Age-related differences in multimodal recipient design: younger, but not older adults, adapt speech and co-speech gestures to common ground

Louise Schubotz, Aslı Özyürek & Judith Holler

View supplementary material

Published online: 04 Oct 2018.

Submit your article to this journal

Article views: 257

View Crossmark data

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS

Check for updates

Age-related differences in multimodal recipient design: younger, but not older adults, adapt speech and co-speech gestures to common ground

Louise Schubotz [a], Aslı Özyürek[a,b,c] and Judith Holler[a,c]

[a]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands; [b]Centre for Language Studies, Radboud University Nijmegen, Nijmegen, The Netherlands; [c]Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

ABSTRACT

Speakers can adapt their speech and co-speech gestures based on knowledge shared with an addressee (common ground-based recipient design). Here, we investigate whether these adaptations are modulated by the speaker's age and cognitive abilities. Younger and older participants narrated six short comic stories to a same-aged addressee. Half of each story was known to both participants, the other half only to the speaker. The two age groups did not differ in terms of the number of words and narrative events mentioned per narration, or in terms of gesture frequency, gesture rate, or percentage of events expressed multimodally. However, only the younger participants reduced the amount of verbal and gestural information when narrating mutually known as opposed to novel story content. Age-related differences in cognitive abilities did not predict these differences in common ground-based recipient design. The older participants' communicative behaviour may therefore also reflect differences in social or pragmatic goals.

1. Introduction

In spite of a growing literature on language and ageing, little is known about the language use of older adults in face-to-face interactions (for comprehensive overviews see e.g. Abrams & Farrell, 2011; Thornton & Light, 2006). This lack of knowledge extends to older adults' use of the gestural modality, a core component of language use in face-to-face settings (Bavelas & Chovil, 2000; Clark, 1996; Kendon, 2004; McNeill, 1992). Considering the prominence of face-to-face interaction in every-day language use, we are thus faced with a serious gap in our understanding of the communicative competencies of older adults as well as the potential role that age-related cognitive changes may play in this respect.

Language used in interaction is produced and tailored for an addressee, shaped by a process called *recipient design* (Sacks, Schegloff, & Jefferson, 1974) or *audience design* (Clark & Murphy, 1983). Recipient design is based on an addressee's communicative needs and affects the way in which language users both speak *and* gesture for others (e.g. Campisi & Özyürek, 2013; de Ruiter, Bangerter, & Dings, 2012; Galati & Brennan, 2014; Hoetjes, Koolen, Goudbeek, Krahmer, & Swerts, 2015; Holler & Stevens, 2007; Holler & Wilkin, 2009). Taking an addressee's perspective into account and designing one's utterances accordingly may be a

cognitively demanding process (e.g. Horton & Gerrig, 2005; Horton & Spieler, 2007; Long, Horton, Rohde, & Sorace, 2018; Wardlow, 2013). Considering that healthy human ageing is frequently associated with changes in cognitive functioning (Salthouse, 1991), systematic age-related changes in multimodal recipient design may be expected. However, although previous studies have investigated older adults' recipient design in speech, as well as their gesture production in general, these two issues have not yet been brought together. It is currently unclear whether, and if so, how older adults use their multiple communicative channels when designing utterances for others and which role general cognitive abilities play in this process. In order to address these issues, we compared younger and older adults' speech and gesture use in a narrative task that required the addressee-based adaptation of utterances, taking cognitive abilities as a potential modulating factor into account.

1.1. Multimodal recipient design in younger and older adults

1.1.1. Verbal recipient design
The ability to engage in recipient design is frequently investigated by manipulating the amount of *common*

Supplemental data for this article can be accessed http://dx.doi.org/10.1080/23273798.2018.1527377.

ground between conversational partners, defined as the knowledge, beliefs and assumptions that conversational partners believe to be mutually shared and that require the appropriate adaptation of utterances (Clark, 1996). Generally, the larger the common ground, i.e. the more information conversational partners mutually share, the less they put into words. This is characterised, for example, by shorter utterances, less complex syntax, or less informational content (e.g. Fussell & Krauss, 1992; Galati & Brennan, 2010; Isaacs & Clark, 1987). Older adults' ability to engage in recipient design based on common ground has previously been compared to that of younger adults using referential communication tasks. Here, participants are required to establish mutual reference to a limited set of objects over the course of several trials, thereby gradually increasing the amount of common ground (Horton & Spieler, 2007; Hupet, Chantraine, & Nef, 1993; Lysander & Horton, 2012). The results of these studies have shown that younger adults' interactions become increasingly more efficient, indicated by shorter utterances and task-completion times on later compared to earlier trials. Older adults, on the other hand, are less efficient than younger adults, indicated by longer utterances, longer task-completion times, and more errors. It thus appears that compared to younger adults, older adults are less successful at interactively designing their utterances for others.

1.1.2. The role of cognitive abilities in verbal recipient design

Horton and Spieler (2007) suggest that older adults' inferior performance on these referential communication tasks may be due to age-related cognitive limitations, specifically difficulties in retrieving partner-specific information from memory (see also Horton & Gerrig, 2005). Additionally, there are indications that working memory may play a role in recipient design: Work on visual perspective-taking abilities in younger (Wardlow, 2013) and older adults (Healey & Grossman, 2016) suggests that working memory plays a significant role when speakers are required to take an addressee's visual perspective into account while formulating their utterances. Older adults perform more poorly on these tasks. Recipient design in conversation similarly requires the awareness that the addressee's perspective may differ from one's own, as well as the ability to incorporate this knowledge during online language processing (see e.g. Brennan, Galati, & Kuhlen, 2010), and should therefore also rely on working memory.

In addition to memory functions, executive control has also been proposed to play a role in verbal recipient design. Hupet et al. (1993) speculate that deficits in

executive control could cause older adults to have difficulties inhibiting irrelevant, egocentric information from entering memory (see also Hasher & Zacks, 1988), which may explain why they have difficulties with partner-specific adaptations in dialogue. Furthermore, executive control has also been related to perspective-taking abilities in younger (Wardlow, 2013) and older adults (Long et al., 2018). Thus, executive control may be underlying the ability to inhibit one's own, egocentric perspective in favour of the addressee's, another crucial component of successful verbal recipient design (Brennan et al., 2010; Keysar, Barr, & Horton, 1998; see also Brown-Schmidt, 2009, for the role of executive function in perspective-taking during language comprehension).

Both working memory and executive control are assumed to decline in healthy ageing (Hasher, Lustig, & Zacks, 2007; Hasher & Zacks, 1988; Salthouse, 1991; but see Verhaeghen, 2011, for a more critical examination of the role of executive functions in age-related cognitive change). One of the aims of the current study was therefore to establish whether these factors contribute to the behavioural differences in verbal recipient design previously observed in younger vs. older adults.

1.1.3. Multimodal recipient design

Most of the studies described above do not consider the multimodal character of face-to-face language use.[1] Yet, information conveyed visually is essential to face-to-face interaction. Especially *representational* co-speech gestures, i.e. "gestures that represent some aspect of the content of speech" (Alibali, Heath, & Myers, 2001, p. 172), contribute crucially to the meaning of a message. For example, speakers can use their hands to indicate the size or shape of an object, to depict specific aspects of an action, or to spatially locate referents that they mention in their speech by pointing. There is a close semantic and temporal alignment between representational co-speech gestures and the speech they accompany (Kendon, 2004; McNeill, 1992; see Özyürek, 2017, for a recent review). However, rather than being fully redundant, gestures often depict information that semantically adds to and complements what is being said (Holler & Beattie, 2003a, 2003b; Rowbotham, Holler, Wearden, & Lloyd, 2016). Moreover, like spoken utterances, co-speech gesture use is sensitive to social context variables. For example, representational gesture rate (e.g. the number of gestures produced per 100 words) is modulated by the visibility between speaker and addressee (e.g. Alibali et al., 2001; Bavelas, Kenwood, Johnson, & Phillips, 2002; Mol, Krahmer, Maes, & Swerts, 2011), as well

as by dialogic interaction (e.g. Bavelas, Gerwing, Sutton, & Prevost, 2008). Addressee location and feedback influence how gestures represent semantic information (Holler & Wilkin, 2011; Kuhlen, Galati, & Brennan, 2012; Özyürek, 2002) and how frequently gestures occur in relation to speech (Jacobs & Garnham, 2007). Hence, for a fuller understanding of older adults' ability to communicate with others, it is necessary to take information conveyed in the gestural modality into account.

