Videos to study Interactions in AGEing (VIntAGE): a longitudinal, ecological and multimodal French corpus

Guillaume Duboisdindien[1][*] & Catherine Bolly[2]

[1]Université Laval, Québec, Canada & Université Lille, France
[2]Université Catholique, Louvain-la-Neuve, Belgium

Abstract. The *Videos to study Interactions in AGEing* (VIntAGE) corpus aims to investigate the complex relationship between language, cognition, and aging, focusing on verbal and non-verbal pragmatic markers in older persons with mild cognitive impairment (MCI). This multimodal and longitudinal corpus incorporates an analysis of gestural and verbal markers in discourse, aligned with neurolinguistic models. It provides a rich dataset for analyzing how aging impacts communicative competence in individuals with MCI. The VIntAGE corpus comprises approximately 18 hours of video recordings from 36 face-to-face interviews conducted by a close acquaintance of each of the nine women, all over 75 years old. Five participants were selected for in-depth analysis due to significant changes in their cognitive status. The participants underwent a series of semi-structured interviews over 15 months. The data were processed using transcription tools (for verbal discourse) and annotation tools (for gestures) and then subjected to Principal Component Analyses to manage each individual's diverse dataset and discursive modalities. The corpus includes the annotation of 6,351 verbal pragmatic markers (VPMs) and 8,044 non-verbal pragmatic markers (NVPMs). The data reveal an average decrease in MoCA scores from 23 / 30 to 20 / 30 over one year, highlighting cognitive decline's effects on verbal and non-verbal communication.

Keywords. multimodal communication, pragmatics, aging, Mild Cognitive Impairment, psycholinguistics

© Guillaume Duboisdindien & Catherine Bolly
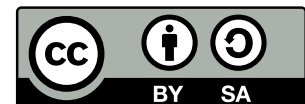
*. Corresponding author, duboisdindien@hotmail.com

OPEN ACCESS

1 Introduction

The aging population, projected to increase significantly by 2050, presents critical challenges due to associated neuroanatomical, metabolic, and cognitive declines. However, these highly heterogeneous changes complicate the distinction between normal and pathological cognitive aging (Delano-Wood et al. 2009). Mild Cognitive Impairment (MCI) exemplifies this complexity, presenting an altered cognitive profile that is not severe enough to qualify as dementia (Petersen et al. 1999; Winblad et al. 2004). Given the variability in cognitive trajectories among older adults with MCI, it is essential to rethink aging through a developmental lens, incorporating innovative humanities and social sciences methodologies (Cummings 2017; Davis and Guendouzi 2014; Mueller et al. 2018). Numerous studies have focused on characterizing the linguistic phenomena observed in the discourse of older individuals and those suffering from Major or Mild Neuro-Cognitive Disorders (MNCD-MNCD / MCI) as senescence indicators. Discourse markers, and more specifically verbal pragmatic markers of affiliation and alignment (such as *mm mm*, *yes*, *of course*, *indeed*), are believed to be widely used in the speech of these populations (Bolly and Boutet 2018; Davis 2005; Hamilton 1994, 2008; Wray 2008). These pragmatic markers are important as they provide significant insights into the interactional resources of these individuals (Hamilton 2005; Pistono et al. 2019). Various research fields have analyzed and defined these occurrences to better understand their function and form in oral discourse (for a comprehensive review, see Bolly and Crible 2015; Crible 2018). However, no studies have explored the interaction between verbal and gestural pragmatic markers, nor have they adopted a functional approach to discourse in light of current neurolinguistic models. Such a perspective could enhance the understanding of the discourse of individuals at risk of developing MNCD, for whom existing standardized tests may lack sensitivity. Discourse analysis through a multimodal corpus warrants further exploration, as it could also provide clinicians, such as speech and language therapists and neuropsychologists, with methodological tools to develop descriptive assessments, or even first-order static criterion-based assessments, in the context of aging.

