Median pause duration ratio, speech rate and F0 variability for the memory and picture conditions*

	BC	**BCcog**	**HC**
Memory			
Pause duration ratio	0.30	0.31	0.23
Speech rate (wpm)	165	173	150
F0 variability (sd)	35.66	38.26	31.91
Picture			
Pause duration ratio	0.23	**0.42**	0.28
Speech rate (wpm)	180	152	162
F0 variability (sd)	34.81	37.72	33.45

Figure 1 shows the violin box distributions of pause duration (a), speech rate (b), and F0 variability (c) between groups for the memory (blue) and picture (orange) conditions. The white line represents the median. A trend can be observed in 1a with a higher pause duration ratio in BCcog than in controls, especially for the picture condition (see Table 2).

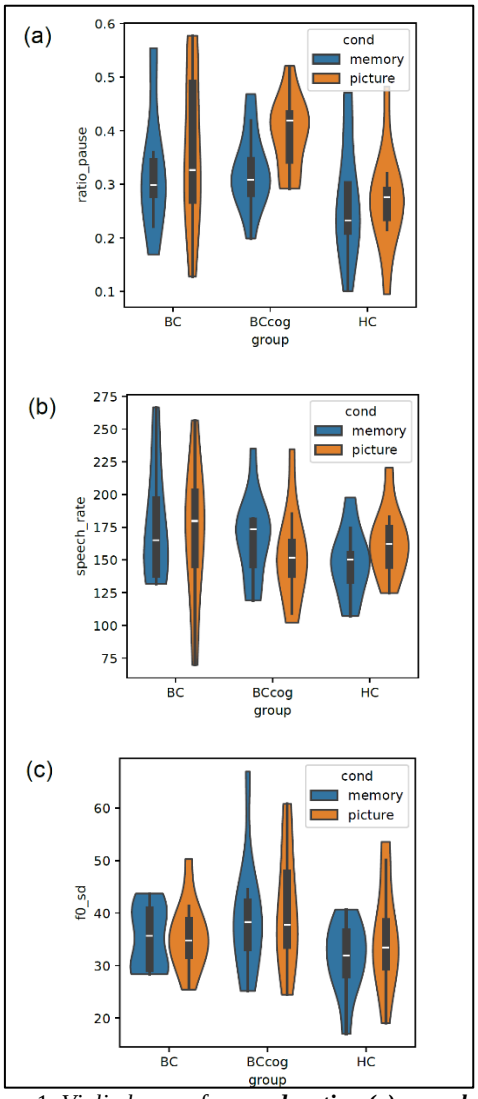

Figure 1: *Violin boxes of **pause duration (a), speech rate (b)**, and **F0 variability (c)** for each group in the memory and picture conditions.*

3.3. Linear and Bayesian analyses

The linear mixed models showed no effect of group, condition, age, anxiety or depression on the three speech markers. The Bayes factor indicates evidence for the null hypothesis (i.e., no difference between conditions and groups), $BF_{H0/H1\ pause\ ratio}$ = $1.63e^7$, $BF_{H0/H1\ speech\ rate}$ = 1.11, $BF_{H0/H1\ F0sd}$ = 3.41, meaning that the data are respectively approximately $1.63e^7$, 1.11, and 3.41 times more likely to occur under H0 than under H1. These results indicate very strong evidence in favor of the null

hypothesis for pause ratio, no evidence for speech rate, and moderate evidence for F0 variability. In other words, the Bayesian analysis suggests that there is no difference between condition and groups for pause ratio, that no conclusion can be made for speech rate, and that there might be no difference for F0.

3.4. Qualitative analysis of pause duration

In order to interpret the trend concerning pause duration ratio alongside the absence of quantitative results, we carried out a finer-grained analysis of pauses. We specifically examined pause duration between two functional groups of pauses, e.g. demarcation and hesitation.

In French, rhythmic groups are marked by final accents [19] with a clear rising or falling F0, coinciding with syntactic boundaries. Pauses that occur after these prosodic boundary cues have a demarcative function [19] and were labeled as such, as shown in Figure 4. Pauses that follow a flat F0 and break syntactic linearity signal hesitation and were labeled as such, as shown in Figure 5. Filled pauses and lengthenings following a silent pause and beginning a clause were also considered hesitation marks. We measured the median duration of both demarcation and hesitation pauses.