Research with younger adults shows that common ground appears to affect speech and gesture in similar ways. In the presence of mutually shared knowledge, when common ground is assumed, gestures often become less informative (e.g. Gerwing & Bavelas, 2004; Hilliard & Cook, 2016; Holler & Stevens, 2007; Parrill, 2010), and/or less frequent, at least in absolute terms. In relative terms, this means that, most commonly, speech and gesture reduce to a comparable degree so that gesture rate does not differ in the presence or absence of mutually shared knowledge (e.g. Campisi & Özyürek, 2013; de Ruiter et al., 2012; Galati & Brennan, 2014; Hilliard & Cook, 2016; Hoetjes et al., 2015).[2] This is in line with the notion that the two modalities operate as a single, integrated system (Kita & Özyürek, 2003; McNeill, 1992; So, Kita, & Goldin-Meadow, 2009), and that this speech-gesture system operates in a coordinated and flexible manner, in response to current communicative demands (e.g. Kendon, 1985, 2004). It is currently unclear however, whether the speech-gesture system is equally flexible in older adults, particularly when designing utterances for others. The present study will address this issue. In doing so, we also to take into account the role of cognitive abilities, as there are indications that gesture production is closely tied to cognitive functions.

1.1.4. The role of cognitive abilities in multimodal utterances and recipient design

Previous research has shown close ties between general cognitive abilities and gesture production. In order to understand whether and how older adults adapt their multimodal utterances to an addressee's needs, we therefore also have to take the cognitive functions of gestures into account.

Generally speaking, gesturing is assumed to provide the speaker with a cognitive benefit. Co-speech gestures may aid the speaker in the speech planning process, e.g. in conceptual planning (Hostetter, Alibali, & Kita, 2007; Kita & Davies, 2009; Melinger & Kita, 2007), or by lightening cognitive load more generally, i.e. freeing up cognitive resources during speaking (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner Cook, Yip, & Goldin-Meadow, 2012). Limited cognitive abilities lead

to an increase in gesture frequency, e.g. lower visual working memory (Chu, Meyer, Foulkes, & Kita, 2014), lower verbal working memory (Gillespie, James, Federmeier, & Watson, 2014), or lower phonemic fluency in combination with higher spatial skills (Hostetter & Alibali, 2008). Although differences in the tasks used to assess cognitive functioning and to elicit gestures make the individual studies difficult to compare, the results can be interpreted as further support for gesticulation as a compensatory mechanism for individuals' weaker cognitive abilities.

Based on the supposed cognitive benefit of gesticulation and the generally assumed age-related declines in working memory and other cognitive functions (Salthouse, 1991), one might expect older adults to gesture more than younger adults. However, the general observation is that older adults produce *fewer* representational co-speech gestures. This has been found for tasks including object (Cohen & Borsoi, 1996) or action descriptions (Feyereisen & Havard, 1999; Theocharopoulou, Cocks, Pring, & Dipper, 2015). Feyereisen and Havard (1999) propose that the observed difference may be due to different speech styles, arguing that there may be a "trade-off between richness of verbal and gestural responses" (p. 169) causing older adults to produce fewer representational gestures when facing the task of speaking and gesturing concurrently. Similarly, Theocharopoulou et al. (2015) suggest that older participants encode information verbally rather than visually, resulting in more verbal elaboration and fewer gestures. These findings suggest an age-related shift in the speech-gesture system, with older adults relying relatively more on speech than on gestures.

However, none of these studies used a communicative paradigm in which older speakers interacted with co-present, non-confederate addressees, a factor that can significantly affect communicative behaviour (e.g. Kuhlen & Brennan, 2013). Thus, whether older adults' decrease in gesture production also manifests itself in contexts where there is a real addressee present and to what extent older adults can then adapt their gestures to the needs of their addressees – given that recipient design itself might be a cognitively demanding task – remains unknown.

1.2. The present study

The main goals of our research are therefore to find out whether, and if so how, younger and older adults differ in their use of speech and co-speech gestures when interacting with an addressee, i.e. whether they adapt their utterances to mutually shared knowledge between

speaker and addressee, and whether differences in addressee-based adaptations may be related to differences in cognitive abilities.

In order to address these issues, we designed a narration task in which a primary participant (the *speaker*) narrated six short comic strips to a secondary participant (the *addressee*), manipulating whether story content was shared (common ground or CG) or not (no common ground or no-CG) between participants. We thus induced a form of personal common ground (Clark, 1996), in which the mutually shared knowledge existed from the outset of the interactions rather than building up incrementally (as in e.g. Horton & Spieler, 2007; Hupet et al., 1993; or Lysander & Horton, 2012).

As for cognitive abilities, we assessed speakers' verbal and visual working memory (verbal and visual WM) as well as executive control and semantic fluency. As summarised above, verbal WM and executive function have previously been related to verbal recipient design (Hupet et al., 1993; Long et al., 2018; Wardlow, 2013). Furthermore, verbal and visual WM have been found to be related to gesticulation in general (e.g. Chu et al., 2014 for visual WM; Gillespie et al., 2014 for verbal WM). Finally, we assessed semantic fluency as an indicator of word finding difficulties, which are thought to increase with increasing age (e.g. Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Burke, MacKay, Worthley, & Wade, 1991), and may be related to gesticulation (Rauscher, Krauss, & Chen, 1996).

Our main dependent variables were the speech-based measures "number of words" and "number of narrative events per narration", and the gesture-based measures "gesture rate per 100 words" as well as the "percentage of narrative events accompanied by a gesture" (or *multimodal events*). We included both speech-based measures in our analysis, as word counts are a global measure of narration length, while number of narrative events serves as a better approximation of the amount of information contained in the narration. Similarly, gesture rate per 100 words globally captures a speaker's relative weighting of gestures to speech, normalising for differences in narration length (e.g. Alibali et al., 2001), whereas the percentage of multimodal events is a closer approximation of the amount of semantic information contained in gesture relative to that contained in speech.

In addition, we coded speakers' explicit references to common ground, as this can provide a further indication of their awareness of mutually shared knowledge. Also, we coded the addressees' verbal and non-verbal feedback in order to control for the possibility that any age-related differences in the speakers' behaviour might be attributable to systematic age-related differences in addressee behaviour.

In line with previous findings, we expected an effect of our common ground manipulation on speech production such that younger adults would use fewer words and include fewer narrative events when relating shared as opposed to novel information (e.g. Campisi & Özyürek, 2013; Fussell & Krauss, 1992; Galati & Brennan, 2010, 2014; Holler & Wilkin, 2009; Isaacs & Clark, 1987). Based on the results obtained by Horton and Spieler (2007), Hupet et al. (1993) and Lysander and Horton (2012), we expected this effect to be significantly smaller in older adults. We additionally aimed to investigate the impact of cognitive abilities on recipient design in speech, expecting that older adults' lower verbal working memory and lower executive control would be associated with a smaller reduction in words and narrative elements (based on the work by e.g. Healey & Grossman, 2016; Horton & Gerrig, 2005; Hupet et al., 1993; Long et al., 2018; Wardlow, 2013).

Regarding the effect of the common ground manipulation on gesture production in younger adults, we expected an overall reduction in gesture frequency and semantic content, in line with the studies cited above. Note that we refrain from making directed predictions for the effect of common ground on gesture rate and multimodal utterances specifically, though, since previous findings vary with respect to the *proportional* reduction of gesture in relation to speech (see Holler & Bavelas, 2017, for an overview). Instead, our focus is the direct comparison between younger and older adults in how they adapt their multimodal utterances to the addressee's knowledge state. Due to the previously found age-related differences in verbal behaviour in relation to common ground (Horton & Spieler, 2007; Hupet et al., 1993; Lysander & Horton, 2012) and due to speech and gesture functioning as one integrated system (Kita & Özyürek, 2003; McNeill, 1992), we predict older adults to be less adaptive to common ground than younger adults, not only in their speech but also in the way they draw on gesture when designing utterances for their recipients.

For a general effect of age on representational gesture production, two possible hypotheses can be formulated considering the literature summarised in the previous section. Based on the findings by Cohen and Borsoi (1996), Feyereisen and Havard (1999), and Theocharopoulou et al. (2015), we might expect older adults to gesture at a lower rate than younger adults. On the other hand, due to potential age-related cognitive limitations, older adults may actually gesture more than younger adults in order to free up cognitive resources (Goldin-Meadow et al., 2001; Wagner Cook et al., 2012) or compensate for weaker cognitive abilities (Chu et al., 2014; Gillespie et al., 2014; Hostetter & Alibali, 2008).

