The multimodal CorpAGEst corpus: Keeping an eye on pragmatic competence in later life

Abstract
The CorpAGEst project aims to study the pragmatic competence of very old people (75 y. old and more), by looking at their use of verbal and gestural pragmatic markers in real-world settings (vs. laboratory conditions). More precisely, we hypothesize that identifying multimodal pragmatic patterns in language use, as produced by older adults at the gesture-speech interface, helps to better characterize language variation and communication abilities in later life. The underlying assumption is that discourse markers (e.g., *tu sais* ‘you know’) and pragmatic gestures (e.g., an exaggerated opening of the eyes) are relevant indicators of stance in discourse. The present paper’s objective is mainly methodological. It aims to demonstrate how the pragmatic profile of older adults can be established by analyzing audio and video data. After a brief theoretical introduction, we describe the annotation protocol that has been developed to explore issues in multimodal pragmatics and aging. Lastly, first results from a corpus-based study are given, showing how multimodal approaches can tackle important aspects of communication abilities, at the crossroads of language and aging research in linguistics.

**Keywords:** corpus, sociopragmatics, aging, stance, multimodality, gesture, discourse markers

1. Introduction

The main objective of the CorpAGEst project, ‘A corpus-based multimodal approach to the pragmatic competence of the elderly’, is to establish the pragmatic profile of very old people, by looking at their use of verbal and gestural pragmatic markers in real-world settings (i.e. in their everyday environment). The underlying assumption is that discourse markers (e.g., *tu sais* ‘you know’) and pragmatic gestures (e.g., an exaggerated opening of the eyes) are relevant indicators of stance in discourse. Stance relates to the cognitive and affective ability to express and understand points of view, beliefs and emotions, as to be in tune with others and to interact with them (Goodwin et al., 2012). As recently stated by Keisanen and Kärkkäinen (2014), stance taking in the embodied interaction is also concerned with the study of multimodal practices (including language, prosody, gesture, body posture, as well as sequential position and timing, activity, and situation settings). We therefore hypothesize that multimodal pragmatic markers of stance play a role in the assessment of speakers’ overall pragmatic competence, which we define as the ability to use language resources in a contextually appropriate manner (Kasper and Rose, 2002). It is our belief that identifying multimodal pragmatic patterns, as produced by older adults, will help to better characterize language variation and communication abilities in later life.

In contrast to methods used in the psychology of aging (mostly based on data elicited under laboratory conditions), the CorpAGEst corpus-based approach aims to reflect the authentic, spontaneous language use of communicating subjects as closely as possible. As stressed by Chafe (1992), language use ‘does provide a complex and subtle, even if imperfect window to the mind’ and, ‘[b]eing natural rather than manipulated, corpora are in that respect closer to reality’ (1992: 88). This project is innovative for (i) its yet unexplored topic which focuses on the pragmatic competence of very old healthy people, (ii) its valuable and replicable methodology based on the multimodal annotation of naturalistic, spontaneous data,

---

1 http://corpagest.org/

2 The approach is corpus-based rather than corpus-driven as far as our primary goal is to systematically analyze variation and regularities of pre-defined language features in use (viz. pragmatic gestures and discourse markers), by confronting their observation in the corpus and theoretical assumptions on age-related phenomena.
and (iii) its multidisciplinary dimension at the crossroads of pragmatics, corpus linguistics, gesture studies, discourse analysis, and language aging. The project makes it possible to provide researchers with a multimodal corpus targeting a population that has rarely been studied in (socio)linguistics (see Davis and Maclagan, 2016). The method also provides an answer to the increasingly evident need to resort to ecological methods in the area of aging.

Being aware of these challengeable dimensions, we developed a coherent chain of audio and video data treatment that takes into account every step of the analysis from pre- to post-processing with respect to the project purposes, in a ‘circular dynamic’ where each research phase ‘not only relies on the previous steps but already begins with them’ (Mondada, 2007: 810). These steps are: (i) elaboration of the interview protocol and corpus design; (ii) data collection, digitalization, sampling, and edition procedure; (iii) transcription and alignment of audio data; (iv) annotation of audio and video data; (v) mono-modal and multimodal corpus-based analyses; and (vi) permanent storage of primary sources and annotation files. Importantly, we aimed at finding a good balance between the ecological dimension of interaction by eliciting spontaneous speech according to ethno-methodological principles (e.g. in the most non intrusive, non obtrusive, and spontaneous manner), and the technological constraints that guarantee representativeness of the targeted population, comparability between subcorpora and tasks, and machine readability to systematize the analysis (following the three main recommendations in corpus linguistics – cf. Kennedy, 1998).

Videos are practical accomplishments within specific contexts and contingent courses of action, adjusting, anticipating, following the dynamics of sequential unfolding of interaction and of changing participation frameworks. Videos produced within the naturalistic requirements of ethnomethodology and conversation analysis both aim at preserving relevant details and phenomenal field features and reflexively contribute to the configuration of the very interactional order they document. (Mondada, 2006: 15)

The present paper first briefly presents some theoretical issues which are the core of the project (section 2.). We next move towards the principles adopted in the corpus design (section 3.) and the annotation method (section 4.) with special attention paid to the gestural annotation procedure. Finally, results from a corpus-based analysis of discourse markers and gestures are highlighted (section 5.) to illustrate some of the many directions that CorpAGEst research is going.

2. Multimodal pragmatics and aging

2.1 Corpus pragmatics: a multimodal and discursive view

A corpus is defined as ‘a collection of texts assumed to be representative of a given language collated so that it can be used for linguistic analysis’ (Tognini-Bonelli, 2001: 2). Corpus linguistics thereby enables the researcher to describe language phenomena by working across vast sets of electronic data, taking into account the context of situation.

In multimodality and gesture studies, several annotation models of language interaction have been developed during the last decades (cf. Allwood et al. 2007; Blache et al. 2010; Colletta et al., 2009). To some extent, we adhere to their common definition of gesture category that includes every visible bodily action that is intentional, meaningful in context (Kendon, 2004), and primarily linked to the speech production process (Krauss et al., 1996). Nevertheless, we broaden the scope of this lexical-based definition by considering every visible bodily action which is likely to convey a pragmatic meaning at the ‘lower limit of
gesture’ (Andrén, 2014). This means that Pragmatic Gestures (henceforth, PGs) are assumed to possibly convey less conventionalized meaning than representational gestures usually do, and that they may include unintentional and unconscious moves, insofar as they contribute to the meaning of the information conveyed at the metadiscursive level of language. In this view, the model takes into account all nonverbal units which have a stressing or mitigating function (see the wide opening of the eyes in Figure 1a and the head tilt in Figure 1b, respectively), a structuring or punctuating function (e.g., beats – McNeill, 1992), an emotional regulation function (e.g., adaptors such as nose-picking or scratching on the body – Ekman and Friesen, 1969), or an interactive function (see the gaze addressed to the interlocutor in Figure 3b – Bavelas et al., 2002).

Figures 1a and 1b to insert here.

In discursive pragmatics, Discourse Markers (henceforth, DMs) are attributed in a relatively consensual manner with a certain number of syntactic, semantic, and functional properties (Schourup, 1999). First, they are highly frequent in use and are most often indicators of the informal character of texts (see the DMs in bold in the example 1).

1) Jeanne – et anorexique je ne parvenais pas à le retenir / j’ai / alors je pense à quelque ch/ je pensais à anus (rires) / comme c’est quand même le tube digestif hein qui est en bas (rires) et ça va depuis lors je n’oublie plus (rires) et encore l’autre jour aussi un mot / tiens je ne sais p/ tu vois / si / j’ai / j’oublie certains mots / enfin / je retombe dessus après hein… (Corpage corpus; Speaker: ageJM1; Age: 90; 2012)

[and anorexic I couldn’t memorize it / I / so I think of someth/ I thought of anus (laughing) / since it’s still the digestive tube right which is at the bottom (laughing) and it’s ok since then I don’t forget anymore (laughing) and again the other day too a word / see I don’t kn/ you see / yes / I / I forget certain words / well / I remember them afterwards right]

Short in form, they are generally autonomous at the prosodic level with a possible phonological reduction (e.g., t’sais ‘y’know’). They usually have a peripheral syntactic position and weak semantic content. Their function is indexical function, since they serve to guide the interlocutor on the way the information is to be organized and manipulated so that it is appropriately interpreted. To date, DMs are scarcely considered by scholars working on dialogue corpus systems (Bunt et al., 2010), and their integration in multimodal models is often reduced to certain category of DMs (Heeman and Allen, 1999).