The French corpus *Videos to study Interaction in AGEing* (VIntAGE 2024) is an innovative and significant experimental resource developed to study the complex relationships between language, cognition, and aging. This multimodal and longitudinal corpus, consisting of approximately 1,080 minutes of recordings from 36 face-to-face interviews with nine female participants aged over 75 (duration: approximately 30 minutes on average), is designed to explore verbal and non-verbal communication, particularly in individuals with MCI. Five participants were selected for in-depth analysis due to significant changes in their MCI status during the recording period (average age 83 years; average MoCA score from 23 / 30 in November 2014 to 18 / 30 in November 2015). This subcorpus of 2 hours and 30 minutes includes the annotation of 6,351 verbal pragmatic markers (VPMs) and 8,044 non-verbal pragmatic markers (NVPMs), providing a comprehensive dataset for a detailed analysis of how aging impacts communicative competence. The VIntAGE corpus is grounded in contemporary neurolinguistic research, specifically focusing on VPMs and NVPMs, which are crucial for maintaining communication despite cognitive decline. By enabling an in-depth investigation of these markers, VIntAGE has the potential to significantly advance our understanding of how older individuals adapt their communication strategies as their

cognitive abilities change. This focus is particularly relevant to psycholinguistics and neurolinguistics, where understanding the compensatory mechanisms employed by individuals with MCI is critical for theoretical and clinical research.

Methodologically, VIntAGE aligns closely with the project CorpAGEst (2021) led by Catherine Bolly (see Bolly and Boutet 2018), which has set high standards for collecting and analyzing ecological and multimodal data. By emphasizing the study of communication in natural settings, VIntAGE ensures that its findings are robust and reflective of real-world interactions.

2 Corpus description

2.1 Participants

The inclusion criteria for the study were based on those established by Petersen et al. (1999) and Winblad et al. (2004). The participants in this study were women. Given its exploratory dimension, we wanted to start with a homogeneous group. To be recruited, participants had to present memory complaints, confirmed by a neurologist and an informant, and objective memory deficits appropriate for their age. Their overall health had to be relatively preserved for their age, with generally intact activities of daily living (level 5 or 6 on the French Autonomie Gérontologique Groupe Iso Ressources (AGGIR) grid, see Service-Public.fr 2024) and the absence of progressive Major NeuroCognitive Disorder (MNCD, score between 22 and 26 on the Montréal Cognitive Assessment (MoCA) test; Nasreddine et al. 2005). Additionally, participants had to meet the following additional criteria: they had to be French-speaking women over 75 years old, without a history of concurrent neurological or psychiatric diseases, and using symptomatic medication for early-stage MNCD. The diagnosis of MCI had to be validated by a physician and regularly monitored by a neurologist or a local health and aging center. Moreover, each recruited person had to be accompanied by a friend or a family member, with whom the interviews were conducted. The choice of this intimate was guided by the principles of social support theory, which postulates a positive correlation between the existence of social relationships with friends or family members and physical and mental health Leahy-Warren 2014. All participants (n = 5) underwent the following tests: an assessment of global cognition (MoCA test; Nasreddine et al. 2005), a French self-assessment of communication in daily life (Echelle de Communication Verbale de Bordeaux (ECVB); Darrigrand and Mazaux 2000), and an assessment of empathy through the French Interpersonal Reactivity Index test (F-IRI; Gilet et al. 2013). Detailed demographic information are presented in the results section in Table 1 (see also Duboisdindien and Bolly 2025).

2.2 Experimental setup and collection of multimodal corpora

Four semi-structured interviews, each lasting an average of 30 minutes, were conducted in the participants' homes over 15 months by an intimate designated by them (for the description of the corpus design see the corpus publication VIntAGE 2024). These recorded interviews covered various autobiographical themes about their past and present life related to aging and life stages. The scientific coordinator discreetly waited for the end of the interview in a nearby room to avoid influencing the exchange.

2.2.1   Structure of the filmed interviews

Tasks: The proposed tasks sequentially alternate between discussions about the past and the present. The past-focused tasks include age-related milestones (in the first interview) and reminiscence tasks for subsequent interviews (approximately 15–20 minutes each). The present-focused tasks involve thematically oriented conversations to encourage spontaneity (approximately 15–20 minutes each). All interviews follow this consistent structure to ensure coherence and comparability across the following levels: (i) between the cross-sectional and the longitudinal subcorpora of CorpAGEst, (ii) within the longitudinal subcorpora of CorpAGEst, and (iii) across linguistic varieties, the French in Belgium (FB) and French in France (FF) subcorpora (for details, see VIntAGE 2024). Reminiscence materials were prepared for past narrative tasks in agreement with the intimate to select personal photographs, odors, and sounds related to the participant's tastes and past.

- Initial Contact Interview: These interviews introduce the participant and include administering the MoCA cognitive test and providing the IRI empathy test.

- Interview 1: For the sake of comparability, the first interview conducted in France aligns with the initial interviews conducted in Belgium (as per the first cross-sectional phase of the CorpAGEst corpus), albeit in a shortened version. This interview follows the established protocol for these sessions.