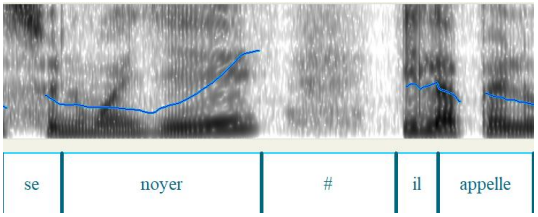

Figure 2: *Silent pause with a demarcative function between two prosodic and syntactic domains ("he's starting to drown # he's calling for help").*

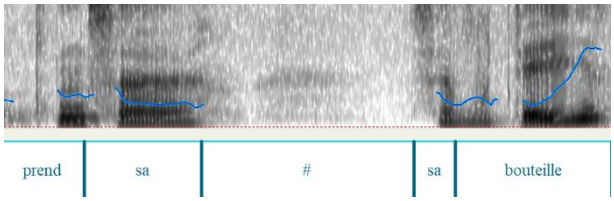

Figure 3: *Silent pause with a hesitation function within a syntactic unit ("he's taking his # his bottle").*

Table 3 shows median pause duration for each pause function between groups in the memory and picture conditions. Demarcation pauses were overall longer than hesitation pauses, except hesitation pauses produced by the BCcog group in the picture condition (1.120 ms).

Table 3: *Median pause duration by function*

Function	**BC**	**BCcog**	**HC**
Memory			
Demarcative pauses	1.065	1.280	0.865
Hesitation pauses	0.460	0.510	0.390
Picture			
Demarcative pauses	0.845	0.851	0.570
Hesitation pauses	0.458	**1.120**	0.410

4. Discussion

The statistical analyses yielded no conclusive results for pause duration, speech rate, and F0 variability. However, a finer-grained investigation of pause functions showed differences between groups. Cancer survivors with a cognitive complaint produced longer hesitation pauses than those without any cognitive complaint and controls. This finding suggests that targeting only hesitation pauses might be relevant for identifying subtle cognitive impairment with a prosodic analysis.

Furthermore, in relation with task type, the descriptive results showed that survivors with a cognitive complaint produced longer pauses in the picture task compared to the memory task. This contradicts our hypothesis stating that survivors would produce longer pauses in the memory-based narrative task due to their reported memory concerns. However, this result is in line with previous research on pauses in individuals with early Alzheimer's Disease, which showed that patients paused more during a picture narrative than in a memory-based narrative [20]. According to the authors, memory-based narratives might be more ecological than picture-based narratives, and might provide an easier context for participants to complete the task. This raises the question of cooccurring cognitive deficits underlying CRCI for cancer survivors. Indeed, a longer hesitation pause duration might reflect the increase of cognitive load due to the nature of the task, which requires visual exploration as well as lexical retrieval and speech planification. Although memory concerns are frequently reported among this population, they might overshadow other deficits than memory such as attentional and executive processes.

Finally, our results may be limited by the small number of participants. To estimate the number of cancer participants with a cognitive complaint and healthy controls required to reach significance for pause duration ratio (with a level of α below 0.05 and a power level above 80%), we performed a power analysis. Twenty participants per group are needed to show a significant difference in the picture condition between BCcog and HC. Otherwise, thirty participants per group would be needed to conclude that there is no difference in memory and picture conditions between BCcog and HC.

5. Conclusions

We conducted a speech analysis using three prosodic markers (pause duration ratio, speech rate and F0 variability) in two narrative tasks (memory-based and picture-based) in an attempt to discriminate cancer survivors with a cognitive complaint from those without any complaint and healthy controls. Although the mixed-effects model and Bayesian analyses showed no significant results, survivors with a cognitive complaint had a higher pause duration ratio than the other groups. A finer-grained analysis of pause function revealed that survivors with a cognitive complaint paused longer especially in the picture-based narrative task. This suggests that identifying hesitation pauses might help detect subtle cognitive impairment in cancer survivors. Furthermore, a longer pause duration in the picture condition might reflect an increase of cognitive load due to underlying deficits such as visual attention and executive processes. Further investigation will thus be needed to confirm that hesitation pauses in picture condition is a good prosodic marker of Cancer-Related Cognitive Impairment.

6. References

[1] O. Ivanova, I. Martínez-Nicolás, and J. J. G. Meilán, 'Speech changes in old age: Methodological considerations for speech-based discrimination of healthy ageing and Alzheimer's disease', *Intl J Lang & Comm Disor*, pp. 1460-6984.12888, May 2023, doi: 10.1111/1460-6984.12888.

[2] J. A. Mefford *et al.*, 'Varied performance of picture description task as a screening tool across MCI subtypes', *PLOS Digit Health*, vol. 2, no. 3, p. e0000197, Mar. 2023, doi: 10.1371/journal.pdig.0000197.

[3] C. De Looze *et al.*, 'Cognitive and Structural Correlates of Conversational Speech Timing in Mild Cognitive Impairment and Mild-to-Moderate Alzheimer's Disease: Relevance for Early Detection Approaches', *Front. Aging Neurosci.*, vol. 13, p. 637404, Apr. 2021, doi: 10.3389/fnagi.2021.637404.