It seems obvious from this brief overview in (multimodal) corpus pragmatics that the annotation of pragmatic resources in language interaction has been a topic of research for a few decades. However, the existing models need to be developed further with a view towards multimodality, exhaustivity and interoperability, since they often include a very fragmented view of Pragmatic Markers (henceforth, PMs) in speech and/or gesture.

2.2 Pragmatics in aging: a multidisciplinary inspiration source

It is worth mentioning that language-related changes affect the aging process at different levels, depending on the speakers’ identity and personality, as well as on their actual physical and psychosocial states (Valdois et al., 1990). Some physical and cognitive competences may be affected (e.g. hearing loss, arthritis, working memory and attention deficits, etc.), while others remain relatively stable or even tend to improve with age (e.g. efficient story-telling, self-efficacy, socio-emotional involvement, neural plasticity, etc.). In line with the Baltes and
Baltes’ idea of ‘successful aging’ (1990), we assume that individual variation also holds for the adaptive mechanisms and compensatory strategies developed by older people to optimize their communication with advancing age.

In spite of the growing interest in corpus-based studies to explore pragmatic issues (e.g., Östman and Verschueren, 2011), very little attention has been paid to date to the study of pragmatic competence of healthy older subjects from the angle of language production in spontaneous conversation (Hamilton, 2001). However, a few studies exist that grant a central role to the analysis of the transcription of oral data in aging research, but most of them do not take into account the pragmatic aspects of language (as stressed by Feyereisen and Hupet, 2002) and usually concentrate more on the pathological side of aging than on its adaptive counterpart (see as an exception Gerstenberg, 2015). The study of language use of the oldest-old people (75 years old and more) is in particular not very widespread, despite the fact that age-related language changes above all concern very old people, since communication skills tends to become problematic only relatively late in life: ‘linguistic problems generally do not have a significant effect on interaction until they become quite severe because most individuals, even those suffering from dementia, develop strategies for maintaining interaction in spite of these problems’ (Pecchioni et al., 2005: 48). It is our aim to uncover these subtle language changes that allow the old speakers to adapt (if needed) their resources to the context of communication, as to their interlocutors.

Several yet unexplored aspects of age-related language phenomena are thus investigated in the CorpAGEst project, arising from multiple fields of research in aging. The focus is on verbal and nonverbal language phenomena that are linked, to some extent, to the pragmatic realization of meaning in a given context of communication. As a consequence, the project is rooted in notions at the intersection of several disciplines, on the basis of which the hypotheses on pragmatic competence in late life have been developed:

- in the field of social cognition and emotion, it has been recognized that the healthy subjects’ advancing age may be accompanied by a loss of empathic ability (Bailey and Henry, 2008) and emotional capacities to recognize, anticipate, and relate to others’ emotions (Magai, 2008: 388), liable to affect their capacity for successful social interaction;
- Davis and her colleagues (Davis et al., 2013; Davis and Maclagan, 2014) pointed out, by means of corpus-based studies, an increasing and repeated use of DMs (e.g., so, oh, well) in the older person’s speech at early stage of dementia; the authors see this process as a compensatory strategy to remain involved in the interaction;
- a decrease in the frequency of use of representational gestures (viz. iconic and abstract deictic gestures that are semantically connected to speech) has been pointed out from experimental studies (Feyereisen and Havard, 1999), coinciding with an increase in beats (viz. non-representational, small rhythmic moves that punctuate speech) among older people; it has been suggested by the authors that a greater mastery of verbal competence at earlier stages implies a functional specialization of beats in later life.

In line with Davis’ view, we strongly believe that ‘[b]efore we can optimize abilities [to communicate in old age], we must describe them more precisely, particularly since we need to understand variation and fluctuation in language production, fluency and communicative force over time’ (Davis, 2005: xiv). In other words, if we want in the end to improve the quality of communication abilities of older people, we should first seek to describe how they actually communicate, before to address the why question. It is thus our aim to explore language at old age by looking at language variation within and between individuals (through small-scale studies), but also to extend the view by investigating language variation across time (through longitudinal studies) and languages (through cross-linguistic studies).
Two main research hypotheses served as a basis to develop our corpus analyses: first, it is hypothesized that clusters of PMs of stance (be they verbal and/or nonverbal) are relevant indicators of the emotional and attitudinal profile of the communicating person (H1); secondly, from a developmental perspective, the hypothesis is that subtle communicative changes in the use of PMs (for instance, a functional specialization of DMs, an imbalanced use of PMs across modalities, a reduction in gesture amplitude) are signals of adaptive strategies developed by the aging person to optimize his/her pragmatic competence in everyday life (H2).

3. Corpus design

The multimodal CorpAGEst corpus is comprised of face-to-face conversations between an adult (young or middle-aged) and a very old person (75 y. old and more). The corpus is part of the international CLAREx3 initiative (‘Corpora for Language and Aging Research’), which combines methods in linguistics and issues in aging, and advocates for more corpus-based naturalistic approaches in the field.

3.1 Data collection

The corpus data consist of semi-directed, face-to-face interviews between an adult and a very old subject that were audio-video recorded. All participants are native-speakers of French and healthy persons, that is, without any major injury or cognitive impairment.

The CorpAGEst corpus is two-fold, including transversal and longitudinal subcorpora:

(i) The transversal corpus4 has been built for intra- and inter-individual testing with the purpose of exploring (non)verbal markers of stance and their combination in language interaction, as they are considered relevant indicators of speakers’ emotional and attitudinal behavior; this part includes 18 interviews in Belgian-French (9 subjects; mean age: 85; sex: 8 F; 1 M; 16.8 hrs.; approx. 250,000 words); each interview was repeated a few weeks later in a slightly adapted manner (interaction with an intimate vs. unknown interviewer) and subdivided into two subtasks (focusing on past events vs. present-day life) (see Table 1);

(ii) The longitudinal corpus has been created with the aim to discover whether any compensatory strategy could be observed in the use of nonverbal and verbal pragmatic cues by older individuals over time; this part of the corpus currently gathers interviews from native-speakers of French-French, where each interview is replicated several times during a year and a half and divided into two subtasks (reminiscence task in relation to past events vs. current topic in relation to present-day life) (see Table 2); to guarantee the comparability of results between pre-

3 http://www.clare-corpora.org
4 Audio recordings: 1 or 2 sound signal(s) (format: .wav, 44.1 Hz, 16 bits, mono); video recordings: 2 digital cameras on the upper body and the whole interaction, respectively (format: .mp4, H264).
5 It has been shown that brain injury frequently causes cognitive, behavioral, and physical impairments, which may in turn negatively impacts the person’s life in several respects (among others, his/her autonomy, social relationships, and emotion regulation) (Schönberger et al., 2009: 2157).
6 This part of the corpus is expected to be part of a larger ‘cross-sequential’ corpus in the future (for a discussion about design in life-span perspective, see Schrauf, 2009).
existing transversal data and longitudinal data, the first interview is based on a shortened protocol from transversal data.

Table 2 to insert here.

The CorpAGEst corpus data and metadata will be disseminated through permanent storage in the Ortolang\(^7\) open-source center, providing the corpus with query facilities through a freely web-accessible interface. The Ortolang model follows the basic structure of the OAIS model under OAI-PMH\(^8\) (‘Open Archive Initiative – Protocol for Metadata Harvesting’), thereby guaranteeing the data correction and enrichment in a long-term perspective. In response to ethical principles, the collection of audio and video data involves informed consent (both oral and written) and a systematic procedure to anonymize the data in order to make them publicly available. The participants have agreed to make the original audio and video data accessible without any restriction if they are used for scientific and didactic purposes. However, they can only be disseminated under the form of brief excerpts, even within the scientific community. When used for a wider, public audience (e.g. on a website) and/or on a permanent, publicly available support (e.g. a folder for nurses with advices on communication with aged people), the primary sources should always be blind, by blurring faces in the video data and replacing proper names by a bip in the audio data. Note that the Ethics Commission of the IPSY Institute (Université catholique de Louvain, Belgium) has validated the data collection, design, and treatment of the present project.

3.2 Metadata and subjects

As stressed by Davis (2005), corpus (meta)data are of invaluable help to understand the older people’s communication behavior with respect to human complexity, diversity and integrity. Contextual variables are part of the corpus design, such as the environment type (private vs. residential home), the social tie between the participants in terms of relationship closeness (intimate vs. unknown interviewer), and the task type (focusing on past events vs. present-day life). Metadata also provide information about the interaction situation (e.g., date, place, duration, quality of the recordings), the interviewer and the interviewee (e.g., sex, education, profession, mother tongue, geographic origin, living environment, social tie between interlocutors, subjective scale of life quality and health, scores from clinical testing, etc.). These psychosocial, situational and clinical features are made available for linguists but, most importantly, also for anyone who is interested in the communication behavior of aging people, including experts in geriatrics, caregivers, psychologists, nurses, social workers, retirement homes directors, etc. In addition to corpus-based approach, psychological evaluation scales were used to serve as a basis for methodological comparison: the Montreal Cognitive Assessment test (MoCA\(^9\) – Nasreddine et al., 2005); and the French version of the Interpersonal Reactivity Index (F-IRI\(^{10}\) – Gilet et al., 2013) for the assessment of empathy. As an illustration, the main characteristics of the 9 Belgian-French old speakers who participated in the transversal corpus are the following:

Table 3 to insert here.