- Interview 2: This interview consists of a reminiscence task based on a personal / family photograph (chosen in consultation with the family if possible), followed by a relatively open conversation on a theme (with follow-up questions and suggested topics).

- Interview 3: This session involves a reminiscence task triggered by a scent (per the list of materials to be prepared), followed by a free conversation task (with follow-up questions and suggested topics). The interviewee can choose from several scents: "Which one evokes something you would like to share?"

- Interview 4: This interview consists of a kinesthetic reminiscence task (blindfolded: objects and materials will be placed in a bag), followed by a free conversation task (with follow-up questions and suggested topics).

Interview 1	Interview 2	Interview 3	Interview 4
Task A Focus on the past	Age milestone	Reminiscence on a personal photograph	Olfactory reminiscence
Task B Focus on the present	Social perceptions of aging	Living environment and daily life	Social and family relationships

Table 1: Interview content

2.2.2   Materials

The audio recordings were made using two digital recorders (Roland R-09HR Ver. 3.0; codec LHDC); two Lavalier microphones (audio-technica ATR3350); four SD / SDHC cards of at least 4 GB each, to have sufficient space if two interviews were to follow

each other; and one audio headset for remote monitoring of the interview and subsequent orthographic transcription and alignment. The recordings were made in .wav format, mono, frequency of 22,000 Hz (16 bits). Video recording was carried out with two non-professional HD digital camcorders (codec H.264), easily transportable and compact to minimize intrusion into the participants' homes (Sony Handycam HDR-CX250E); four SD / SDHC cards of 16 or 32 GB each, to have sufficient space if two interviews were to follow each other; two lightweight semi-professional tripods for ease of transport (Bilora PRO 25); two power strips, one of which was reserved in case there were not enough power outlets or if the outlets were too far from the cameras. In our case, the video recording format complied with international standards adopted by the French-speaking corpus repository center Ortolang (Pierrel et al. 2017).

2.2.3 Arrangement of equipment in the space

The recording setup consisted of two cameras mounted on tripods to ensure they remained stationary (Figure 1).

Figure 1: Arrangement of Cameras 1 and 2 during the VIntAGE recording. (a) Camera 1: medium shot, focused solely on the older participant (here, Tristane); (b) Camera 2: wide shot, presenting the interaction scene between the close companion and the participant, full body view

Camera 1 was positioned facing the participant, capturing a medium shot that included the head, torso, and upper thighs. This setup allowed for a close enough view to capture facial details (facial expressions and eye movements) while covering the range of manual gestures. The participant was consistently instructed to perform various movements in front of the camera to ensure that no movement would go unrecorded.

Camera 2 was oriented to capture a comprehensive view of the interaction, focusing on both participants (the participant and the close companion) face-to-face. The camera was angled slightly to ensure their faces and limbs were visible, capturing them from head to toe. This setup was specifically designed to facilitate the annotation of movements and / or foot positioning.

2.3 Data processing

We developed a systematic chain of multimodal data processing for our audio and video samples, both pre- and post-analysis, in line with our research objectives and for corpus representativeness (Arbach and Ali 2013; Mondada 2007) (Figure 2). In this study, we adopted the recommendations of Kennedy (2014) for corpus linguistic analysis, emphasising the balance between ecological approaches (i.e., non-intrusive and spontaneous), technical constraints, and the comparability of the tools used. The goal was to ensure (i) adequate representativeness of the studied population, (ii) comparability between sub-corpora and proposed tasks, and (iii) systematization of the analysis for future projects.

Multimodal data processing included four main phases. First was the stage of processing the audio and video recordings of the interviews, which involved sampling and the calibration of verbal and non-verbal data (Arbach and Ali 2013; Sinclair 2004). Another stage was dedicated to the processing of gestural data. There was the stage of processing verbal data, including orthographic transcription of data by using Praat software (Boersma and Van Heuven 2001) and their formatting, as well as the annotation of VPMs (Bolly and Crible 2015; Crible 2018) involving their formatting in Elan software and the annotation of NVPMs (Allwood 2008; Allwood et al. 2007; Bolly and Boutet 2018).