2. Method

2.1. Participants

Thirty-two younger adults (16 women) between 21 and 30 years old (M_{age} = 24.31, SD = 2.91) and 32 older adults (16 women) between 64 and 73 years old (M_{age} = 67.69, SD = 2.43) participated in the study. All participants were native Dutch speakers with self-reported normal or corrected-to-normal vision and hearing and no known history of neurological impairment. Each participant was allocated to a same-age and same-sex pairing. The role of speaker or addressee was randomly assigned and kept constant across the entire experiment. Only the speaker data were analysed here. All participants in the role of speaker had minimally secondary school education, except for one older participant who only had primary school education. Participants were recruited from the participant pool of the Max Planck Institute for Psycholinguistics and received between € 8 and € 16 for their participation, depending on the duration of the session. The experiment was approved by the Ethics Commission for Behavioural Research from the Radboud University Nijmegen.

2.2. Materials

Six black-and-white comic strips from the series "Vater und Sohn" (by cartoonist e.o. plauen, for an example see appendix A) were used to elicit narratives. Each strip consisted of a self-contained story, which centred on the activities of a father and a son. Half of the strips consisted of four frames, the other half of six frames. The strips did not contain any writing but consisted of black and white drawings only and were not known to the participants beforehand. Four experimental lists determined the order in which the different strips were presented. Initially, we created two orders of presentation for the six stories, one being the reverse of the

other. In doing this, we alternated between four- and six-frame stories. In a second step, we assigned the condition in which the stories occurred. For each story, either the first or the second half (corresponding to two or three frames, depending on story length) could be presented in common ground. We alternated between which half of each story would be presented in common ground (e.g. first story – first half, second story – second half, third story – first half, etc.). Counterbalancing the order of common ground presentation across lists ultimately resulted in four experimental lists. Each list was tested eight times, distributed evenly across age groups and sexes.

2.3. Procedure and common ground manipulation

Upon arrival, the speaker and the addressee were asked to sit in designated chairs at a table at 90° from each other. Two video cameras were set up on tripods at a small distance from the table, one of them getting a frontal view of the speaker, and the other one positioned such that it captured both speaker and addressee (see Figure 1 for stills from the two cameras). Sound was recorded with an additional microphone suspended from the ceiling over the table and connected to the speaker camera.

Participants were introduced to each other and received a description of the experiment. This and all subsequent instructions were given both in writing and verbally to ensure that all participants received and understood the information necessary to successfully participate in the experiment. Signed consent was acquired from all participants.

For the narration task, all participants completed one practice trial and six experimental trials, narrating a total of seven stories. At the beginning of each trial, both participants were presented with either the first or the second half of the comic strip and were instructed to look at it together for a limited amount of time without

Figure 1. Example of the lateral (left panel) and frontal (right panel) views of the speaker in the experimental set-up. In this frame, the speaker refers to "a really big fish", both in her speech and in her gesture.

talking, with the aim to experimentally induce common ground about this part of the story. Hence, in each trial there was both CG and no-CG content. Subsequently, the drawings were removed and a screen was put up on the table between speaker and addressee. The speaker then received the full story to look at, with no time limit imposed. Once the speaker signalled that she had understood and memorised the story, drawings and screen were removed again and the speaker narrated the entire story to the addressee. She was instructed to narrate the full story, keeping in mind that the addressee had already seen part of it. Addressees were instructed to listen to the narrations and ask all clarification questions at the end. Then the screen was put back up and the addressee answered a question about the story in writing.[3] Participants received no feedback about the accuracy of these answers so as to not influence speakers' communicative behaviour. Depending on the pair, the task took about 20–30 min. After the experimental tasks were completed, the addressee was allowed to leave, while the speaker performed the cognitive tests.

2.4. Transcription and coding

2.4.1. Speech coding
All recordings from the two cameras were synchronised and subsequently segmented into trials. Transcription of speech and annotation of gestures was done in Elan (Version 4.9.4, 2016; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). For all segments, the speaker's initial narration, i.e. the first retelling of the full story without potential subsequent repetitions, was identified. All analyses reported here are based on these initial narrations only, discarding repetitions or clarifications elicited by the addressee following the initial narration. This is motivated by the fact that the focus of our study was the effect of our experimental manipulations on the speakers' behaviour rather than the impact of speaker-addressee interaction (for a similar argument see Horton & Gerrig, 2005). Speech from the speaker was transcribed verbatim, including disfluencies such as filled pauses and word fragments. However, disfluencies were excluded from the word counts presented in the results section, as we were mainly interested in speech content and did not want potential age-related differences in the number of disfluencies to influence the word count (e.g. Mortensen, Meyer, & Humphreys, 2006). For this reason, we also distinguished between speech belonging to the narrative proper (i.e. relating to story content) and non-narrative speech such as statements about the task or comments relating to the speaker or the addressee (for this distinction see McNeill, 1992).

2.4.2. Explicit references to common ground
Among the non-narrative speech, we identified explicit references to common ground, i.e. statements such as "this time we saw the first half together". These explicit references to common ground give additional insight into whether participants were aware of the shared knowledge or not and will be reported separately in the results section.

2.4.3. Narrative event coding
For the narrative event coding, we roughly followed the procedure described in Galati and Brennan (2010). We devised a narrative event script for each of the six stories, containing all elements that we deemed necessary in order to narrate the story accurately and fully (for an example see appendix B). For the largest part, these were observable events that advanced the plot, with the exception of a few inferences on the intentions of the stories' characters. One event roughly consisted of one "idea unit" (Butterworth, 1975) and frequently corresponded to one syntactic clause. We then checked these scripts against the actual narrations, including additional events in the script if they were included by a substantial number of participants across both age groups. On average, the 4-frame stories contained a total of 18.67 ($SD = .6$) events and the 6-frame stories contained a total of 27.67 events ($SD = .6$). Collapsed across both story types, each story contained 4.63 events per frame ($SD = .11$), with the actual number of events per frame ranging from 1 to 7.

In a subsequent step, we scored each participant's narration based on these fixed scripts for whether the scripted event was contained in the narration or not (note that we only took into consideration the spoken part of the narrations here). In cases where only part of the event was included in the narration, the participant received half a score. A second coder blind to the experimental hypothesis coded 10% of the trials (N = 20). Inter-rater agreement on narrative event scoring was 94% overall.

2.4.4. Gesture coding
For the gesture coding, we first identified all co-speech gestures produced by the speaker during narrative speech, disregarding non-gesture movements as well as gestures accompanying non-narrative speech. Our unit of analyses was the gestural *stroke*, i.e. the most effortful part of the gesture determined according to criteria established in previous co-speech gesture research (Kendon, 2004; Kita, van Gijn, & van der Hulst, 1998; McNeill, 1992). We then categorised these strokes as *representational* and *non-representational* gestures (see

Alibali et al., 2001). For our purposes, representational gestures include *iconic* gestures, which iconically depict shape or size of concrete referents or represent physical movements or actions;[4] *metaphoric* gestures, which resemble iconic gestures but relate to speech in a metaphorical manner (e.g. a rotating movement of the hand to indicate the passing of time); and pointing gestures or *deictics*, i.e. finger points to a specific location in imaginary space, e.g. that of a story character (McNeill, 1992).

All other gestures were considered non-representational and include what are frequently called *beat* gestures, i.e. biphasic movements of the hand, for example to add emphasis, as well as *pragmatic* gestures (Kendon, 2004), i.e. gestures which have pragmatic functions, for example to convey information about how an utterance should be interpreted, or relating to managing the interaction more generally (Bavelas, Chovil, Lawrie, & Wade, 1992, 1995).

A second coder blind to the experimental hypotheses coded 10% of the trials for stroke identification, and another 10% of the trials for gesture categorisation. Inter-rater agreement on stroke identification, based on stroke onsets and offsets, was 92.3%. Inter-rater agreement on gesture categorisation was 97.9%, Cohen's Kappa = .95.

2.4.5. Gesture rates

As we were mainly interested in the semantic content of the narratives and the accompanying gestures, in our analyses we focus exclusively on the representational gestures (i.e. iconic, metaphoric, and abstract deictic gestures). In addition to reporting the raw representational gesture frequency as a descriptive measure, we used two different measures of gesture production in relation to speech in our main analyses.

2.4.5.1. Representational gesture rate (gestures per 100 words).
We computed a gesture rate per 100 words (see above for criteria on word count) by dividing the number of gestures by the number of words a given participant produced for each condition within each trial separately and multiplied this by 100.

2.4.5.2. Percentage of multimodal events.
We computed a percentage of multimodal events for each participant by dividing the number of narrative events accompanied by a gesture by the total number of narrative events per condition within each trial and multiplied this by 100.

In appendix C, we additionally report the analyses of gesture frequencies in order to be able to draw direct comparisons between our study and previous studies on gesticulation in older adults (Cohen & Borsoi, 1996;

Feyereisen & Havard, 1999; Theocharopoulou et al., 2015), as well as the analyses of gesture rate per narrative event, as used e.g. by Galati and Brennan (2014).