---

\(^7\) http://www.ortolang.fr and http://sldr.org

\(^8\) https://www.openarchives.org/pmh/

\(^9\) http://www.mocatest.org/

\(^{10}\) http://www.nbarraco.com/papers/Giletetal_2012_CJBS.pdf
4. Method

The multimodal approach adopted seeks to understand language interaction as a whole, including both verbal and nonverbal contexts of use, by questioning the way in which speech and gesture interact to make sense in real-world settings. The various steps in data treatment and annotation procedure are described in the following sections. The annotation protocol has been developed to detect any pragmatic expression of emotions and attitudes in the everyday communication of the old people. It thus includes every minimal unit of meaning that is less visible or audible (such as adaptors and beats in gesture, or fillers and breaths in speech), which can nevertheless play a role in the on-line planning and transmission of information during interaction.

4.1 A multimodal and multi-level model

The annotation procedure of nonverbal data in the CorpAGEst project is a form-based one (see Müller et al., 2013), extended and applied to facial expressions, gaze, hand gestures, and body gestures (including moves from the head, shoulders, torso, legs, and feet). First, every candidate PG is selected in the samples among all the identified nonverbal units on the basis of their kinetic features (see 4.3). Second, it is the functional annotation of these units (be they representational or not) that will allow specifying their actual role in the language interaction. Thus, we do not a fortiori exclude, at the first step of the analysis, any unintentional or unconscious nonverbal cue that can fulfill a structuring, expressive, or interactive function (e.g., self-adaptors, stress mitigating smiles, or planning devices). According to this two-step procedure, which moves from a kinetic analysis towards a functional one, the question whether the nonverbal cues under scrutiny are considered as playing a role in nonverbal communication at large (Hall and Knapp, 2013: 6) will become of interest only at the end of the analysis process when correlations between kinesis and function would be established.

In order to improve the replicability of the model and favor inter-coder agreement, a detailed annotation guide was established and the annotators were trained. We also adopted a multi-coder approach to cross-check the annotations during the coding procedure. The checking procedure was roughly the same for every unit of analysis (in gesture and speech), with nevertheless some variation between articulators according to their degree of complexity: (i) a short piece of data (about 20% of the sample) was annotated by two annotators with the stated objective to verify the model validity, thus reducing in turn the distance between their respective interpretation; (ii) every coder then annotated the samples independently; (iii) after coding, a close observation of the recurrent cases of disagreement served to solve major cases of ambiguity and uncertainty (using the ELAN function ‘compare annotators’). Notably, one coder has at least partly annotated both speech and gesture at every level, thus guaranteeing the interoperability between modes and articulators (for instance, the values for feet moves were adapted from those for the head, given their similar axial possibilities).

The transcription standards adopted for the oral component are mainly inspired from those of the Valibel research center (Dister et al., 2007)11 and Ciel-F corpus (Gadet et al., 2012) (for further detail, see the CorpAGEst guide for speech transcription and alignment on the project Website). Once transcribed12 and aligned13 to the sound signal, DMs were semi-automatically retrieved from speech and aligned to the video signal in the ELAN annotation

11 http://www.uclouvain.be/valibel.html
12 Using the Praat program (Boersma and Weernink, 2014).
13 Using the EasyAlign plugin for Praat (Goldman, 2011).
files. The protocol for DMs’ annotation follows the one developed within the MDMA\textsuperscript{14} research project (‘Model for Discourse Marker Annotation’) (Bolly et al., 2015a, in press), which aims at covering every step of the analysis from DMs’ identification to their parameter and functional description in context.

Starting with mono-modal analyses (gesture and speech, respectively) and focusing on one group of articulators at a time within each modality (e.g., facial displays for the nonverbal mode), the annotation procedure next moved to a multimodal and functional perspective on pragmatic cues. The multimodal data (text, sound and video) were aligned to the sound signal in partition mode, using the ELAN software (Wittenburg et al., 2006 – version 4.6.2.). ELAN is a tool that has been developed for multimedia annotation which ‘is especially designed to encode and display the multilayer activity we can observe in visual data, whether stemming from hearing or deaf communication’ (Crasborn et al., 2013). It can run on Mac, Windows, or Linux platforms and is freely downloadable from the Language Archive website\textsuperscript{15}.

### 4.2 Sampling

The annotation procedure required selecting and sampling the primary audio and video sources. Regarding the transversal part of the corpus, this was done with respect to the following methodological principles:

- **Sample 1 (Interview n°1)**: it consists in the selection of the first 5 min. of every first interview, with the aim of exploring the way older people actually manage their language competence in a new communication situation;
- **Samples 2 and 3 (Interview n°1)**: they consist in one excerpt of 5 min. each occurring respectively in the middle of the first part (Task 1A: focus on the past) and second part (Task 1B: focus on the present-day time) of the interview; the aim was to build comparable samples taking sub-tasks as dependent variables;
- **Sample 4 (Interview n°2)**: consists of one excerpt of approx. 5 min. taken from the second part (Task 2B) of the second interview, whose thematic content must be on the perception of places to live, with the aim to compare, on the one hand, the samples 4 to each other (dependent variable: individuals) and, on the other, to their corresponding Sample 3 (dependent variable: type of social tie between interlocutors).

As a result, 14 samples have been created by means of the video editor Adobe Premiere Elements (see Table 4).

Table 4 to insert here.

The corpus in its current state shows an imbalance and incomplete picture of the annotated data in terms of participants and articulators at stake. However, there is some consistency in the procedure. For instance, one sample has been fully annotated in order to explore how physiological features and pragmatic functions multimodally combine in one single individual (sample 3 of Nadine’s speech), while 8 samples among four healthy participants have been in-depth analyzed to explore the link between emotions and facial displays.

### 4.3 Gestural annotation scheme

Mainly inspired from the MUMIN project (Allwood et al., 2007), the nonverbal annotation scheme resulted in the creation of a list of physiological parameters and tags for annotating

\textsuperscript{14} http://www.uclouvain.be/467911.html

\textsuperscript{15} https://tla.mpi.nl/tools/tla-tools/elan/
gesture in the ELAN software. The scheme takes into account several physical articulators for the nonverbal mode (see Table 5 below).

Table 5 to insert here.

The originality of the classification, compared to previous annotation models, resides in its exhaustivity combined with an operationalization procedure, which allows to compare the different articulators with each others. It has also the advantage to maximize interoperability by allowing comparisons across modalities and languages, since labels are organized and assigned using comparable controlled vocabulary among variables (in speech and gesture).

In order to make the ELAN annotation schemes (called ‘templates’ in ELAN) easily usable from one coder to another and transposable from one recording to another, the templates were organized and grouped as follows: facial displays and gaze (section 4.3.1), hand gestures (section 4.3.2), upper-body gestures (section 4.3.3), and lower-body gestures (section 4.3.4).

4.3.1 Facial displays and gaze

The decoding of facial displays is of great importance to favor the mutual understanding of speakers who are involved in social interaction (among others, in care settings or during physician-patient encounter). We know, for instance, that facial expressions are impacted by social factors, such as the closeness of relationship between the participants in conversation (Yamamoto and N. Suzuki, 2006). Facial cues also play an important role in managing conversation when they provide listener response, attunement by means of behavioral mimicry, or facilitate the flow of interaction (Chovil, 1991). They can also be viewed as co-speech syntactic devices that punctuate spoken words and sentences, thus helping in the organization of the conversation (for instance, by raising and lowering the eyebrows). Other facial expressions are directly connected with the semantic content of the information (for instance, a smile can mitigate the content of bad memories during a reminiscence task).

Eye-gaze also plays a major role by providing feedback and establishing or sustaining the focus of shared attention, thus mirroring reciprocal arrangement during social interaction (Kendon and Cook, 1969; Rimé and McCusker, 1976). In addition, the quality of a gaze may indicate the degree of involvement of the speaker, the nature of his emotional or psychological state (self-esteem, self-confidence, anxiety), or a cognitive effort in planning or processing information. For instance, gazing away or avoiding eye contact may signal a difficulty to process complex ideas, thus reflecting ‘a shift in attention from external to internal matters’ (Knapp et al., 2014: 301).