2.3.1 Sampling of audio and video data

The VIntAGE corpus was compiled over 15 months and required regular sampling phases. In addition to the internal criteria previously described (including gender, age, cognitive profile, and the systematic conditions and locations of recordings), external criteria were established to address specific constraints. Two main dimensions are considered in the standardization of analysis procedures and their application to other types of studies using these corpora. To characterize these sampling criteria in advance, we examine:

1. A horizontal axis, which concerns representations of language. This includes the representativeness of the three modalities (written, audio, and audio-visual), the representativeness of discourse types, the representativeness of sociolinguistic variation, and the representativeness of languages within the corpus (including atypical and pathological productions in our case). These criteria result in categories that can be further refined into subcategories according to the specificities of VIntAGE and based on the experience gained from the CorpAGEst project (Bolly and Boutet 2018).

2. A vertical axis, which relates to the representativeness derived from enough occurrences and word types, i.e., the corpus size. For oral corpora, it is crucial to account for the idiosyncratic behaviors of speakers and to reflect on how to minimize their effects—however relevant they may be—to support our hypotheses

Participant selection and inclusion

- ✓ Women – Age ≥ 75 Y. old – no diagnosis of AD
- ✓ French-speaking
- ✓ No concomitant history of neurological or psychatric disease / absence of symptomatic drug treatments
- ✓ Diagnosis of MCI validated by a clinical exam
- ✓ MoCA-score between : 22 – 26 + goof level of autonomy (5-6 AGGIR)

Experimental design

- ⮌ A participant's relative conducted semi-directed longitudinal interviews (15 months)
- ⮌ 4 filmed interviews (2 cameras) l 30 minutes l every quarter
- ⮌ 15 minutes of interviews based on past autobiographical data (multisensory reminiscences, autobiographical questions) and 15 minutes on current identity data (themes : technology, living place to old age, age benefits)
- ⮌ At the speaker's home, in a quiet environment
- ⮌ Relative training beforehand to conduct the interviews
- ⮌ Supervised by ethics committee + informed consent (written) + verbal agreement reiterated on camera

Sampling

- A. Selection and sampling (10 minutes/sample)
- B. Data cleaning/formatting/synchronisation of the media

Verbal data processing

- ✓ Full orthographic transcription
- ✓ Inter-annotator judgment phase
- ✓ Clean orthographic transcription
- ✓ Alignement to text <> sound (words – syllables – phonemes)

Hand gestures processing

- ✓ Identification and segmentation of strokes gestures
- ✓ Inter-annotator judgment phase
- ✓ Cleaning and adjustments

Verbal pragmatic markers processing

- ✓ Identification of VPMs
- ✓ Cleaning and inter-rater agreement
- ✓ Inter-annotator judgment phase
- ✓ Cleaning and adjustments

Non verbal pragmatic markers processing

- ✓ Functional labeling ot strokes
- ✓ Labeling of gestures symmetries (counting)
- ✓ Inter-annotator judgment phase
- ✓ Cleaning and adjustments