2.4.6. Addressee feedback

As stated in the introduction, gesture production has been found to be sensitive to addressee feedback (e.g. Holler & Wilkin, 2011; Jacobs & Garnham, 2007; Kuhlen et al., 2012). In order to ensure that any potential difference in gesture production between younger and older adults would not be due to systematic differences in addressee behaviour, we coded the addressees' verbal (backchannels, questions, other verbal remarks) and non-verbal feedback (head movements, manual gestures) for two of the six stories. An analysis of this addressee behaviour is reported in the results section.

2.5. Cognitive measures

Participants performed the Operation Span Task (Ospan) as a measure of verbal WM, the Corsi Block Task (CBT) as a measure of visuo-sequential WM, the Visual Patterns Test (VPT) as a measure of visuo-spatial WM, the Trail Making Test (TMT) as a measure of executive function, and the animal naming task to assess semantic fluency. Detailed descriptions of these cognitive tasks, how they were administered, and how the scores were computed can be found in appendix D.

2.6. Statistical methods

To investigate the influence of age and the common ground manipulation on the main speech- and gesture-based measures (word and narrative event count, gesture rate and percentage of multimodal events), as well as on explicit reference to common ground and addressee feedback, we fitted linear mixed-effect models in R version 3.2.1 (R Development Core Team, 2015), using the package lme4 (Bates, Maechler, & Bolker, 2017). We only report best-fitting models established via likelihood ratio tests for model comparisons, eliminating all non-significant predictors in the model comparison process. All the models reported contain random intercepts for participants and items (story), as well as by-participant random slopes for the common ground manipulation unless explicitly stated otherwise. Reported *p*-values were obtained via the package lmerTest (Kuznetsova, Brockhoff, & Bojesen Christensen, 2016). The function lsmeans from the package emmeans (Lenth, 2018) was used to test linear contrasts among predictors for the individual models.

To investigate the influence of cognitive abilities on our main dependent measures, and to test whether

potential age-related differences in verbal and gestural behaviour could be attributed to age-related differences in cognitive abilities, we applied the same basic procedure as described above. We built on the best-fitting models established in the previous analyses and created separate models for each cognitive predictor. As the analyses were exploratory, we performed a backwards-model-stripping procedure, starting out with a full model including the cognitive predictor of interest, age, and the common ground manipulation, as well as all their interaction terms, eliminating non-significant interactions and predictors in the model comparison process.

3. Results

3.1. Gesture frequency and gesture types per age group

Younger adults produced 849 gestures accompanying narrative speech, out of which 542 were iconic gestures (63.84%), 7 metaphoric gestures (0.82%), 104 deictic gestures (12.25%), and 196 non-representational gestures (23.09%). Older adults produced 673 gestures accompanying narrative speech, out of which 479 were iconic gestures (71.17%), 13 metaphoric gestures (1.93%), 60 deictic gestures (8.92%), and 121 non-representational gestures (17.98%). Note again that only representational gestures were included to compute the dependent measures gesture frequency, gesture rate, and percentage of multimodal events reported in the following sections.

3.2. Effects of age and common ground on speech and co-speech gesture

Mean values and standard deviations for the various dependent measures by age group and condition are listed in Table 1. The distribution of observations for word count, narrative event count, gesture rate, and percentage of multimodal events is displayed in Figure 2.

3.2.1. Words and narrative events
As described in Section 2.6, we fitted linear mixed effects models to evaluate the effects of age and common

ground manipulation, as well as their interaction, on word count and narrative event production. The models are summarised in Table 2.

In order to obtain the simple main effects of the two predictors we compared nested models to the omnibus models via likelihood ratio tests, excluding only the predictor variable of interest, one at a time, but keeping the respective other predictor as well as the interaction term. There was no main effect for age, such that younger and older adults did not differ in the overall number of words and narrative events they produced (both p's > .05). There was an effect of common ground manipulation, significant for word count ($\chi^2(1)$ = 15.88, $p < .001$) but not for narrative event count ($\chi^2(1) = 3.59$, $p = .06$), such that participants produced fewer words in the CG as opposed to the no-CG condition. However, this effect was modulated by age, as there were significant interactions between age group and common ground manipulation.

Individual contrasts revealed that only younger adults produced significantly more words and narrative events in the no-CG as opposed to the CG condition (β = 20.47, $SE = 3.65$, $t(34.13) = 5.60$, $p < .0001$ and $\beta = 1.17$, $SE = .43$, $t(34.10) = 2.75$, $p = .01$ respectively), whereas this difference was not significant for older adults (both p's > .05). Younger adults did not differ from older adults in the number of words and narrative events produced in the CG and no-CG conditions (all p's > .05).

To summarise, younger and older adults did not differ in the overall number of words and narrative events they produced. However, a significant effect of our common ground manipulation was only present in the younger adults, i.e. they used more words and more narrative events when talking about novel as opposed to shared story content.

3.2.2. Representational gesture rate and percentage of multimodal events
As for the speech-based measures, we fitted linear mixed effects models to evaluate the impact of age and common ground manipulation on gesture rate per 100 words and percentage of multimodal events. Note that

Table 1. Means (and SD) for the speech- and gesture-based dependent measures for each age group and condition.

	Younger		Older	
	CG	No-CG	CG	No-CG
Number of words	44.63 (21.35)	65.09 (19.93)	52.39 (12.45)	54.59 (12.47)
No. of narrative events	4.58 (2.07)	5.75 (1.07)	5.37 (.94)	5.35 (1.1)
Gesture frequency	2.02 (1.39)	4.78 (2.39)	3.01 (2.09)	2.74 (2.18)
Gestures/100 words	5.89 (4.5)	7.73 (3.94)	5.96 (3.94)	4.88 (3.86)
% Multimodal events	34.57 (21.56)	54.63 (19.95)	39.59 (24.13)	31.56 (23.64)

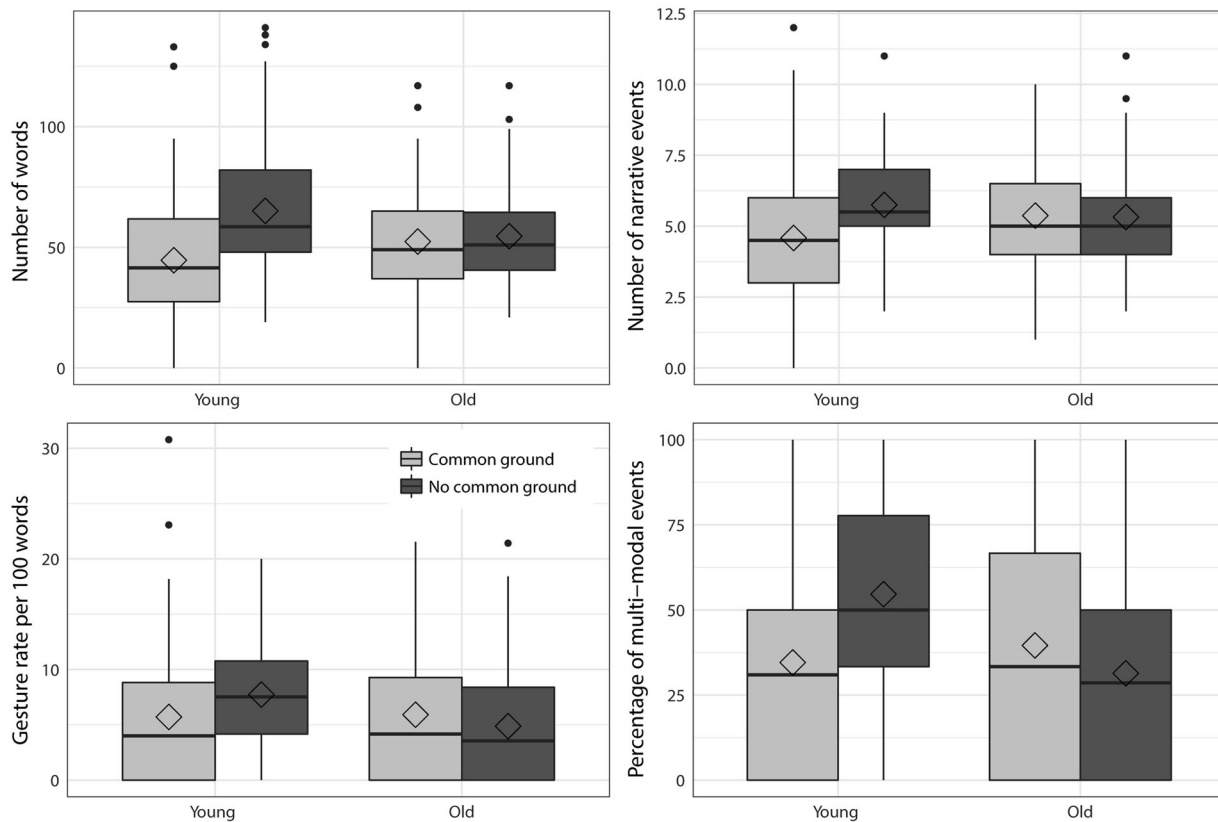Note: CG = common ground condition; no-CG = no common ground condition.