We therefore assume that facial expressions and gaze are ‘a major conveyance of both affective and cognitive stance, that is, of intersubjective evaluation, positioning, and alignment of language users in a situation of collaborative interaction’ (Bolly and Thomas, 2015: 26). In order to investigate these pragmatic functions in late life, the physical features that are likely to correlate with them must be first described in a systematic manner. Facial displays (including gaze) were identified according to their location in the face (eyebrows, eyes, gaze, mouth) and then annotated in terms of physiological features (e.g., closed-both for the eyes, corners up or retracted for the lips). The ELAN annotation scheme dedicated to the physiological description of facial expressions (including gaze) is comprised of 7 variables (see Table 6 below).

Table 6 to insert here.
The corresponding annotation file is comprised of 9 annotation lines (‘tiers’) in relation to the 4 physiological articulators under scrutiny (viz. eyebrow, eye, gaze, and mouth). In addition to the kinetic description of these articulators, emotions perceived from the face were annotated according to their emotion category (see Bolly and Thomas, 2015, for further detail). For instance, in the example below (Figure 2), Nadine\textsuperscript{16} has just finished raising her eyebrows (‘Rais’), is about to repeatedly close her eyes (‘Close-R’) and produces a vague\textsuperscript{17}, upward gaze to her left side (‘Vague’, ‘Up-L’), while searching her words (see the ‘euh’ editing term) and expressing surprise through the face.

\textit{Figure 2 to insert here.}

Notably, the in-depth study of Nadine’s interactions includes an analysis of the type of semantic relationship between emotions perceived from the face and their context of appearance, which allowed to distinguish between five types of semantic relationships: facial expressions could be redundant with speech, complementary to speech, contradictory to speech, independent to speech, or accordant with extralinguistic information.

To date, more than 1 hour of video data (66 min. 12 sec.) has been fully annotated on the basis of the facial and emotional annotation scheme (4 speakers: Nadine, Albertine, Anne-Marie, Louise; samples 1, 2 and 3).

4.3.2 Hand gestures

The ELAN template for the notation of hand gestures is comprised of 21 annotation lines (see Figure 3), describing the hand moves according to their manual segmentation into phases, to their form – based on the description of the four traditional parameters in sign language and gesture studies (viz. shape, position, movement, and orientation in space) (Stokoe, 1960) –, and to the contact that often accompanies self-adaptors (e.g., touching one’s nose, rubbing the hands each other) or hetero-adaptors (e.g. manipulating some object such as a tissue or a ring) (Ekman and Friesen, 1969; Ekman, 2004). These parameters are applied to the right and left hand, respectively (see Table 7). The last parameter describes the type of symmetry for the hands, if any occurred. As gestures are best captured in terms of time sequences, they are usually described by segmenting the gesture unit into successive phases, from the beginning of its preparation to the end of its retraction, before going back to a neutral, static position or initiating another gesture unit (Kita et al., 1998). Following Bressem and Ladewig’s work (2011, 2013), we define gesture phases according to their articulatory features as ‘minimal units of analysis, which can be described on their own and in relation to each other’ (2011: 55).

\textit{Figure 3 to insert here.}

To segment hand moves into possibly meaningful units, the basic principle in the CorpAGEst project is that if only one change in one articulatory feature (e.g., change in ‘Shape’ but not in ‘Orientation’, ‘Position’ or ‘Movement’), then the move has to be considered as a whole gesture phase. But if there is a change in at least two parameters, then it has to be considered as two consecutive phases. It is also of importance to stress our semantic-pragmatic definition of ‘Stroke’, as being the most potentially meaningful part of the move, that is, which is

\textsuperscript{16} Nadine is the alias attributed to the ageBN1 speaker (cf. Table 4).

\textsuperscript{17} In our model, a \textit{vague} gaze is defined by its lack of expression and the absence of precise target, often accompanied (when noticeable) with pupil dilation.
supposed to convey meaning in the language interaction (Kendon, 2004). In this project, we adhere to the context-sensitive definition of meaning potentials of language units as ‘affordances […] to combine with (dynamic) properties of contexts in order for situated meanings or interpretations to be constituted’ (Norén and Linell, 2007: 389-390). Given its pragmatic anchor point, the CorpAGEst project therefore considers a move to presumably play a role in the interaction when it transmits at least (partial) semantic-conceptual (iconic, metaphoric, symbolic) or pragmatic-procedural meaning (beats, adaptors, interactive gestures, etc.) given its meaning actualized in context.

For instance, the palm-up family of gestures (Kendon, 2004; Müller, 2004) comprises gestures with the following kinetic features: an open lax handshape with extended fingers, a supine forearm, and an upward facing of the hand (cf. Figure 2 for a prototypical case of palm-up gesture by the left hand). They are said to be PGs (Kendon, 2004), as they contribute to the meaning of the utterance in fulfilling a modal (e.g. by intensifying the expressive content), a performative (e.g. by highlighting a question), or a parsing function (e.g. by marking the discourse’s structure) in combination with the verbal utterance and its context. To date, about 40 min. of video data have been fully annotated on the basis of the annotation scheme for hand gestures (4 speakers: Nadine, Albertine, Anne-Marie, Louise; 7 samples; duration: 39 min. 07 sec.)18.

4.3.3 Upper-body gestures

Nonmanuals (Herrmann and Steinbach, 2013) are of great importance for a better understanding and exhaustive analysis of the emotional and attitudinal behavior of aging people. Gestures from the upper part of the body – such as shoulder shrugs, for instance – can convey disengagement and acquire an epistemic-evidential function in specific contexts (see Debras and Cienki, 2012): the speaker is then positioning him/herself with regard to what is said, taking a multimodal stance in the interaction. Moreover, as stated in Kendon (2004: 265), the more extensive and salient these nonmanuals are, the more expressive the information conveyed by the gesture may be.

The CorpAGEst annotation scheme describes all potentially meaningful bodily actions that originate from the upper-body parts, including head, shoulders, and torso moves (see Table 8).

Head moves
The scheme for head moves has been adapted in such a way to adhere more accurately to the objectivity principle of form-based approaches to gesture, according to which what is annotated first is the physical move (without any interpretation of its potential meaning or function at that stage). In line with systematic coding of body posture (e.g. Dael et al., 2012), head moves are described according to the position and direction of the head in a three-dimensional space, with respect to the three orthogonal body planes (viz. frontal, sagittal, and horizontal).

---

18 The annotation of hand gestures is still ongoing for the other samples.
A change observed in the form and direction (e.g. ‘TiltRight’, ‘TurnLeft’, ‘Back’, ‘Up’, etc.) is usually sufficient to distinguish between two consecutive single head gestures. Ideally, one complex move cannot consist of more than two different moves. But, if two external movements simultaneously occur without being able to decide between one or the other (that is, both seem to be meaningful), then they are both noted by alphabetical order in the annotation span (e.g., ‘Back+Down’). When a single or binary move simultaneously combines with repeated internal movements (not necessarily salient), the annotation of the repeated moves must be added at the end of the tag (e.g., ‘Tilt+DownUp-R’).

**Torso moves**

A torso move is defined as a visible action that originates in a movement of the whole trunk. They are described according to the form and direction of the move. The following categories are distinguished: forward and backward moves, moves from side to side, moves with a rotation of the body, and any other types of moves.

**Shoulder moves**

The notation of shoulders’ moves distinguishes between gestures from the right and left shoulder, respectively. The ELAN tag-set includes every upward-downward and forward-backward movement, which can be a single (e.g. ‘Down’), binary (e.g. ‘DownUp’) or repeated gesture (e.g. ‘DownUp-R’). A second variable is dedicated to the notation of any symmetric shoulders’ movement.

To date, about 45 min. of video data have been at least partly annotated on the basis of the annotation scheme for the upper-body gestures (4 speakers: Nadine, Albertine, Anne-Marie, Louise; samples n°2 and n°3).

### 4.3.4 Lower-body gestures

In the field of nonverbal communication, the focus has traditionally been on the most meaningful part of the body (viz. hand gestures), as on the expressive power (viz. facial displays) and interactive role of nonverbal cues (viz. eye-gazing and head moves). Hence, the lower parts of the body are most of the time absent from the annotation schemes in gesture studies. However, they are recognized to play a role in the expression of the subjective positioning of the speaker as they can reflect his/her emotional or affective state (Ekman and Friesen, 1967: 720; Mehrabian, 1969). To our knowledge, foot gestures have been mostly studied in human computer interactions or virtual environments, by means of motion capture technologies (see Scott et al., 2010, among others). In the CorpAGEst model, feet moves are described by adopting a controlled vocabulary similar to the one given to head moves (see Table 9), due to their shared articulatory possibilities along one main axis (the neck and the ankle, respectively).