Statistical processing of multimodal data

- ⮌ Data pre-processing (optimization, agrégation and completeness)
- ⮌ Univariate processing (distriution and counting)
- ⮌ Unimodal multivariate processing
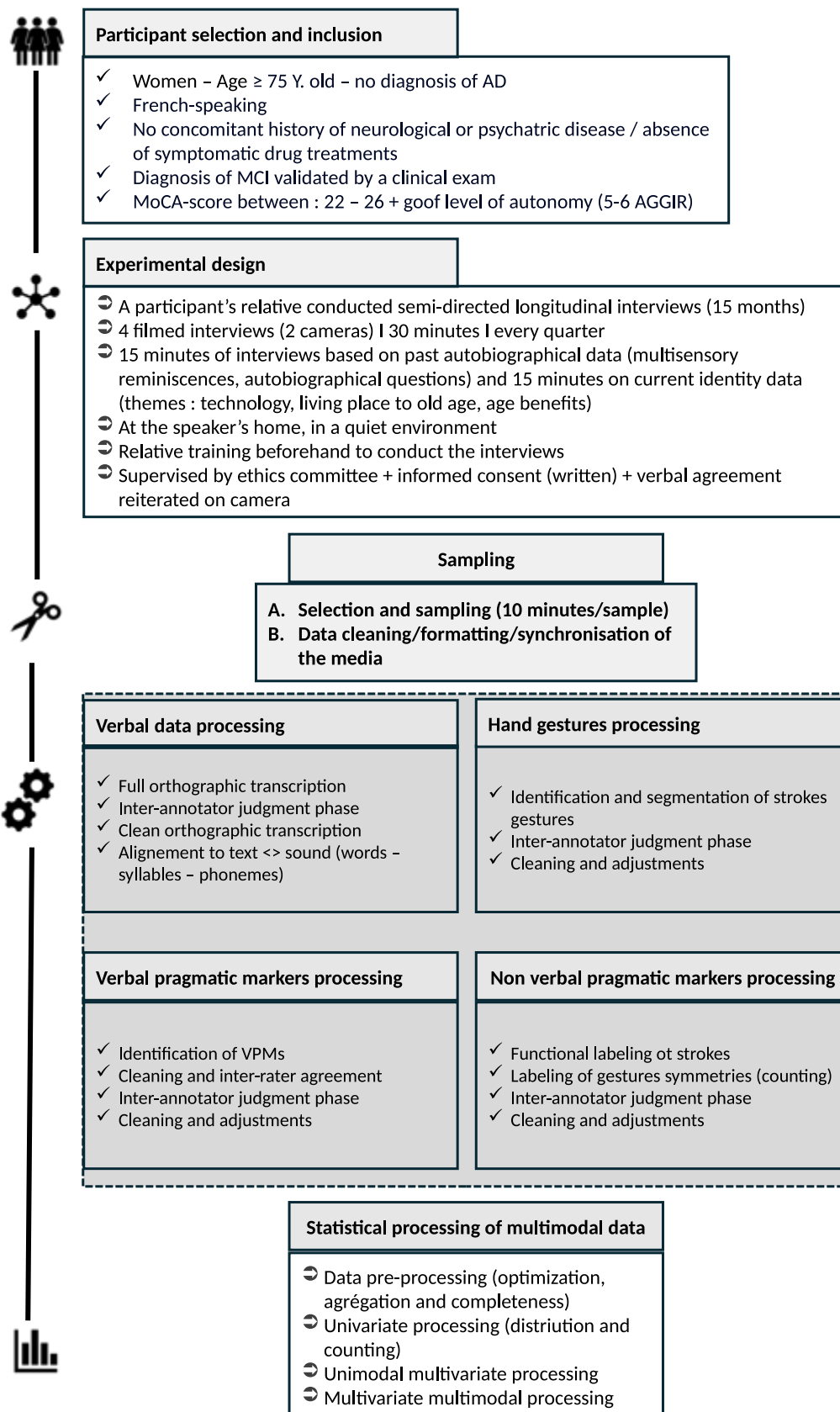- ⮌ Multivariate multimodal processing

Figure 2: Flow chart of the experimental design and treatment chain of the French VIntAGE multimodal analysis procedures (VIntAGE 2024)

on the heterogeneity of speaker profiles, particularly in relation to future anno-
tations and statistical analyses.

2.3.2 Sampling criteria

The duration of a sample is approximately 5 minutes. This duration was validated
during previous interviews conducted within CorpAGEst and during a pilot study
(Duboisdindien, Bolly, and Lacheret 2017), considering the number of hours required
for the future processing of verbal and non-verbal samples for 5 minutes of analysis
by an annotator. This duration may reasonably be exceeded (by up to 1 minute
30 seconds) according to three flexibility criteria: the loquacity of the interviewee, the
impact of the task on the interviewee's production, and a contextualized approach
that respects the interaction initiated by the speakers. If an external event disrupts
the sample, adjusting the sampling boundaries to a different theme being discussed
is preferable. However, it was possible to retain the external event if it occurred
at a particularly poignant or relevant moment for the participant. In such cases, an
extension corresponding to the duration of the event was made, considering the theme
and the conclusion of the speaker's utterance.

2.3.3 Transcription of verbal data and gesture annotation

The transcription conventions, tools employed, and all procedures for gesture annota-
tion and pragmatic markers tagging used in this study are not only available online in
open access, in both English and French, but also easily accessible. Users are encour-
aged to consult these resources to enhance their understanding of the research. The
manual annotation guidelines, corpus design, research reports, and ethical frame-
work are detailed and freely available in English as part of the CorpAGEst (2021)
corpus on Ortolang, including the following.

1. *Annotation manual I: Gestural annotation guidelines*, presenting the detailed method-
   ology on how these variables were processed and analyzed (e. g., coding schemes,
   articulators by articulators);

2. *Annotation manual II*, presenting the speech annotation guidelines.

The reference section indicates the principal reference study of the VIntAGE corpus
(Duboisdindien 2024). For French readers, the entire protocol and detailed method-
ology in French are available online (Duboisdindien 2019, Chapter 5).

