Age-related differences in multi-modal audience design: Young, but not old speakers, adapt speech and gestures to their addressee's knowledge

Louise Schubotz[1,2], Judith Holler[1], Aslı Özyurek[1,3]

[1] Max Planck Institute for Psycholinguistics, Nijmegen, the Netherlands
[2] Max Planck International Research Network on Aging
[3] Radboud University Nijmegen, Nijmegen, the Netherlands
Louise.Schubotz@mpi.nl, Judith.Holler@mpi.nl, Asli.Ozyurek@mpi.nl

Abstract

Speakers can adapt their speech and co-speech gestures for addressees. Here, we investigate whether this ability is modulated by age. Younger and older adults participated in a comic narration task in which one participant (the *speaker*) narrated six short comic stories to another participant (the *addressee*). One half of each story was known to both participants, the other half only to the speaker. Younger but not older speakers used more words and gestures when narrating novel story content as opposed to known content. We discuss cognitive and pragmatic explanations of these findings and relate them to theories of gesture production.
Index Terms: co-speech gesture, aging, audience design, common ground

1. Introduction

The assumption that communicative competence deteriorates with advancing age, due to cognitive or biological decline, is widespread both among younger and older adults, and also in the scientific community, see [1]. Yet, to date little is known about the every-day language use of older adults in personal, face-to-face interactions. Importantly, face-to-face interaction has two inherent features which are frequently overlooked in laboratory investigations of language production: face-to-face language use is a *multi-modal* activity in the sense that it comprises communicative channels beyond the mere speech signal, such as manual co-speech gestures; and it is produced for and targeted at an addressee, shaped by a process called *audience design* [2]. Previous research with younger adults shows that speakers adapt both their speech and their co-speech gestures to an addressee's perceived communicative needs (e.g. [3], [4]). Likewise, addressees are able to perceive, integrate, and interpret information that is presented in these two modalities (e.g. [5], [6]). It is currently unclear whether, and if so, how older adults use these multiple communicative channels when designing utterances for others. However, the findings of previous research presented in the following paragraphs suggest that language produced by older adults may differ systematically from that of younger adults.

1.1 Audience design in speech and co-speech gesture

In younger adults, effects of audience design are frequently investigated by manipulating the amount of conversational *common ground*, i.e. knowledge, beliefs and assumptions that conversational partners believe to be mutually shared and that allows for the appropriate adaptation of utterances [7]. Generally, the more information is shared between interactants, the less needs to be put into words, characterized e.g. by shorter utterances, less complex syntax, or less informational content (e.g. [8], [9]).

Older adults' ability to engage in audience design based on mutually shared knowledge has been addressed in a number of studies employing a director-matcher card game:

Participants are required to establish mutual reference to a limited set of objects over the course of several trials, thereby building up *local* or *emerging* common ground ([10], [11], [12]). Younger adults' interactions become increasingly more efficient, indicated by shorter utterances and task-completion times. Older adults are also able to establish common reference, however, they are less efficient, indicated by longer utterances, longer task-completion times, more errors produced, and more idiosyncrasy when compared to younger adults. [10] suggest that this may be due to age-related cognitive limitations, specifically difficulties in retrieving partner-specific information from memory.

Findings from studies manipulating global addressee characteristics such as age ([13], [14]) or mental retardation [15] suggest that older adults are able to adapt their speech based on these *a priori* or *global* aspects of common ground and do not differ significantly from younger adults. Arguably, memory demands are much lower in these paradigms as opposed to the director-matcher tasks, which may account for older adults' better performance here.

One shortcoming of nearly all of these studies is that they ignore the multi-modal character of face-to-face language use.[1] Yet, information conveyed in the visual domain is essential to face-to-face interaction. Especially *representational* co-speech gestures, i.e. gestures that depict information imagistically, contribute to the semantic content of a message and are sensitive to social context variables. For example, speakers can produce representational gestures to clarify verbal ambiguity for their addressee [16], and representational gesture rate (i.e. the number of gestures produced per 100 words) is sensitive to visibility between speaker and addressee (e.g. [17], [18]), as well as to dialogue and addressee feedback (e.g. [19], [20]), suggesting that speakers take their addressee's communicative needs and abilities into account when designing multi-modal utterances.

Studies investigating the effect of common ground on gesture production often obtain effects that parallel the findings for speech. For example [4] used a cartoon narration task in which a speaker narrated one story three times, first to a naïve addressee, and then again to either the same addressee (common ground) or a to different addressee (no common ground). In second narrations, speakers produced significantly fewer, smaller, and less precise representational gestures for same addressees than for different addressees. Using a similar paradigm, [21], Exp. 1, also found that speakers produce fewer representational gestures when narrating the same comic story three times to either the same addressee or an addressee who could also see the story (common ground) as opposed to addressees who were not familiar with the story (no common ground), again indicated by a decrease in gesture rate. Similar effects for common ground on gesture rate or quality have been obtained by e.g. [22], [23], and [24] amongst others. However, others have found no effects of common ground on

---

[1] With the exception of [12], who take eye-gaze into consideration.

gesture rate ([25], [26], [27]), or even opposite effects, such that participants gesture more in relation to speech when common ground was present ([28], [29]).