**Table 9 to insert here.**

Leg movements have been considered according to the presence or absence of any observable action located in the area including the thigh, the knee and the calf. Given the role of legs’ position in speakers’ stance and emotional state, a specific variable has been added to distinguish between crossed and uncrossed position of these. In order to investigate the way all these articulators interact in real-world settings, one 5 min. sample has been fully annotated taking into account facial expressions and gaze, hand gestures, upper-body and lower-body gestures (speaker: Nadine; sample n°3).
To sum up, we can say that the originality of the CorpAGEst method – compared to existing multimodal models – is in its integrative and comprehensive approach, which tends to reach maximal exhaustivity, systematicity and interoperability between modes and languages. It also adopts an extended view of pragmatics by pushing the boundaries of the so-called ‘pragmatic units’ at their lower limit in speech, including, among others, filled pauses and breath-taking, and gesture, including, among others, adaptors and beats. Following Bressem’s recommendation (2008), the two-step annotation procedure has also been developed to avoid an interpretative bias at any level of analysis: starting from a form-based mono-modal approach to spoken and gestural data, respectively, the analysis then moves to a multimodal functional annotation that takes the overall context of interaction into account.

5. Towards Multimodal Pragmatic Constructions in Nadine’s speech

In order to illustrate the corpus-based methodology, the most striking results that emerged from a multimodal, functional analysis (Bolly, 2015) are highlighted here.

5.1 A multimodal view of pragmatic constructions

This study aims at highlighting the combinatory nature of ‘multimodal pragmatic constructions’ at the intersection of speech and gesture in real-life interaction. As mentioned above, every pragmatic marker (PM) – including discourse markers (DMs) and pragmatic gesture (PGs) – is considered as potentially conveying (at least) one pragmatic meaning in the particular context of its realization. In other words, the purpose is here to discover if there would be any recurrent combinations of nonverbal cues and verbal markers to convey pragmatic functions in multimodal interaction (thus addressing the first hypothesis [H1] formulated in 2.2).

In line with emergentist constructionist approaches to language use (Goldberg, 2006), linguistic units (including the pragmatic ones) are best defined as conventionalized pairings of form and meaning/function, which must be conceived on a continuum between lexis and syntax thus contrasting with the modular view of the linguistic system. Our main hypothesis is that PMs are multimodal constructions where one form (or one pattern of features) is regularly associated with one pragmatic function. Since conversational gestures – as opposed to representational gestures – are traditionally considered idiosyncratic and not conventionalized, the central question addressed is whether there could be more regularity in pragmatic (non)verbal phenomena than what is usually expected: Are there any form-function patterns for (non)verbal PMs emerging from multimodal corpus data? This question is of primary importance for the pragmatic profiling of communicating people, even more when studying language in later life: older people are expected to develop adaptive strategies by means of pragmatic devices while aging (for instance, by producing smiles to remain involved in the communication in spite of hearing loss, or by using nonverbal devices instead of spoken words to reduce the cost of language processing) (cf. our second hypothesis [H2]).

5.2 Annotation procedure for pragmatic markers functions

In the present study, all Pragmatic Hand Gestures (henceforth, PHGs) and DMs have been examined in one single speaker’s language (Nadine; sample n°3), taking into account their synchronous combination with physiological features from the other body parts. A new model for the annotation of pragmatic functions in speech and gesture has been built (see Bolly and Crible, 2015), which allows for a detailed description of the functions of (non)verbal PMs in a multimodal perspective. According to this model, PMs play a role at the
metadiscursive level of language (vs. ideational), helping the addressee to ‘connect, organise, interpret, evaluate, and develop attitudes’ towards the information conveyed (Vande Kopple, 2002: 93). Adapted from Crible’s taxonomy for DMs in speech (2014) and inspired by taxonomies for co-speech gestures (e.g. Bavelas et al., 1992; Colletta et al., 2009), the resulting multimodal annotation scheme currently comprises 44 functions grouped by language domains (based on Halliday, 1970). As CorpAGEst’s focus is on the pragmatic competence in later life, particular attention is paid to the textual and interpersonal functions of pragmatic units, making a distinction between: (i) the structuring function (text-oriented), serving the organization and the cohesion of speech (e.g., bon ‘well’; beats); (ii) the expressive function (speaker-oriented), conveying the speaker’s attitude, feelings, emotions, value judgments, or epistemic stance (e.g., vraiment ‘really’; exaggerated opening of the eyes); and (iii) the interactive function (addressee-oriented), helping to achieve cooperation, to create shared knowledge or intimacy (e.g., tu sais ‘you know’; eye gaze; open-palm, upper-oriented hand gesture, in the extreme-peripheral space of the communicating person).

Given the heterogeneity of the category, we advocate for a bottom-up, inclusive approach that takes every candidate PM into account in the perspective of corpus annotation (without any a priori definition of what should be a PM or not). Next, we attributed specific functions (no more than 2 per unit) to all these DM and PHG candidates. As already mentioned, every visible bodily action that was potentially meaningful in context has therefore been identified in the sample at stake and described in a previous phase, according to the CorpAGEst sets of physiological features. Notably, every gesture phase – including typical strokes and peripheric phases (with the exception of holds that are static) – was considered in our study as possibly conveying a pragmatic meaning. The extraction of DMs in speech was made on the basis of a closed list of markers (detailed in 5.3.2.), including discourse particles, adverbials, parentheticals, comment clauses, connectives, and interjections. All tokens were then manually disambiguated in context and missing DMs were added to the list in a second step. All in all, two groups of two coders each independently annotated the identified units by taking into consideration the entire context of the interaction situation, with one coder who has annotated both speech and gesture to guarantee the interoperability between modes.

5.3 Results

The results reported aim at giving insight in the way Nadine combines speech and gesture units to convey pragmatic meanings in face-to-face interaction with her daughter (for a comparison of Nadine’s interactions with the intimate vs. unknown interviewer, see Lepeut and Bolly, 2016). Nadine has been chosen to serve as a study subject in this study for several reasons: (i) the audio and video material is of exceptionally high quality (bright light in the room, contrasted dressing colors, non-creaky voice, etc.); (ii) Nadine obtained a normal score at the cognitive test, thus indicating that she is undergoing a healthy process of aging.

5.3.1 Overall distribution of Nadine’ gestures

Taking into account Nadine’s entire body (from head to feet), we observed that physiological features attributed to her gesturing correspond to 39 tiers and 2.349 physiological tags in the 5 minutes video sample (sample n°3, 447 tags/min.). As shown in Figure 4, the richest and most detailed body part having been described is the central part of the body including the hands and the torso (with about 60% of the tags; 1,422 tags), then the upper body parts with about 30% of the tags (incl. face, gaze, and head; 687 tags) and, finally, the lower body part with legs and feet moves (10% of the tags; 240 tags).
Among all these physiological features, we focus in the next sections on hand gestures. The identification procedure of gestures yielded a total of 175 potentially meaningful gestures for both hands (87 in the right hand and 88 in the left hand), including ideational, structural, expressive and interactive gestures. In the sample at stake (see Figure 5), most of them are playing (at least partly) a role at the interactive level of language, including 41 cases of partly interactive (e.g., IDE+INT) and 65 cases of fully interactive gestures (INT). Then, 49 gestures are considered as having an expressive function, with 25 cases of partly expressive (e.g. EXPR+INT) and 24 cases of fully expressive gestures (EXPR). We also observed 39 structuring gestures in the data – including 19 cases with mixed domains (e.g., INT+STR) and 20 cases of strict structuring markers (STR). Lastly, we counted 30 gestures functioning, at least partly, at the ideational level of language (13 cases of mixed functions [e.g. IDE+STR] and 17 cases of strict ideational gestures [IDE]).

At the final step, only gestures that fulfill a pragmatic, metadiscursive function (that is, excluding the 17 cases of strict ideational units) were counted as being PGs, leading to a total of 158 PGs for the two hands.

### 5.3.2 Functions of discourse markers and pragmatic hand gestures

For the purpose of the present study, three categories of pragmatic units have been analyzed, which comprise 79 PHGs with the right hand, 79 PHGs with the left hand, and 92 DMs in speech (see Table 10). Note that among the PHGs tagged, 47 gestures were considered symmetric, that is, simultaneously produced by both hands with a similar handshape and a parallel or alternate movement.

Considering DMs with respect to their syntactic category, they were mostly discourse particles (41 tokens among the 12 DM types: ah, pf, bè, ben, euh, euhm, ff, hein, mm, oh, oh là là, pf, quoi), then conjunctions (24 tokens among the 5 DM types et, mais, malgré que, parce que, si), adverbs or adverbial phrases (17 tokens among the 9 DM types: alors, en tout cas, etcétera, là, non, oui, par exemple, quand même, voilà), parenthetical clauses (10 tokens among the 6 DM types: je dis, je me dis, je sais pas moi, je te dis, tu sais, tu vois), and adjectives or pronouns (1 case of each, respectively bon and ça).