3 Example analysis

We conducted a comprehensive statistical analysis of the data. We chose unsuper-
vised multivariate analysis, specifically fuzzy c-means cluster analysis (FCA). The
main advantage of this statistical method is that it allows the analysis and visualiza-
tion, without a priori, of high-dimensional datasets containing individuals described
by several heterogeneous, qualitative, quantitative, or combined variables. The un-
supervised method allows exploring multivariate data (i. e., several variables n used
to describe several individuals p in a matrix of dimension $n \times p$, with $n \ll p$ in gen-
eral) and identifying underlying structures: proximity / distance between individuals,

functions of the modalities of the variables describing them, and correlation / anti-correlation between the variables as a function of the values they take among different individuals (Husson, Lê, and Pagès 2016).

Procedures:

- Transfer to contingency table: After curation, consolidation, and aggregation and before proceeding with multivariate analysis, the data were compiled in the form of a reference table listing and filling in each occurrence of verbal pragmatic markers (VPM) associated with the speaker, the so-called temporal information (interview, themes, time interval of emission), and the pragmatic function attached to it.

- χ^2 test of independence: We then used the contingency table to assess whether a statistical association existed between the two discrete features of interest (speaker and PM annotation). In this study, we wanted to determine whether female speakers showed a similar distribution of PMs in the different functional annotation classes (or not). We formulated the test hypotheses as follows:

  - Null hypothesis (H_0): The speaker does not influence the distribution of PMs in the different functional annotation classes.
  - Alternative hypothesis (H_1): The distribution of PMs in the different functional annotation classes depends on the speaker.

To evaluate these hypotheses, we employed the χ^2 test of independence, which involves comparing the distance between the observed n_{ij} frequencies and the expected e_{ij} frequencies under the assumption of independence.

Mathematically, we divided the VIntAGE corpus and subcorpus data into two main categories:

1. VIntAGE speaker characteristics providing information on the age of the subjects, and their cognitive and empathic abilities and communication skills.

2. Annotation data on verbal and gestural PMs.

4 Reuse potential

Additionally, we have guaranteed the systematic storage and preservation of the sources and the annotated and anonymized files, as well as their future editing, on the Ortolang platform for the benefit of the scientific and clinical community. The research team is currently working on this corpus, with a planned deposit anticipated around April 2025. Due to the nature of the clinical data in the VIntAGE research, access and reusability may be restricted in accordance with privacy laws, and this will be clearly stated on the Ortolang platform. For any partnerships with the authors of this corpus or use for research purposes, it is necessary to contact the principal author at the email address provided. The corpus publication CorpAGEst (2021) is available in open access with Ortolang. The types of linguistic functions and structures analyzed and the non-verbal / gestural variables (e.g., types of gestures, their meanings, and contexts) are based on Bolly and Crible (2015), Crible (2018), and Duboisdindien (2019) and are explicitly described in Duboisdindien (2024, Appendix A, Tables A.1., A.2., A.3, Figure A.1.).

5   Scientific and clinical trajectories

In conclusion, the VIntAGE corpus highlights the importance of a multimodal and experimental approach to understanding communication in older adults. By integrating verbal and non-verbal data, this research sheds light on the adaptive strategies individuals with MCI use to maintain interaction despite cognitive decline. Consequently, VIntAGE substantially contributes to experimental psycholinguistics and neurolinguistics applied to clinical research, offering a rich resource for investigating the cognitive and linguistic processes involved in aging. Its findings have significant implications for research and clinical practice, offering hope for improved understanding and management of cognitive decline in aging adults. For interested readers of this introductory article, the authors of this study will be delighted to carry out work and any other form of partnership.

6   Acknowledgments

Funding

Ethics statement

The entire experimental protocol of the VIntAGE study was approved by the ethics committee of the University of Louvain-la-Neuve in 2013, under reference of the IPSY Institute (Catholic University of Louvain-la-Neuve, Belgium) for the broader Marie

Curie project CorpAGEst, on December 18, 2012 (Grant Agreement No. [PIEF-GA-2012328282], Bolly). The consent form was provided to the study's participants (i.e., speakers and their intimates). The researcher assigns a pseudonym in the recording transcripts (e.g., Leopold for Leon) and replaces the names of individuals and places with fictitious names.

Informed Consent: To adhere to the ethical principles of scientific research and following European legislation, the recordings are subject to a two-step informed consent process (oral and then written): (1) Oral informed consent (recorded at the beginning of each interview): agreement to being recorded and filmed. (2) Written informed consent (signed at the beginning of the second interview): agreement concerning the individual's rights, the guarantee of anonymity, and the dissemination of the research.