Figure 2. Distribution for the speech- and gesture-based dependent measures summarised by age group and condition (boxplots display six [story] * two [condition manipulation] data points per participant). The black line represents the median; the diamond represents the mean; the two hinges represent the 1st and 3rd quartile; the whiskers capture the largest and smallest observation but extend no further than 1.5 * IQR (data points outside 1.5 * IQR are represented by dots).

we did not include a by-participant random slope in the models predicting gesture rate, as this yielded a perfect correlation for the random effects. The final models are summarised in Table 3.

Again, we used likelihood ratio tests to compare nested models in order to obtain the simple main effect of age and common ground manipulation. This yielded no main effects for age or common ground manipulation for both measures (all p's > .05). However, the model summaries (Table 3) show that for the reference group of the younger adults, the effect of common ground was significant, such that participants gestured at a higher rate and produced more multimodal events in the no-CG as opposed to the CG condition. This

effect was modulated by age, as the significant interactions between age group and common ground manipulation show.

Individual contrasts confirmed that younger adults gestured at a significantly higher rate in the no-CG as opposed to the CG condition ($\beta = 1.86$, $SE = .54$, $t(343.70) = 3.46$, $p = .0006$), whereas older adults showed the reverse trend ($\beta = -1.04$, $SE = .54$, $t(342.69) = -1.96$, $p = .051$). Younger adults also produced significantly more multimodal events in the no-CG as compared to the CG condition ($\beta = 20.06$, $SE = 3.58$, $t(33.86) = 5.61$, $p < .0001$), whereas older adults showed the reverse pattern ($\beta = -8.18$, $SE = 3.59$, $t(34.36) = 2.28$, $p = .029$). Contrasts further revealed that younger and older

Table 2. Linear mixed-effects models for the effects of age and common ground manipulation on word count and number of narrative events mentioned. Age group = young and Condition = CG[a] are on the intercept. $N = 32$.[b]

	Words				Narrative events			
	β	SE	t	p	β	SE	t	p
Intercept	44.63	5.37	8.31	<.001	4.58	.53	8.68	<.001
Age group_old	7.76	6.03	1.29	.207	.78	.55	1.42	.17
Condition_no-CG[a]	20.47	3.57	5.79	<.001	1.17	.41	2.84	.008
Age group_old : Condition_no-CG	−18.26	5.00	−3.65	<.001	−.2	.58	−2.06	.048

[a]CG = common ground; no-CG = no common ground.
[b]Both models contain random intercepts for participants and items and by-participant random slopes for the common ground manipulation.

Table 3. Linear mixed-effects models for the effects of age and common ground manipulation on gesture rate per 100 words and percentage of multimodal events. Age group = young and Condition = CG[a] are on the intercept. $N = 32$.[b]

	Gesture rate per 100 words				Percentage multimodal events			
	β	SE	t	p	β	SE	t	p
Intercept	5.87	1.18	5.00	<.001	34.57	7.19	4.81	<.001
Age group$_{old}$	−.04	1.40	−.03	.98	5.25	7.91	.66	.51
Condition$_{no-CG}$[a]	1.86	.54	3.46	.0006	20.06	3.46	5.8	<.0001
Age group$_{old}$: Condition$_{no-CG}$	−2.9	.75	−3.84	.0001	−28.245	4.91	−5.76	<.0001

[a]CG = common ground; no-CG = no common ground.
[b]Both models contain random intercepts for participants and items. The model predicting the percentage of multimodal events includes by-participant random slopes for the common ground manipulation.

adults did not differ in the rate at which they gestured and in the percentage of multimodal events in the CG condition (both p's > .05). However, there was an age-related difference in the no-CG condition that approached significance for gesture rate ($\beta = 2.86$, $SE = 1.43$, $t(38.28) = 1.97$, $p = .053$) and was significant for percentage of multimodal events ($\beta = 22.99$, $SE = 7.75$, $t(32.95) = 2.97$, $p = .006$). That is, younger adults trended towards gesturing at a higher rate and produced a larger percentage of multimodal events than older adults in the no-CG condition.

To summarise, older and younger adults did not differ in their gesture rate and the percentage of multimodal events overall. However, we found different effects of our common ground manipulation for younger vs. older adults. While younger adults gestured at a higher rate and produced more multimodal events when narrating novel as opposed to known story content for their addressees, the opposite was the case for the older adults.

3.2.3. Explicit reference to common ground and addressee feedback

In addition to the main analyses reported above, we explored the influence of age and common ground manipulation on the frequency of speakers' explicit references to common ground, and on the frequency of addressee feedback. Explicit references to common ground can serve as an additional indicator of whether speakers were aware of their addressees' knowledge state. Controlling for addressee feedback is necessary in order to preclude the possibility that younger and older speakers' verbal and gestural behaviour differs due to differences in addressee behaviour. We fitted linear mixed effect models as described in Section 2.6. Note that we did not include by-participant random slopes in the models, as this yielded a perfect correlation for the random effects. Full model summaries are provided in appendix E.

3.2.3.1. Explicit reference to common ground. Per story, younger adults made on average .72 explicit references to common ground in the CG condition ($SD = .59$) and .03 ($SD = .09$) in the no-CG condition. Older adults

made on average .11 explicit references in the CG condition ($SD = .23$) and zero in the no-CG condition per story. With age group = young and common ground condition = CG mapped onto the intercept, the best fitting model contained effects for age ($\beta = −.41$, $SE = .1$, $t(50.8) = −3.87$, $p < .001$), common ground condition ($\beta = −.67$, $SE = .07$, $t(352) = −9.84$, $p < .001$), as well as the significant interaction term ($\beta = .51$, $SE = .1$, $t(352) = 5.33$, $p < .001$). Likelihood ratio tests showed that there was no overall main effect for age ($\chi^2(1) = 2.52$, $p = .11$), but only for common ground condition ($\chi^2(1) = 66.95$, $p < .001$). Thus, the two age groups did not differ significantly from each other in the overall number of explicit references to common ground they made. However, in the CG condition, younger adults produced significantly more explicit references than older adults. Hence, younger adults provided stronger indications of their awareness of the addressee's knowledge state than older adults.

3.2.3.2. Addressee feedback. We divided the amount of addressee feedback by the number of words per narration to account for differences in narration length. Both younger and older addressees produced numerically more feedback in the CG condition ($M_{young} = .07$, $SD = .04$; $M_{old} = .06$, $SD = .04$) than in the no-CG condition ($M_{young} = .05$, $SD = .03$; $M_{old} = .04$, $SD = .03$). The best fitting model contained a significant main effect for the common ground condition ($\beta = −.02$, $SE = .006$, $t(93.52) = −2.96$, $p = .004$), confirming the significance of this difference. The main effect for age approached significance ($\beta = −.02$, $SE = .009$, $t(31.01) = −1.99$, $p = .06$) such that older adults produced marginally less feedback overall than younger adults. Importantly, the interaction term of age and common ground condition did not improve the model fit, indicating that there was no systematic difference in the amount of feedback that younger and older addressees gave based on common ground condition. Hence, the observed age-related differences in common ground-based adaptation of speech and gesture reported above are unlikely to be due to differences in addressee feedback.

3.2.3.3. Effects of addressee feedback on verbal and gestural behaviour. We followed this analysis up by entering addressee feedback as a predictor into the previously reported models on word and narrative event count, gesture rate, and percentage of multimodal events, drawing on the subset of data for which feedback was coded. This was done in order to test whether accounting for feedback would modulate the effect of common ground for the younger adults that we established in the main analyses. We found that including feedback did not improve the models predicting word count or percentage of multimodal events. For narrative event count, there was no effect of the common ground manipulation in this subset, but addressee feedback had a significant effect such that more feedback predicted a reduction in narrative events ($\beta = -9.91$, $SE = 4.32$, $t(119.58) = -2.29$, $p = .02$). This effect appears to be driven more by the younger than by the older adults, but the interaction was not statistically significant. Finally, for gesture rate, feedback had a significant effect ($\beta = 36.16$, $SE = 9.9$, $t(113.41) = 3.65$, $p < .001$) such that more feedback predicted a higher gesture rate. However, crucially for our study, the effect of feedback did not influence the effect of common ground or its interaction with age. Overall then, taking addressee feedback into consideration did not eliminate the effect of the common ground manipulation observed in the speech- and gesture-based measures.