Two sets of tags were used distinguishing between the functional domain (or macro-function) and the functional category (or micro-function) of these language units. Comparing the three groups of PMs (viz. right hand PGs, left hand PGs, and DMs in speech) in terms of their functional domains, a statistical difference has been found between the three groups by calculating and comparing the rate of function tags per category in the sample (χ^2 = 14.88; df = 6; p < .05). This significant difference stresses the fact that interactive functions, in contrast to structuring, expressive and mixed functions (combining at least two dimensions), are much more frequent in gesture than in speech. In contrast, DMs show a strong potential to convey expressive meaning (e.g. pf in [1], indicating emotional state and uncertainty) and, less strikingly, to structure discourse (e.g. quoi in [1], indicating the closing of an informational meaningful unit).
Furthermore, a more detailed look at the distribution of interactive micro-functions in the sample data (viz. self-adaptors, common-ground, monitoring, and planning/punctuating markers) reveals the predominant use of adaptors in the gestural mode (37/38 cases) and the very few cases of monitoring gestures in the right hand (1/13 cases) (see Figure 6).

This leads us to consider the former type of adaptive PMs as specific to the gestural mode (vs. speech), independently from the hand at stake, and the monitoring gestures as specific to both DMs and to the weaker hand of the speaker (compared to the right hand, which is the dominant hand of Nadine).

5.3.3 Pairing of physiological patterning and pragmatic functions

Another group of results directly addresses the hypothesis of the emergence of regular multimodal patterns of PMs in speech and gesture. To investigate this, we have automatically retrieved from the annotation file every PM (including PHGs and DMs) that was characterized by an overlap with at least one DM in speech (when considering PHGs) or with at least one nonverbal move from another part of the body, such as a feet move or a closing of the eyes (when considering both PHGs and DMs). This was done by means of the export function of ELAN, which allows extraction of overlapping annotations from different tiers and/or files (‘Export Multiple File As’). The purpose was to analyze every pragmatic multimodal pattern from the data, to uncover what language levels (gesture and/or speech) and what type of gestures (manuals vs. nonmanuals – cf. Herrmann and Steinbach, 2013) are involved in the transmission of pragmatic function by hands and words.

The following tendencies emerge from observation of co-occurring features, formulated here with regard to macro-functions where relevant in the sample:

(i) PHGs always co-occur with one move in the other hand (being not necessarily a stroke or a symmetric move), or with at least one move in one hand in the case of DMs;

(ii) most of the time, PHGs and DMs are simultaneously produced with at least one move in the face, be it a move of the eye, eyebrow, or mouth (80% of the cases; 202 out of 250 PMs); notably, structuring PHGs always co-occur with a facial movement (20 cases); when compared to other domains, interactive PMs less frequently combine with facial expressions with 66 cases out of 89 interactive PMs (74%), against 36 cases out of 43 expressive PMs (84%), and 34 out of 38 cases of structuring PMs (89%) (still in more than 70% of the cases, irrespective of the mode);

(iii) when there is a co-occurrence with a head move, there is more chance that the PM (be it verbal or gestural) fulfills an expressive function (in more than 88% of the cases; 39 out of 43 expressive PMs), and less chance to be interactive (69% of the cases; 62 out of 89 interactive PMs); this tendency is even more significant for DMs (Pearson $X^2$: p < .05) that are more likely to be expressive (89% of the cases; 17 out of 18 expressive DMs) than interactive (46% of the cases; 11 out of 24 interactive DMs) when a head move accompanies the marker;
PMs are less often accompanied by shoulders and torso moves in the sample data (less than 20% of the cases, with only 35 shoulder moves and 5 torso moves), both in speech and gesture.

Going one step further into the analysis, the most frequent functions involving the dominant hand of the speaker (viz. her right hand) have been investigated. Results show that some multimodal patterns combining physiological features and DMs are more prone to be associated with particular micro-functions. To illustrate this, we focus on two sub-categories of particularly frequent interactive functions, which indicate either (a) a planning process at play in the speaker’s mental language processing, or (b) a common-ground effect targeted by the speaker while interacting with the addressee.

Figure 7 to insert here.

To sum up, prototypical planning gestures appear to preferably cluster with fillers and interjections (e.g., *pf*, *euh*), while common-ground gestures mostly co-occur with parentheticals or connectives (e.g., *je (te) dis, et*) (for a more in-depth study of multimodal pragmatic functions, see Bolly and Crible, 2015). Planning gestures also frequently consist in (or integrate) micro-movements, whereas common-ground hand gestures seemed to be wider external moves (notably with a side-in orientation, possibly in peripheral subjective space, and also with flat-lax configuration of the hand) (for more detail, see the on-line annotation guide of the project). Again, planning hand gestures differ from common-ground gestures in Nadine’s interaction, insofar as they often co-occur with self-contact with another body part or an object, head turns, and vague gaze, whereas common-ground gestures are mostly produced with simultaneous gaze addressed to the interlocutor in a straightforward direction.

These results aim at giving a first insight into multimodal patterning of PMs produced by an old speaker, considered as relevant indicators of attitudinal and emotional states in real-life interaction. The next step in the analysis could be to examine whether any effect of the type of move (e.g. head tilt vs. head turn, exaggerated opening vs. closing of the eyes) on the frequency and strength of the multimodal co-occurrence. A multivariate analysis would also be of great value to measure the relative impact of every type of move on the pragmatic pattern taken as a whole. Further analysis should also provide more grounded observations, extending the study to several speakers in order to put the results to the test and widening the scope of PMs (beyond simultaneous co-occurring features) by looking at their previous/left and ulterior/right context.

**Conclusion**

Still in its infancy, the linguistics of aging, especially when based on multimodal data, is a very promising field of research to foster knowledge about language use of older people in real-world settings. Despite their inevitably exploratory dimension, several corpus-based studies have been carried out and others are still ongoing within the framework of the CorpAGEst project. Among others, Lepeut and Bolly (2016) have explored the interactive functions of PHGs in the intersubjective space, focusing on the adaptive behavior of one single old speaker when communicating either with an intimate person or with an unknown person. This study is fully in line with sociolinguistic variationist approaches to language (Coupland, 2007), which assume that older speakers’ communication style varies according to the situational context and under specific psychosocial constraints (e.g., interacting with a nurse or with a relative, in a more or less stressful situation, at their private home or in residential home). From a developmental perspective, Duboisindien’s ongoing work (2015) explores the impact of (non)verbal PMs on the communicative competence of very old people.
over time (in situation of cognitive frailty). The interoperability and transferability of the CorpAGEst multimodal model is also being put to the test by collaborating with specialists in sign languages, with the shared objective to reach a better understanding of the way PGs and signs combine to make sense in deaf and hearing older people’s language (see Bolly et al. 2015b).

In response to socio-economic concerns about the aging population, the projects’ contribution is multiple: (i) contribution to improving knowledge of language competence of healthy elderly in a natural environment and, based thereon, the informed enrichment of the discussion of the concrete strategies to be implemented to promote their ‘aging well’; (ii) enrichment of discourse and multimodal annotation systems, which is a key issue for linguistic description and more generally for understanding language mechanisms; (iii) provision of a multimodal corpus and annotation interaction system that may serve as a basis for further studies on language competence in later life, by bringing to light both individual variations and language regularities. With regard to these three points, the present paper mainly served to highlight the second and third points of the CorpAGEst program: our methodological objective was mainly to demonstrate how the pragmatic profile of older adults can be established by adopting a corpus-based approach to audio and video data. We have to keep in mind, however, that the first point – directed toward the aging well – is the ultimate goal to achieve: to our view, it should guide every scholar who wishes to explore issues in aging through the lens of applied linguistics.

Acknowledgment
The CorpAGEst project has received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n° [PIEF-GA-2012-328282].