Conflict of interest

We confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

Allwood, Jens. 2008. "Multimodal corpora." In *Corpus linguistics: An international handbook,* edited by Anke Lüdeling and Merja Kytö, 207–225. Berlin: Mouton de Gruyter.

Allwood, Jens, LLoredana Cerrato, Kristiina Jokinen, Costanza Navarretta, and Patrizia Paggio. 2007. "The MUMIN coding scheme for the annotation of feedback, turn management and sequencing phenomena." *Language Resources and Evaluation* 41 (3–4): 273–287. https://doi.org/10.1007/s10579-007-9061-5.

Arbach, Najib, and Saandia Ali. 2013. "Aspects théoriques et méthodologiques de la représentativité des corpus." *Corela* HS-13. https://doi.org/10.4000/corela.3029.

Boersma, Paul, and V. Van Heuven. 2001. "Speak and unSpeak with PRAAT." *Glot International* 5 (9/10): 341–347.

Bolly, Catherine, and Dominique Boutet. 2018. "The multimodal CorpAGEst corpus: Keeping an eye on pragmatic competence in later life." *Corpora* 13 (2). https://doi.org/10.3366/cor.2018.0151.

Bolly, Catherine, and Ludivine Crible. 2015. "From context to functions and back again: Disambiguating pragmatic uses of discourse markers." Presented at the 14th International Pragmatics Conference, Antwerp.

CorpAGEst. 2021. *CorpAGEst.* V. 1, 2021-02-07. Edited by Catherine Bolly. Ortolang. Corpus. https://hdl.handle.net/11403/corpagest/v1.

Crible, Ludivine. 2018. *Discourse markers and (dis)fluency: Forms and functions across languages and registers.* Amsterdam: Benjamins. https://doi.org/10.1075/pbns.286.

Cummings, Louise, ed. 2017. *Research in clinical pragmatics.* Cham: Springer.

Darrigrand, Bénédicte, and Jean-Michel Mazaux. 2000. "L'échelle de communication verbale de Bordeaux: Une évaluation des compétences communicatives des personnes aphasiques." *Glossa,* 4–14.

Davis, Boyd, ed. 2005. *Alzheimer talk, text and context: Enhancing communication.* Basingstoke: Palgrave.

Davis, Boyd H., and Jacqueline Guendouzi. 2014. *Pragmatics in dementia discourse.* Newcastle upon Tyne: Cambridge Scholars.

Delano-Wood, Lisa, Mark W. Bondi, Joshu Sacco, Norm Abeles, Amy Jak, David Libon, and Andrea Bozoki. 2009. "Heterogeneity in mild cognitive impairment: Differences in neuropsychological profile and associated white matter lesion pathology." *Journal of the International Neuropsychological Society* 15 (6): 906–914. https://doi.org/10.1017/S1355617709990257.

Duboisdindien, Guillaume. 2019. "Analyse multimodale des marqueurs pragmatiques au sein du vieillissement langagier en situation de Trouble Cognitif Léger." Ph.D. dissertation, University of Paris Nanterre. https://hal.science/tel-02288120.

———. 2024. "The analysis of gestural and verbal pragmatic markers produced by Mild Cognitive Impaired participants during longitudinal and autobiographical interviews." *Clinical Linguistics & Phonetics* 38 (2): 116–137. https://doi.org/10.1080/02699206.2023.2174450.

Duboisdindien, Guillaume, Catherine Bolly, and Anne Lacheret. 2017. "De la sémiologie linguistique à la sémiologie clinique: L'intérêt d'une approche multimodale auprès des personnes âgèes les plus vulnérables." In *Dimensions multimodales des pratiques discursives: Une perspective actuelle pour les linguistes,* edited by Agathe Cormier, Adèle Dé'soyer, Guillaume Duboisdindien, Emmanuelle Eggington, and Véra Kryshtaleva. Nanterre: University of Paris Nanterre.

Duboisdindien, Guillaume, and Catherine T. Bolly. 2025. "Heterogeneity of verbal and gestural functions in spoken discourse with MCI." *Journal of Neurolinguistics* 73:101227. https://doi.org/10.1016/j.jneuroling.2024.101227.

Gilet, Anne-Laure, Nathalie Mella, Joseph Studer, Daniel Grühn, and Gisela Labouvie-Vief. 2013. "Assessing dispositional empathy in adults: A French validation of the Interpersonal Reactivity Index (IRI)." *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences du Comportement* 45 (1): 42–48. https://doi.org/10.1037/a0030425.