3.3. Effects of cognitive abilities on verbal recipient design and co-speech gesture

As we were also interested in the influence of cognitive abilities on verbal recipient design and on gesture production, we next turned to these factors. Particularly, we wanted to test whether the age-related differences in verbal and gestural behaviour could be attributed to age-related differences in cognitive functioning. As a group, younger adults significantly outperformed older adults on all cognitive tests with the exception of the semantic fluency task, see Table 4. For subsequent analyses, we standardised each task's scores by z-scoring. Correlations between cognitive predictors and dependent measures are reported in the supplementary materials, appendix F.

3.3.1. Words and narrative events

First, we tested our hypothesis that verbal WM and executive control influence verbal recipient design, by including these cognitive variables in the models predicting the number of words and narrative events

Table 4. Mean scores (and SD) per age group on cognitive tests, plus statistical comparisons (independent t-tests and Mann-Whitney tests where appropriate).

	Younger	Older	Test statistic
Verbal WM (Operation span task)	44.06 (8.39)	34.73 (8.92)	$t(28.5) = 2.99^{**}$
Semantic Fluency (Animal naming test)	31.5 (9.4)	27.5 (5.99)	$t(25.465) = 1.35$
Inhibitory control (Trail Making Test, TMT)[a,b]	14.5 (27)	21.5 (62)	$W = 65.5^{*}$
Visuo-spatial WM (Visual Patterns Test, VPT)[a]	13 (4)	10 (8)	$W = 187.5^{***}$
Visuo-sequential WM (Corsi Block Task, CBT)[a]	54 (42)	37.5 (34)	$W = 215^{***}$

$^{***}p < .001$.
$^{**}p < .01$.
$^{*}p < .05$.
[a]Owing to the non-normality of the data, the figures represent Median (and Range).
[b]Note that smaller numbers indicate better performance on this task.

produced per narration. We fitted linear mixed effects models, applying a backwards model-stripping procedure as described in Section 2.6. Both cognitive measures did not significantly improve the models fit for word and narrative event count, either as main effects or in interaction with age and common ground.

3.3.2. Representational gesture rate and percentage of multimodal events

Next, we tested the hypothesis that lower visuo-spatial or visuo-sequential WM, verbal WM, or semantic fluency are associated with an increase in gesticulation, and whether this affects the two age groups differently, by including these cognitive variables in the models predicting gesture rate and percentage of multimodal events. As for the previous analysis, none of the cognitive measures significantly contributed to models predicting either of the two gesture-based measures.

To summarise, we could not find any evidence that the observed age-related differences in cognitive abilities were predictive of the age-related differences in verbal and gestural behaviour as reported in Section 3.1. Furthermore, individual differences in cognitive abilities also could not predict verbal or gestural behaviour more generally, regardless of age group or common ground condition.

3.4. Summary of results

Overall, there were no age-related differences in how much participants spoke and gestured. However, in the presence of common ground, only younger adults used fewer words, fewer narrative events, gestured at a lower rate, and produced fewer multimodal events as compared to when there was no common ground. Older adults, on the other hand, did not adapt their

speech to common ground. Also, they gestured less in relation to speech when there was no common ground as compared to when there was common ground. Additionally, in spite of the absence of a general effect of age on gesticulation, in the no common ground condition, older adults produced fewer multimodal events than younger adults.

Furthermore, younger adults made more explicit references to common ground than older adults in the CG condition, overtly indicating their awareness of the mutually shared knowledge.

Crucially there were no age-related differences in the amount of addressee feedback, making this an unlikely explanation for differences in verbal and gestural behaviour between the two age groups. Additionally, we found that more addressee feedback was predictive of a reduction in narrative events and an increase in gesture rate, regardless of age and common ground.

Finally, although we found significant age-related differences in cognitive abilities, these did not explain the age-related differences in verbal and gestural adaptation to the common ground manipulation.

4. Discussion

The present study provides a first insight into how younger and older adults adapt their speech and co-speech gestures to an addressee's knowledge state when narrating short stories, and whether this verbal and gestural behaviour is affected by cognitive abilities. We found that younger and older adults did not differ in the number of words and narrative events they used, or in their representational gesture rate and percentage of multimodal utterances *overall*. However, adaptations of both speech and co-speech gestures based on mutually shared knowledge between speaker and addressee occurred only in the younger, but not in the older adults. Age-related differences in cognitive abilities did not predict these differences in behaviour, nor did addressee feedback behaviour modulate the observed effects. The individual results will be discussed in more detail below.

4.1. Effects of age and common ground on verbal recipient design

Overall, there were no age-related differences in the number of words and narrative events produced per narration. This suggests that younger and older adults were able to remember and reproduce approximately the same amount of information.

Crucially, with respect to our hypotheses concerning the adaptation to mutually shared knowledge, we found that younger adults showed a stronger effect of common ground on speech than older adults. That is, younger adults used fewer words and narrative events to narrate known story content compared to novel content. This is in line with previous findings for younger adults in similar narration tasks (e.g. Galati & Brennan, 2010; Holler & Wilkin, 2009). It shows that the more knowledge speakers assume to be mutually shared, the less verbal information is conveyed. The fact that younger adults frequently referred to common ground explicitly when relating familiar content, e.g. by stating "you've already seen the first half so I'll go through it quickly" similarly shows that they were aware of their addressee's knowledge state.

Furthermore, we found indications that younger adults were not only aware of the addressee's knowledge state as a function of the common ground manipulation, but that they were also sensitive to the addressees' verbal and visual backchannel signals. In the present study, addressees provided more backchannel signals in the presence of shared knowledge, which in turn predicted a decrease in narrative events. Note that previous research, for example Galati and Brennan (2014), found shared knowledge to be associated with a *reduction* in addressee feedback. However, in their task, addressees listened to the retelling of the same story twice, which may have caused the addressee to be less involved and less responsive during the second retelling. In the present task, on the other hand, common ground was manipulated within each story, and even though the addressee had seen part of the story already (thus constituting common ground), they had not spoken about it or heard the speaker narrate the content previously. The purpose of the increased feedback during common ground content may have been to actively indicate to the speaker that the addressee recognised the content and to affirm that it was mutually shared. This additional finding highlights the important influence of the addressee's behaviour on the speaker's language use (Bavelas, Coates, & Johnson, 2000).

In contrast to the younger adults, older adults hardly differed in the number of words and narrative events they used to talk about known versus novel story content. Also, they made fewer explicit references to common ground in the CG condition than the younger adults, meaning they were less likely to verbally mark mutually shared knowledge for their addressee. We had expected this effect of age on verbal recipient design based on earlier studies showing that older adults are less good at establishing conversational common ground than younger adults (Horton & Spieler, 2007; Hupet et al., 1993; Lysander & Horton, 2012). In principle, two kinds of explanations for these

behavioural differences are conceivable: Older adults may not be able to engage in common ground-related recipient design as we induced it here due to age-related cognitive limitations, but they may also respond differently to the communicative situation than younger adults due to other factors. We explore both options in the following paragraphs.

Based on previous research (e.g. Healey & Grossman, 2016; Horton & Spieler, 2007; Hupet et al., 1993; Long et al., 2018; Wardlow, 2013), we had hypothesised that verbal WM and executive control influence the ability to engage in recipient design. Deficits in verbal WM may limit the extent to which speakers can focus their resources on considering which information is or is not mutually shared when designing their utterances, and on adapting the utterances accordingly. Deficits in executive control may be related to difficulties in inhibiting the speaker's own, egocentric perspective or suppressing irrelevant information, both of which are necessary for recipient design to occur (Brennan et al., 2010; Keysar et al., 1998). As older adults in the present study had significantly lower verbal WM and executive control than younger adults, this might have contributed to their lack of verbal addressee-based adaptations. However, we could find no support for this hypothesis, as neither of the two cognitive abilities could predict differences in verbal behaviour. Of course, the small sample size employed in the present study limits our ability to interpret this absence of an effect. Additionally, it is possible that the particular tasks that we used to assess verbal WM and executive control do not tap into the actual processes involved in verbal recipient design.