References


Figure 1. Pragmatic gestures: expressive (a) eyebrow raising and (b) head tilt, respectively.
<table>
<thead>
<tr>
<th>Task Type</th>
<th>Interview n°1 (with an intimate person)</th>
<th>Interview n°2 (with an unknown person)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A: Descriptive task with a focus on past events</td>
<td>Task 1A: Milestones in aging</td>
<td>Task 2A: Milestones in progress</td>
</tr>
<tr>
<td>Task B: Explicative task with a focus on present-day life</td>
<td>Task 1B: Self-perception of aging</td>
<td>Task 2B: Self-perception of every-day environment</td>
</tr>
</tbody>
</table>

**Table 1.** Tasks for the transversal corpus data collection
<table>
<thead>
<tr>
<th>Task Type</th>
<th>(Interview n°1)</th>
<th>Interview n°2</th>
<th>Interview n°3</th>
<th>Interview n°4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task A: Focus on past events</td>
<td>Task 1A: Milestones in aging</td>
<td>Task 2A: Visual reminiscence from a personal picture</td>
<td>Task 3A: Olfactory reminiscence</td>
<td>Task 4A: Auditory reminiscence</td>
</tr>
<tr>
<td>Task B: Focus on present-day life</td>
<td>Task 1B: Society’s perception of aging</td>
<td>Task 2B: Self-perception of everyday environment</td>
<td>Task 3B: Family and social relationships</td>
<td>Task 4B: Self-perception of aging</td>
</tr>
</tbody>
</table>

**Table 2.** Tasks for the longitudinal corpus data collection
<table>
<thead>
<tr>
<th>Recordings</th>
<th>hh:mm:ss</th>
<th>Speaker</th>
<th>Pseudo</th>
<th>Age</th>
<th>Birth</th>
<th>Sex</th>
<th>Education (n years)</th>
<th>Cognition (MoCA)</th>
<th>Empathy (F-IRI %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ageBN1r-1</td>
<td>1:01:14</td>
<td>ageBN1</td>
<td>Nadine</td>
<td>75</td>
<td>1938</td>
<td>F</td>
<td>12</td>
<td>29</td>
<td>63.57</td>
</tr>
<tr>
<td>ageBN1r-2</td>
<td>0:49:02</td>
<td>ageBN1</td>
<td>Nadine</td>
<td>75</td>
<td>1938</td>
<td>F</td>
<td>12</td>
<td>29</td>
<td>63.57</td>
</tr>
<tr>
<td>ageLL1r-1</td>
<td>1:13:41</td>
<td>ageLL1</td>
<td>Louise</td>
<td>79</td>
<td>1933</td>
<td>F</td>
<td>12</td>
<td>26</td>
<td>66.43</td>
</tr>
<tr>
<td>ageLL1r-2</td>
<td>1:14:25</td>
<td>ageLL1</td>
<td>Louise</td>
<td>79</td>
<td>1933</td>
<td>F</td>
<td>12</td>
<td>26</td>
<td>66.43</td>
</tr>
<tr>
<td>ageBM1r-1</td>
<td>0:59:02</td>
<td>ageBM1</td>
<td>Anne-Marie</td>
<td>82</td>
<td>1932</td>
<td>F</td>
<td>12</td>
<td>28</td>
<td>61.43</td>
</tr>
<tr>
<td>ageBM1r-2</td>
<td>0:50:36</td>
<td>ageBM1</td>
<td>Anne-Marie</td>
<td>82</td>
<td>1932</td>
<td>F</td>
<td>12</td>
<td>28</td>
<td>61.43</td>
</tr>
<tr>
<td>ageDA1r-1</td>
<td>0:59:07</td>
<td>ageDA1</td>
<td>Albertine</td>
<td>84</td>
<td>1929</td>
<td>F</td>
<td>14</td>
<td>29</td>
<td>61.43</td>
</tr>
<tr>
<td>ageDA1r-2</td>
<td>0:52:41</td>
<td>ageDA1</td>
<td>Albertine</td>
<td>84</td>
<td>1929</td>
<td>F</td>
<td>14</td>
<td>29</td>
<td>61.43</td>
</tr>
<tr>
<td>ageMM1r-1</td>
<td>1:20:40</td>
<td>ageMM1</td>
<td>Marie-Louise</td>
<td>84</td>
<td>1928</td>
<td>F</td>
<td>12</td>
<td>23</td>
<td>71.43</td>
</tr>
<tr>
<td>ageMM1r-2</td>
<td>0:57:06</td>
<td>ageMM1</td>
<td>Marie-Louise</td>
<td>84</td>
<td>1928</td>
<td>F</td>
<td>12</td>
<td>23</td>
<td>71.43</td>
</tr>
<tr>
<td>ageAE1r-1</td>
<td>0:41:35</td>
<td>ageAE1</td>
<td>Emile</td>
<td>86</td>
<td>1927</td>
<td>M</td>
<td>15</td>
<td>30</td>
<td>55.00</td>
</tr>
<tr>
<td>ageAE1r-2</td>
<td>0:47:00</td>
<td>ageAE1</td>
<td>Emile</td>
<td>86</td>
<td>1927</td>
<td>M</td>
<td>15</td>
<td>30</td>
<td>55.00</td>
</tr>
<tr>
<td>ageSM1r-1</td>
<td>0:51:14</td>
<td>ageSM1</td>
<td>Marie-Thérèse</td>
<td>89</td>
<td>1924</td>
<td>F</td>
<td>9</td>
<td>23</td>
<td>57.86</td>
</tr>
<tr>
<td>ageSM1r-2</td>
<td>0:58:38</td>
<td>ageSM1</td>
<td>Marie-Thérèse</td>
<td>89</td>
<td>1924</td>
<td>F</td>
<td>9</td>
<td>23</td>
<td>57.86</td>
</tr>
<tr>
<td>ageTL1r-1</td>
<td>0:49:56</td>
<td>ageTL1</td>
<td>Lucie †</td>
<td>92</td>
<td>1920</td>
<td>F</td>
<td>6</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>ageTL1r-2</td>
<td>0:12:47</td>
<td>ageTL1</td>
<td>Lucie †</td>
<td>92</td>
<td>1920</td>
<td>F</td>
<td>6</td>
<td>n.a.</td>
<td>n.a.</td>
</tr>
<tr>
<td>ageDI1r-1</td>
<td>1:25:34</td>
<td>ageDI1</td>
<td>Irène †</td>
<td>94</td>
<td>1919</td>
<td>F</td>
<td>8</td>
<td>13</td>
<td>75.71</td>
</tr>
<tr>
<td>ageDI1r-2</td>
<td>0:51:14</td>
<td>ageDI1</td>
<td>Irène †</td>
<td>95</td>
<td>1919</td>
<td>F</td>
<td>8</td>
<td>13</td>
<td>75.71</td>
</tr>
</tbody>
</table>

**Table 3.** Main characteristics of the study subjects by chronological age (transversal corpus)
<table>
<thead>
<tr>
<th>Speaker</th>
<th>Pseudo</th>
<th>Recordings</th>
<th>Samples</th>
<th>hh:mm:ss.ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>ageBN1</td>
<td>Nadine</td>
<td>ageBN1r-1</td>
<td>Sample 1</td>
<td>00:05:05.00</td>
</tr>
<tr>
<td>ageBN1</td>
<td>Nadine</td>
<td>ageBN1r-1</td>
<td>Sample 2</td>
<td>00:06:05.01</td>
</tr>
<tr>
<td>ageBN1</td>
<td>Nadine</td>
<td>ageBN1r-1</td>
<td>Sample 3</td>
<td>00:05:14.01</td>
</tr>
<tr>
<td>ageBN1</td>
<td>Nadine</td>
<td>ageBN1r-2</td>
<td>Sample 4</td>
<td>00:07:59.02</td>
</tr>
<tr>
<td>ageLL1</td>
<td>Louise</td>
<td>ageLL1r-1</td>
<td>Sample 1</td>
<td>00:05:40.17</td>
</tr>
<tr>
<td>ageLL1</td>
<td>Louise</td>
<td>ageLL1r-1</td>
<td>Sample 2</td>
<td>00:06:38.02</td>
</tr>
<tr>
<td>ageLL1</td>
<td>Louise</td>
<td>ageLL1r-1</td>
<td>Sample 3</td>
<td>00:05:33.13</td>
</tr>
<tr>
<td>ageBM1</td>
<td>Anne-Marie</td>
<td>ageBM1r-1</td>
<td>Sample 1</td>
<td>00:05:34.14</td>
</tr>
<tr>
<td>ageBM1</td>
<td>Anne-Marie</td>
<td>ageBM1r-1</td>
<td>Sample 2</td>
<td>00:06:26.01</td>
</tr>
<tr>
<td>ageBM1</td>
<td>Anne-Marie</td>
<td>ageBM1r-1</td>
<td>Sample 3</td>
<td>00:05:01.11</td>
</tr>
<tr>
<td>ageDA1</td>
<td>Albertine</td>
<td>ageDA1r-1</td>
<td>Sample 1</td>
<td>00:05:10.04</td>
</tr>
<tr>
<td>ageDA1</td>
<td>Albertine</td>
<td>ageDA1r-1</td>
<td>Sample 2</td>
<td>00:04:43.10</td>
</tr>
<tr>
<td>ageDA1</td>
<td>Albertine</td>
<td>ageDA1r-1</td>
<td>Sample 3</td>
<td>00:05:03.12</td>
</tr>
<tr>
<td>ageDA1</td>
<td>Albertine</td>
<td>ageDA1r-2</td>
<td>Sample 4</td>
<td>00:05:59.22</td>
</tr>
</tbody>
</table>