[4] J. E. Mack *et al.*, 'What do pauses in narrative production reveal about the nature of word retrieval deficits in PPA?', *Neuropsychologia*, vol. 77, pp. 211–222, Oct. 2015, doi: 10.1016/j.neuropsychologia.2015.08.019.

[5] M. Belz, 'Defining Filler Particles: A Phonetic Account of the Terminology, Form, and Grammatical Classification of "Filled Pauses"', *Languages*, vol. 8, no. 1, p. 57, Feb. 2023, doi: 10.3390/languages8010057.

[6] G. Bolton and A. Isaacs, 'Women's experiences of cancer-related cognitive impairment, its impact on daily life and care received for it following treatment for breast cancer', *Psychology, Health & Medicine*, vol. 23, no. 10, pp. 1261–1274, 2018, doi: https://doi.org/10.1080/13548506.2018.1500023.

[7] F. M. Henderson, A. J. Cross, and A. R. Baraniak, '"A new normal with chemobrain": Experiences of the impact of chemotherapy-related cognitive deficits in long-term breast cancer survivors', *Health Psychology Open*, vol. 6, no. 1, p. 205510291983223, Jan. 2019, doi: 10.1177/2055102919832234.

[8] V. J. Bray, H. M. Dhillon, and J. L. Vardy, 'Systematic review of self-reported cognitive function in cancer patients following chemotherapy treatment', *J Cancer Surviv*, vol. 12, no. 4, pp. 537–559, Aug. 2018, doi: 10.1007/s11764-018-0692-x.

[9] I. Martínez-Nicolás, T. E. Llorente, F. Martínez-Sánchez, and J. J. G. Meilán, 'Ten Years of Research on Automatic Voice and Speech Analysis of People With Alzheimer's Disease and Mild Cognitive Impairment: A Systematic Review Article', *Front. Psychol.*, vol. 12, p. 620251, Mar. 2021, doi: 10.3389/fpsyg.2021.620251.

[10] J. E. Fardell, V. Bray, M. L. Bell, B. Rabe, H. Dhillon, and J. L. Vardy, 'Screening for cognitive symptoms among cancer patients during chemotherapy: Sensitivity and specificity of a single item self-report cognitive change score', *Psycho-Oncology*, vol. 31, no. 8, pp. 1294–1301, Aug. 2022, doi: 10.1002/pon.5928.

[11] C. Bézy, A. Renard, and J. Pariente, *GREMOTS Batterie d'évaluation des troubles du langage dans les maladies neurodégénératives*. Louvain-la-Neuve, Belgique: De Boeck Supérieur, 2016.

[12] Y. Joanette, B. Ska, and H. Côté, *Protocole Montréal d'Evaluation de la Communication*. Isbergues, France : Ortho Edition, 2004.

[13] E. Campione and J. Véronis, 'A Large-Scale Multilingual Study of Silent Pause Duration', presented at the SPeech Prosody, Aix-En-Provence, 2002.

[14] P. Boersma and D. Weenink, 'Praat: Doing Phonetics by Computer'. 2024. Accessed: Oct. 26, 2021. [Online]. Available: www.praat.org

[15] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, 'Robust Speech Recognition via Large-Scale Weak Supervision', 2022, doi: 10.48550/ARXIV.2212.04356.

[16] B. Bigi, 'SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech', *the Phonetician" - International Society of Phonetic Science*, vol. 111–112, no. 2015-1–11, pp. 5–69, 2015.

[17] D. Bates, M. Mächler, B. Bolker, and S. Walker, 'Fitting Linear Mixed-Effects Models Using **lme4**', *J. Stat. Soft.*, vol. 67, no. 1, 2015, doi: 10.18637/jss.v067.i01.

[18] P.-C. Bürkner, 'Bayesian Item Response Modeling in *R* with **brms** and *Stan*', *J. Stat. Soft.*, vol. 100, no. 5, 2021, doi: 10.18637/jss.v100.i05.

[19] C. Astésano, 'Prosodic characteristics of Reference French', in *Varieties of Spoken French*, 1st ed., S. Detey, J. Durand, B. Laks, and C. Lyche, Eds., Oxford University PressOxford, 2016, pp. 68–85. doi: 10.1093/acprof:oso/9780199573714.003.0006.

[20] A. Pistono, J. Pariente, C. Bézy, B. Lemesle, J. Le Men, and M. Jucla, 'What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease', *Neuropsychologia*, vol. 124, pp. 133–143, Feb. 2019, doi: 10.1016/j.neuropsychologia.2018.12.018.