Nevertheless, as we found no support for the cognitive account, it is necessary to consider alternative explanations for the older adults' behaviour. Previous research suggests that age-related differences in communicative behaviour may also be related to differences in social or pragmatic goals (e.g. Adams, Smith, Pasupathi, & Vitolo, 2002; Horton & Spieler, 2007; James, Burke, Austin, & Hulme, 1998). For example, older adults may have had the primary goal of narrating the story "well", therefore giving equal weight to both known and unknown story content in their narrations, whereas younger adults may have focused primarily on being concise and providing information that the addressee did not yet have (see e.g. James et al., 1998, who found that older adults are judged to be better at story telling than younger adults). Another possibility is that older adults may have wished to demonstrate that they remembered all parts of the story well and thus could perform well on the story telling task in general, as beliefs about age-related memory decline are widespread, also among older adults (e.g. Lineweaver &

Hertzog, 1998). This desire may have overruled any common ground-based adaptations of their speech. Finally, the fact that older speakers always narrated the stories for older addressees may also have influenced their verbal behaviour. Potentially, older speakers may have thought that their addressees could not remember all of the mutually shared content due to memory limitations and therefore refrained from reducing verbal content in the CG condition. Previous research shows that older adults adapt their verbal utterances based on addressee characteristics such as age (Adams et al., 2002; Keller-Cohen, 2015) or mental retardation (Gould & Shaleen, 1999). Future research could address this possibility by testing mixed age pairs in order to see whether older speakers adapt their speech differently for younger addressees (and younger speakers differently for older addressees).

4.2. Effects of age and common ground on multimodal recipient design

As in verbal recipient design, younger and older adults also differed in how they adapted their representational gesture use to their addressee's knowledge state. Younger adults gestured at a higher rate and produced more multimodal utterances when communicating novel as opposed to mutually shared content, similar to the findings by Jacobs and Garnham (2007) (but see Holler & Bavelas, 2017, for a summary of the range of different effects common ground can have on gesture). This reduction in multimodal information appears to be a direct effect of speakers adapting to the speakers knowledge state, providing the addressee with a comprehensive verbal and visual representation of the novel part of the story, and a verbally and especially visually reduced representation when talking about familiar content. It is additionally interesting to note that even though the CG condition was associated with an increased amount of addressee feedback, which in turn predicted an increase in gesture rate, this did not eliminate the effect of common ground on gesture rate. Taken together, these findings illustrate that younger adults could flexibly adapt not only their speech, but also their gestures to the communicative requirements of the situation (Kendon, 1985, 2004).

For older adults, we observed a pattern opposite to that of the younger adults: They tended to gesture at a lower rate and produced fewer multimodal events when talking about novel content, both compared to their own production for shared content and compared to younger adults' production for novel content. We had expected that older adults would show a smaller common ground effect on gesture production than

younger adults, based on our predictions for verbal audience design and on the hypothesis that speech and gesture function as one integrated system (Kita & Özyürek, 2003; McNeill, 1992). Therefore, it is surprising that we found common ground to influence older adults' gesture production in this opposite direction, also considering the absence of an effect on their speech. One possible explanation for this finding is that relating novel story content required more cognitive effort than relating mutually shared content. Older adults may have been aware that they should provide more information, yet failed to do so verbally, potentially due to memory limitations. This presumed increase in cognitive load associated with the novel content condition may have led to a reduction in multimodal events, as gestures produced primarily for the benefit of an addressee may actually be cognitively costly to the speaker (Mol, Krahmer, Maes, & Swerts, 2009). However, this speculation rests on the assumptions that the gestures produced during this narrative task were primarily intended to illustrate the story for the addressee, and that older adults failed to engage in verbal recipient design due to cognitive limitations, which we could not find evidence for (but due to our sample size, this needs to be followed up with future research, see previous section).

The present study shows that younger and older adults differ in how they adapt speech and gestures to the common ground shared with an addressee. Ultimately, it seems likely this behaviour is determined by a combination of cognitive and social or pragmatic factors (see also Horton & Spieler, 2007). Based on the design of the present study, however, we cannot tease the individual contributions of these two factors apart. First of all, our ability to interpret the absence of cognitive effects is limited by the sample used in our study. Additionally, it might be that the cognitive tests we used did not capture the abilities that are involved in recipient design. Also, in this study, we did not assess what the speakers' goals and intentions were, and whether there were systematic differences between younger and older adults with respect to this. Thus, while the present study provides clear evidence of age-related differences in multimodal recipient design, we currently can only provide some preliminary ideas on what causes these. Future studies are needed which include larger samples and a broader range of interactive tasks and measures.

4.3. General effects of age and cognitive abilities on gesticulation

Despite the age-related difference in how speakers adapted multimodally to common ground, younger and older adults did not differ in terms of representational gesture rate or the percentage of multimodal narrative events they produced overall. The analyses of gesture frequency and gesture rate per narrative event yielded identical results (see appendix C). Thus, our results are not in line with the earlier finding that older adults gesture less than younger adults overall (Cohen & Borsoi, 1996; Feyereisen & Havard, 1999; Theocharopoulou et al., 2015). We would like to propose that the difference in findings is due to the communicative paradigm we employed. Whereas participants in the previous studies on gesture production in ageing either had no addressee at all or an experimenter-addressee, in the present study we used co-present, non-confederate addressees. Previous research with younger adults indicates that the presence of a visible, attentive addressee increases the relative frequency of representational gestures (e.g. Jacobs & Garnham, 2007; Kuhlen et al., 2012). In the current study, older and younger addressees differed only marginally with respect to the amount of feedback they gave, and in both age groups, an increase in addressee feedback was predictive of an increase in gesture rate. Certainly, this should be considered gestural recipient design (as has been argued for effects of addressee feedback on gesture form, Holler & Wilkin, 2011), albeit not the kind of common ground-based recipient design that we intended to investigate through our experimental manipulation.

As younger and older adults did not differ in how much they gestured in relation to speech, there was also no support for the hypothesis that older adults produce more gestures than younger adults in order to compensate for their relative deficit in cognitive abilities, based on accounts of gestures being cognitively beneficial (Chu et al., 2014; Gillespie et al., 2014; Goldin-Meadow et al., 2001; Hostetter & Alibali, 2008; Wagner Cook et al., 2012). Additionally, we found no associations between verbal WM, visuo-sequential WM, or semantic fluency and gesticulation as we assessed them. The field would benefit from a broader investigation of the relationship between cognitive abilities and gesticulation in older adults, using a wider range of gesture elicitation tasks and of cognitive measures (as well as the large sample required for investigating individual differences), similar to previous work with younger adults (Chu et al., 2014; Gillespie et al., 2014).

Nevertheless, the fact that in the *absence* of shared knowledge, older adults gestured *less* than younger adults might be an indication that older adults reduce their gesture production in contexts that induce a higher cognitive load. Future work is needed to test this possibility.

5. Conclusion

The present study offers a first glimpse of how ageing affects multimodal recipient design in the context of common ground. In an interactive setting, older adults spoke as much and gestured as frequently in relation to speech as younger adults, and were similarly sensitive to addressee feedback on the whole. However, only younger adults adapted both their speech and gesture use for their addressee based on the mutually shared knowledge established at the outset of the interaction, such that they provided relatively less multimodal information when there was shared knowledge, and relatively more multimodal information when there was not. Older adults did not adapt their speech based on the addressee's knowledge state and conveyed less, rather than more, multimodal information in the absence of shared knowledge.

If we take younger adults' behaviour in this task as the baseline against which to compare the older adults, we must conclude that older adults failed to engage in successful common ground-based recipient design. That is, while younger adults flexibly adapted both their speech and their gestures to the communicative requirements of the situation, older adults appeared less flexible in the way they drew on their different communicative modalities. We attribute these behavioural differences at least in part to age-related changes in social or pragmatic goals, as they were not reliably predicted by the significant age-related differences in cognitive abilities. Yet, we acknowledge our limited sample size and do not want to exclude the possibility of a cognitive explanation for some findings, such as that older adults produced fewer multimodal events in the absence of shared knowledge.

Our findings raise the question of whether the age-related differences in verbal and gestural patterns found here persist in other types of communicative tasks where common ground builds up incrementally, and whether they have an impact on how older adults are comprehended by others, both young and old.

Notes

1. With the exception of Lysander and Horton (2012), who take eye-gaze into consideration.
2. Note, however, that the proportional relation of speech and gesture, expressed in measures of relative frequency, such as gesture rate (e.g. the number of gestures per 100 words), may vary considerably, depending on whether the two modalities are reduced to the same extent, or whether the reduction in one modality is stronger than in the other; for a detailed discussion of this issue see Holler & Bavelas, 2017).
3. Note that the questions did not target common ground vs. no common ground information systematically and can therefore unfortunately not give any insights into the addressee's information uptake as based on the speaker's narration.
4. "Re-enactments", i.e. movements of the body that represented specific actions of the stories' characters, were also coded as iconic gestures, even if they did not include manual movements.

ORCID

Louise Schubotz 🄳 http://orcid.org/0000-0002-1311-8855

References

Abrams, L., & Farrell, M. T. (2011). Language processing in normal aging. In J. Guendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders* (pp. 49–73). New York, NY: Psychology Press.