**Total duration:** 01:20:11.10

**Table 4.** Audio-video samples for the first annotation phase (transversal corpus)
<table>
<thead>
<tr>
<th>Nonverbal, gestural</th>
<th>Articulators</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Face and gaze</td>
<td>Eyebrows</td>
</tr>
<tr>
<td></td>
<td>Eyes</td>
</tr>
<tr>
<td></td>
<td>Gaze</td>
</tr>
<tr>
<td></td>
<td>Mouth</td>
</tr>
<tr>
<td>2. Gesture</td>
<td>Hands</td>
</tr>
<tr>
<td></td>
<td>Head</td>
</tr>
<tr>
<td></td>
<td>Shoulders</td>
</tr>
<tr>
<td></td>
<td>Torso</td>
</tr>
<tr>
<td></td>
<td>Legs</td>
</tr>
<tr>
<td></td>
<td>Feet</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verbal</th>
<th>Levels of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Lexis</td>
<td>Orthographic transcription</td>
</tr>
<tr>
<td></td>
<td>Word segmentation and alignment</td>
</tr>
<tr>
<td>2. Pragmatics</td>
<td>Discourse markers identification</td>
</tr>
</tbody>
</table>

**Multimodal, function-based analyses**
- Multimodal annotation of emotions
- Multimodal annotation of pragmatic functions

*Table 5.* Modalities, articulators and levels of analysis in CorpAGEst
<table>
<thead>
<tr>
<th>Articulator</th>
<th>Variable</th>
<th>Tag sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyebrows</td>
<td>Form</td>
<td>Frowning, Raising, Other</td>
</tr>
<tr>
<td>Eyes</td>
<td>Form</td>
<td>Exaggerated Opening, Closing-Both, Closing-One, Closing-Repeated, Other</td>
</tr>
<tr>
<td>Gaze</td>
<td>Direction</td>
<td>Forward-Front, Forward-Right, Forward-Left, Up-Front, Up-Right, Up-Left, Down-Front, Down-Right, Down-Left, Other</td>
</tr>
<tr>
<td></td>
<td>Target</td>
<td>Addressee, Other participant, Vague, Object, Body part, Camera, Other</td>
</tr>
<tr>
<td>Mouth</td>
<td>Openness</td>
<td>Open</td>
</tr>
<tr>
<td></td>
<td>Lips’ corners</td>
<td>Up, Down, Other</td>
</tr>
<tr>
<td></td>
<td>Lips’ shape</td>
<td>Protruded, Retracted, Other</td>
</tr>
</tbody>
</table>

**Table 6.** Articulators and physiological parameters for facial expressions
Figure 2. Annotation of facial displays (ELAN file: ageBN1r-1_sample2)
Figure 3. Annotation of hand gestures (ELAN file: ageDA1r-1_sample2)
<table>
<thead>
<tr>
<th><strong>Variable</strong></th>
<th><strong>Tag set</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hand form (Right, Left)</td>
<td>Phases(^{19}) Prepa, Stroke, Hold/Rest, Return, Partial return, Chain/Transition</td>
</tr>
<tr>
<td></td>
<td>Shape(^{20}) Flat hand closed, Flat hand lax, Flat hand spread, Fist, 1 stretched, 2 stretched, 1+2 stretched, 2 bent, 2+3 stretched, 1-3 stretched, 1-4 stretched, 2-4 stretched, 1+2 connected, 1+2 bent, 1-5 bent, 1-5 spread bent, 2-5 flapped down, 2-5 bent, 1-5 connected, Other</td>
</tr>
<tr>
<td></td>
<td>Orientation(^{21}) CENTER-CENTER, CENTER, PERI left, PERI lower left, PERI lower right, PERI lower right, PERI upper right, PERI upper left, EXTR left, EXTR lower left, EXTR lower right, EXTR upper right, EXTR upper right, EXTR upper left, Other</td>
</tr>
<tr>
<td></td>
<td>Position Up, Down, Back, Forward, Side-in, Side-out, Invisible</td>
</tr>
<tr>
<td></td>
<td>Movement Single external, Single internal, Repeated external, Repeated internal, External+internal, Other, Invisible</td>
</tr>
<tr>
<td>Hand symmetry (Both)</td>
<td>Plane(^{22}) Frontal, Sagittal, Horizontal, Point</td>
</tr>
<tr>
<td></td>
<td>Time Parallel, Alternate</td>
</tr>
<tr>
<td>Hand contact (Right, Left)</td>
<td>Target Self, Self-O, Partner, Object</td>
</tr>
<tr>
<td></td>
<td>Body / Object Forehead, Hair, Cheek, Chin, Eyes, Eyebrow, Nose, Ear, Mouth, Neck, Shoulder, Chest, Abdomen, Arm, Hand, Fingers, Leg, Knee, Lap, Wrist, Object:[name], Other</td>
</tr>
<tr>
<td></td>
<td>Activity Rest, Touch, Percuss-R, Percuss-D, Manip, Move, Rub, Roll, Scratch, Other</td>
</tr>
</tbody>
</table>

**Table 7.** Articulators and physiological parameters for hand gestures

\(^{19}\) In line with McNeill’s (1992) and Kendon’s (2004) seminal works on gesture units.


\(^{21}\) Following the McNeill (1992) bi-dimensional approach to gesture space.

\(^{22}\) Following Boutet’s description of hands’ gesture symmetry (plane and time) (Boutet, 2012).
<table>
<thead>
<tr>
<th>Articulator</th>
<th>Variable</th>
<th>Tag set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>Direction and</td>
<td>Down, DownUp, DownUp-R, Up, UpDown, UpDown-R, Forward,</td>
</tr>
<tr>
<td></td>
<td>movement</td>
<td>ForwBack, ForwBack-R, Back, BackForw, BackForw-R, TiltRight,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TiltLeft, Tilt, Tilt-R, TurnRight, TurnLeft, Turn, Turn-R, Waggle,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
</tr>
<tr>
<td>Torso</td>
<td>Direction</td>
<td>Forward, ForwBack, ForwBack-R, Back, BackForw, BackForw-R, TurnRight,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TurnLeft, Turn, Turn-R, Rotation, Other</td>
</tr>
<tr>
<td>Shoulder (Right,</td>
<td>Direction</td>
<td>Down, DownUp, DownUp-R, Up, UpDown, UpDown-R, Forward, ForwBack,</td>
</tr>
<tr>
<td>Shoulders (Both)</td>
<td>Symmetry</td>
<td>Parralel, Alternate</td>
</tr>
</tbody>
</table>

Table 8. Articulators and physiological parameters for upper body parts’ moves
<table>
<thead>
<tr>
<th>Articulator</th>
<th>Variable</th>
<th>Tag set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feet (Right, Left)</td>
<td>Direction</td>
<td>Down, DownUp, DownUp-R, Up, UpDown, UpDown-R, TurnRight, TurnLeft, Turn, Turn-R, Other</td>
</tr>
<tr>
<td>Leg (Right, Left)</td>
<td>Move</td>
<td>Move</td>
</tr>
<tr>
<td>Legs (Both)</td>
<td>Position</td>
<td>Crossed, Uncrossed, Other</td>
</tr>
</tbody>
</table>

**Table 9.** Articulators and physiological parameters for other body parts’ moves
Figure 4. Distribution of physiological tags among Nadine’s speech (5 min.)
Figure 5. Distribution of functional domains among Nadine’s gestures (5 min.)
<table>
<thead>
<tr>
<th>Domains</th>
<th>Right Hand</th>
<th>Left Hand</th>
<th>Discourse Marker</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structuring</td>
<td>11</td>
<td>9</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>Expressive</td>
<td>14</td>
<td>11</td>
<td>18</td>
<td>43</td>
</tr>
<tr>
<td>Interactive</td>
<td>32</td>
<td>33</td>
<td>24</td>
<td>89</td>
</tr>
<tr>
<td>Mixed</td>
<td>22</td>
<td>26</td>
<td>22</td>
<td>80</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>79</strong></td>
<td><strong>79</strong></td>
<td><strong>92</strong></td>
<td><strong>250</strong></td>
</tr>
</tbody>
</table>

**Table 10.** Pragmatic gestures and discourse markers according to their main functional domain
Figure 6. Interactive micro-functions
Figure 7. Prototypical planning (a) and common-ground (b) hand gestures