Hamilton, Heidi E. 1994. "Requests for clarification as evidence of pragmatic comprehension difficulty: The case of Alzheimer's Disease." In *Discourse analysis and applications: Studies in adult clinical populations,* edited by Ronald Bloom, Loraine Obler, Susan De Santi, and Jonathan Erlich, 185–199. Hillsdale (NJ): Erlbaum.

———. 2005. *Conversations with an Alzheimer's patient: An interactional sociolinguistic study.* Cambridge.

Hamilton, Heidi E. 2008. "Discourse and aging." In *The handbook of discourse analysis,* edited by Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton, 568–589. Malden: Wiley. https://doi.org/10.1002/9780470753460.ch30.

Husson, Franois, Sébastien Lê, and Jérôme Pagès. 2016. *Analyse de données avec R.* Rennes: Presses universitaires de Rennes.

Kennedy, Graeme. 2014. *An introduction to corpus linguistics.* Abingdon: Routledge.

Leahy-Warren, P. 2014. "Social support theory." In *Theories guiding nursing research and practice: Making nursing knowledge development explicit,* edited by Joyce J. Fitzpatrick and Geraldine McCarthy, 85–101. New York: Springer.

Mondada, Lorenza. 2007. "Enjeux des corpus d'oral en interaction: Re-temporaliser et re-situer le langage." *Langage et Société* 3:143–160. https://doi.org/doi.org/10.3917/ls.121.0143.

Mueller, Kimberly D., Bruce Hermann, Jonilda Mecollari, and Lyn S. Turkstra. 2018. "Connected speech and language in mild cognitive impairment and Alzheimer's disease: A review of picture description tasks." *Journal of Clinical and Experimental Neuropsychology* 40 (9): 917–939. https://doi.org/10.1080/13803395.2018.1446513.

Nasreddine, Ziad S., Natalie A. Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings, and Howard Chertkow. 2005. "The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment." *Journal of the American Geriatrics Society* 53 (4): 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x.

Petersen, Ronald C., Glenn E. Smith, Stephen C. Waring, Robert J. Ivnik, Eric G. Tangalos, and Emre Kokmen. 1999. "Mild Cognitive Impairment." *Archives of Neurology* 56 (3): 303. https://doi.org/10.1001/archneur.56.3.303.

Pierrel, Jean-Marie., Christophe Parisse, Jérôme Blanchard, Etienne Petitjean, and Frédéric Pierre. 2017. "ORTOLANG: A French infrastructure for Open Resources and TOols for LANGuage." *Linköping Electronic Conference Proceedings* 136:102–112. https://doi.org/10.3384/ecp136.

Pistono, Aurélie, Mélanie Jucla, C. Bézy, Béatrice Lemesle, Johanne Le Men, and Jérémie Pariente. 2019. "Discourse macrolinguistic impairment as a marker of linguistic and extralinguistic functions decline in early Alzheimer's disease." *International Journal of Language & Communication Disorders* 54 (3): 390–400. https://doi.org/10.1111/1460-6984.12444.

Service-Public.fr. 2024. *Apa: Qu'est-ce que le Gir 1, 2, 3 ou 4 de la grille Aggir?* Paris: Direction de l'information légale et administrative (Premier ministre) / Caisse nationale de solidarité pour l'autonomie (CNSA). https://www.service-public.fr/particuliers/vosdroits/F1229.

Sinclair, John. 2004. "The search for units of meaning." In *Trust the text: Language, corpus and discourse,* edited by John Sinclair and Ronald Carter, 24–48. London: Routledge.

VIntAGE. 2024. *Videos to study Interaction in AGEing.* V. 3, 2024-10-16. Edited by Guillaume Duboisdindien. Ortolang. Corpus. https://hdl.handle.net/11403/vintage/v3.

Winblad, Brengt, Kaitlin Palmer, Miia Kivipelto, V. Jelic, L. Fratiglioni, L. O. Wahlund, A. Nordberg, et al. 2004. "Mild Cognitive Impairment — Beyond controversies, towards a consensus: Report of the International Working Group on Mild Cognitive Impairment." *Journal of Internal Medicine* 256 (3): 240–246. https://doi.org/10.1111/j.1365-2796.2004.01380.x.

Wray, Alison. 2008. *Formulaic language: Pushing the boundaries.* Oxford: Oxford University Press.