Adams, C., Smith, M. C., Pasupathi, M., & Vitolo, L. (2002). Social context effects on story recall in older and younger women: Does the listener make a difference? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57B*(1), P28–P40.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*, 169–188.

Bates, D., Maechler, M., & Bolker, B. (2017). lme4: Linear mixed-effects models using "Eigen" and S4. R package version 1.1–13. Retrieved from https://cran.r-project.org/web/packages/lme4/

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning. An integrated model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*(2), 163–194.

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, *21*(4), 394–405.

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, 15, 469–489.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, 79(6), 941–952.

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495–520.

Bavelas, J. B., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, 2(1), 1–17.

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech*, 44, 123–147.

Brennan, S., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialogue: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, 53, 301–344.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, 16(5), 893–900.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*, 30, 542–579.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic Research*, 4(1), 75–87.

Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, 47, 14–27.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: The role of cognitive ability and empathy. *Journal of Experimental Psychology: General*, 143(2), 694–709.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Murphy, G. L. (1983). Audience design in meaning and reference. In J. F. Leny & W. Kintsch (Eds.), *Language and comprehension* (pp. 287–299). Amsterdam: North-Holland Publishing.

Cohen, R. L., & Borsoi, D. (1996). The role of gestures in description-communication: A cross-sectional study of aging. *Journal of Nonverbal Behavior*, 20(1), 45–63.

de Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science*, 4(2), 232–248.

ELAN (Version 4.9.4) [Computer software]. (2016, May 19). Nijmegen: Max Planck institute for psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Feyereisen, P., & Havard, I. (1999). Mental imagery and production of hand gestures while speaking in younger and older adults. *Journal of Nonverbal Behavior*, 23(2), 153–171.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, 62, 378–391.

Galati, A., & Brennan, S. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, 62, 35–51.

Galati, A., & Brennan, S. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, 29(4), 435–451.

Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture*, 4(2), 157–195.

Gillespie, M., James, A. N., Federmeier, K. D., & Watson, D. G. (2014). Verbal working memory predicts co-speech gesture: Evidence from individual differences. *Cognition*, 132, 174–180.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, 12, 516–522.

Gould, O., & Shaleen, L. (1999). Collaboration with diverse partners: How older women adapt their speech. *Journal of Language and Social Psychology*, 18(4), 395–418.

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). New York, NY: Oxford University Press.

Hasher, L., & Zacks, R. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 2, pp. 193–225). San Diego, CA: Academic Press.

Healey, H., & Grossman, M. (2016). Social coordination in older adulthood: A dual-process model. *Experimental Aging Research*, 42(1), 112–127.

Hilliard, C., & Cook, S. W. (2016). Bridging gaps in common ground: Speakers design their gestures for their listeners. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 91–103.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (2015). Reduction in gesture during the production of repeated references. *Journal of Memory and Language*, 79–80, 1–17.

Holler, J., & Bavelas, J. (2017). Multi-modal communication of common ground. In R. Breckinridge Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 213–240). Amsterdam: John Benjamins.

Holler, J., & Beattie, G. (2003a). How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process? *Semiotica*, 146, 81–116.

Holler, J., & Beattie, G. (2003b). Pragmatic aspects of representational gestures: Do speakers use them to clarify verbal ambiguity for the listener? *Gesture*, 3, 127–154.

Holler, J., & Stevens, R. (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26(1), 4–27.

Holler, J., & Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24(2), 267–289.

Holler, J., & Wilkin, K. (2011). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics*, 43, 3522–3536.

Horton, W. S., & Gerrig, R. (2005). The impact of memory demands on audience design during language production. *Cognition*, 96, 127–142.

Horton, W. S., & Spieler, D. H. (2007). Age-related differences in communication and audience design. *Psychology and Aging*, *22*(2), 281–290.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514.

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007). I see it in my hands' eye: Representational gestures reflect conceptual demands. *Language and Cognitive Processes*, *22*(3), 313–336.

Hupet, M., Chantraine, Y., & Nef, F. (1993). References in conversation between young and old normal adults. *Psychology and Aging*, *8*(3), 339–346.

Isaacs, E. A., & Clark, H. H. (1987). References in conversations between experts and novices. *Journal of Experimental Psychology: General*, *116*, 26–37.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language*, *56*, 291–303.

James, L. E., Burke, D. M., Austin, A., & Hulme, E. (1998). Production and perception of "verbosity" in younger and older adults. *Psychology and Aging*, *13*, 355–367.

Keller-Cohen, D. (2015). Audience design and social relations in aging. *Research on Aging*, *37*(7), 741–762.

Kendon, A. (1985). Some uses of gesture. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 215–234). Norwood, MA: Ablex.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric bias of language use: Insights from a processing approach. *Current Directions in Psychological Science*, *7*(2), 46–49.

Kita, S., & Davies, T. S. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language and Cognitive Processes*, *24*(5), 761–775.

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction. GW 1997. Lecture notes in computer science* (Vol. 1371, pp. 23–35). Berlin: Springer.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*, 54–72.

Kuhlen, A. K., Galati, A., & Brennan, S. E. (2012). Gesturing integrates top-down and bottom-up information: Joint effects of speakers' expectations and addressees' feedback. *Language and Cognition*, *4*, 17–41.

Kuznetsova, A., Brockhoff, P. B., & Bojesen Christensen, R. H. (2016). lmerTest: Tests in linear mixed effects models. R package version 2.0-33. Retrieved from http://cran.r-project.org/web/packages/lmerTest/

Lenth, R. (2018). emmeans: Estimated marginal means, aka least-squares means. R package version 1.1.3. Retrieved from https://cran.r-project.org/web/packages/emmeans/

Lineweaver, T. T., & Hertzog, C. (1998). Adults' efficacy and control beliefs regarding memory and aging: Separating general from personal beliefs. *Aging, Neuropsychology, and Cognition*, *5*, 264–296.

Long, M. R., Horton, W. S., Rohde, H., & Sorace, A. (2018). Individual differences in switching and inhibition predict perspective-taking across the lifespan. *Cognition*, *170*, 25–30.

Lysander, K., & Horton, W. S. (2012). Conversational grounding in younger and older adults: The effect of partner visibility and referent abstractness in task-oriented dialogue. *Discourse Processes*, *49*(1), 29–60.

McNeill, D. (1992). *Hand and mind*. Chicago, IL: The Chicago University Press.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, *22*(4), 473–500.

Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2009). Communicative gestures and memory load. In N. Taatgen (Ed.), *Proceedings of the 31st annual conference of the Cognitive Science Society (CogSci 2009)* (pp. 1569–1574). Amsterdam: Cognitive Science Society.

Mol, L., Krahmer, E., Maes, A., & Swerts, M. (2011). Seeing and being seen: The effects on gesture production. *Journal of Computer-Mediated Communication*, *17*(1), 77–100.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, *21*, 238–290.

Özyürek, A. (2002). Do speakers design their co-speech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, *46*(4), 688–704.

Özyürek, A. (2017). Function and processing of gesture in the context of language. In R. B. Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 39–58). Amsterdam: John Benjamins Publishing.

Parrill, F. (2010). The hands are part of the package: Gesture, common ground and information packaging. In S. Rice & J. Newman (Eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 285–302). Stanford, CA: CSLI Publications.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science*, *7*, 226–231.

R Development Core Team. (2015). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from http://www.R-project.org

Rowbotham, S. J., Holler, J., Wearden, A., & Lloyd, D. M. (2016). I see how you feel: Recipients obtain additional information from speakers' gestures about pain. *Patient Education and Counseling*, *99*(8), 1333–1342.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum.

So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, *33*, 115–125.

Theocharopoulou, F., Cocks, N., Pring, T., & Dipper, L. T. (2015). TOT phenomena: Gesture production in younger and older adults. *Psychology and Aging*, *30*(2), 245–252.

Thornton, R., & Light, L. L. (2006). Language comprehension and production in normal aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 261–287). San Diego, CA: Elsevier Academic Press.

Verhaeghen, P. (2011). Aging and executive control: Reports of a demise greatly exaggerated. *Current Directions in Psychological Science*, *20*(3), 174–180.

Wagner Cook, S., Yip, T. K., & Goldin-Meadow, S. (2012). Gestures, but not meaningless movements, lighten working memory load when explaining math. *Language and Cognitive Processes*, *27*(4), 594–610.

Wardlow, L. (2013). Individual differences in speaker's perspective taking: The roles of executive control and working memory. *Psychonomic Bulletin & Review*, *20*(4), 766–772.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multi-modality research. In *Proceedings of LREC 2006, fifth international conference on language resources and evaluation*. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from http://tla.mpi.nl/tools/tla-tools/elan/