1.2 Co-speech gesture production in older adults

Research on co-speech gesture production in older adults is to date limited to the two studies summarized in the following. [30] asked younger and older women to describe four physically co-present objects to a video camera and found that older women produced representational co-speech gestures at a significantly lower frequency than younger women. The authors explain this significant age difference by referring to the idea that older adults are less involved with mental imagery during speaking. This assumption was explicitly tested by [31]. In their study, younger and older participants responded to questions thought to evoke mental imagery (visual and motor). Again, older adults produced representational gestures at a significantly lower rate (computed as number of gestures per five-second time window) than the younger experimental group, but the imagery content of their speech was comparable to that of younger adults. Hence, a lack of mental imagery seems to be an unlikely explanation for older adults' decreased use of representational gestures. Rather, [31] argue that older adults prefer simpler gestural forms when facing the task of speaking and gesturing concurrently, possibly due to cognitive limitations, although the authors do not elaborate on this issue further.

However, neither of these studies used an interactive paradigm in which an addressee's knowledge state must be taken into account for successful communication. It is therefore unclear how older adults use speech and co-speech gestures concurrently in these types of situations.

1.3 The present study

The main aim of our research was to find out whether, and if so, how older adults adapt their speech and their gestures to mutually shared knowledge between speaker and addressee. In order to investigate this, we designed a comic narration task in which a primary participant (the *speaker*) narrates six short comic strips to a secondary participant (the *addressee*) who would then answer a question based on this narration. Common ground was manipulated by showing both participants one half of each strip (either the first or the second half) at the beginning of each trial. Only the speaker would subsequently see the full story, meaning that one half of the story content was mutually known to both participants (common ground or CG), while the other half was only known to the speaker (no common ground or no-CG).

Our two main dependent measures were number of words and the gesture rate per condition/narration. 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 when narrating shared story content and more words when narrating novel content (e.g. [3]). Based on the results obtained by [11], [12], and [13] we expected this effect to be smaller in older adults. Given the mixed findings in the gesture literature on common ground, gesture rates could decrease (e.g. [4], [21]), stay constant (e.g. [25]), or increase (e.g. [28]) with an increase in common ground. In analogy to our predictions for an age effect on speech, we did hypothesize that if there is an effect for younger adults on gesture production, this should be smaller for older adults. Also, we expected older adults to produce fewer representational gestures overall, in line with [30] and [31].

2. Method

2.1 Participants

64 participants took part in the study, 32 younger adults (16 women) between 21 and 30 years old (M = 24.31 years, SD = 2.91 years) and 32 older adults (16 women) between 64 and 73 years old (M = 67.69 years, SD = 2.43 years). All 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. All participants were native Dutch speakers with self-reported normal or corrected-to-normal vision and hearing.

2.3 Design

We employed a 2 x 2 design, with the between-participant variable age (young vs. old), and the within-participant variable common ground (CG vs. no-CG).

2.3 Materials

Seven black-and-white comic strips from the series "Vater und Sohn" were used to elicit narratives. Each strip consisted of a self-contained story, either four or six frames long, and centered around a father and a son and their activities.

2.4 Procedure

Participants came to the lab in pairs. We tested same age and same sex pairs only. The role of speaker and addressee were pre-assigned randomly and kept constant across the entire experiment. Upon arrival, speaker and 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 full frontal view of the speaker, and the other one positioned such that it captured both speaker and addressee (see Figure 1 for a still from the second camera). Sound was recorded with an additional microphone suspended from the ceiling over the table and connected to the speaker camera.

Figure 1. *Speaker (left) and addressee (right) seated at the table.*

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 necessary information to successfully participate in the experiment. Signed consent was acquired from all participants. Before the start of the actual experimental sessions, participants played a warm-up game to get familiar with each other as well as the experimental set-up. Following the warm-up game, the experiment continued with one of two experimental tasks: a comic narration task (present experiment) and a building block task (reported elsewhere), the order of the two tasks was counterbalanced across dyads.

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 (counterbalanced across the experiment) and were instructed to look at it together for 10 seconds without talking. Eight experimental lists determined the order of story presentation. Each list was tested four times, once for each age/sex pair. 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. Once the speaker signaled that she had understood and memorized the story, both drawings and screen were removed again and the speaker narrated the entire story to the addressee. The speaker 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 content of the story in writing. Depending on the dyad, the task took about 20 to 30 minutes.

2.5 Transcription and coding

All recordings from the two cameras were synchronized and subsequently segmented into trials. Transcription of speech and annotation of gestures was done in Elan [32]. For all segments, the speaker's initial narration was identified. All analyses reported here are based on these initial narrations only, discarding repetitions or clarifications elicited by the addressee. Speech from the speaker was transcribed verbatim, including disfluencies such as filled pauses and word fragments. However, these disfluencies are excluded from the word counts presented in the results section. We also distinguished between speech belonging to the narrative proper, i.e. relating story content, and non-narrative speech such as statements about the task or comments relating to the speaker or the addressee. 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".

For the gesture coding, we first identified all co-speech gestures produced by the speaker and accompanying narrative speech, disregarding irrelevant movements that were not gestures as well as gestures accompanying non-narrative speech. We then categorized these gestures according to their function. Globally, we distinguished between *representational* and *non-representational* gestures (see [17]).

For our purposes, representational gestures include *iconic* gestures, which depict shape or size of concrete referents or represent specific physical movements or actions; [2] furthermore *metaphoric* gestures, which relate to speech in a more abstract manner; and finally pointing gestures or *deictics*. We distinguished between *concrete deictics*, i.e. finger points to a physically co-present referent, and *abstract deictics*, i.e. finger points to a specific location in space, e.g. that of a story character.

All other gestures were considered non-representational and include what are frequently called *beat* gestures, i.e. biphasic movements of the hand e.g. to add emphasis, furthermore *interactive* gestures relating to the structuring of the conversation. As non-representational co-speech gestures occurred very infrequently and were not the primary interest of the present study, we decided to not investigate them further here.

A second coder blind to the experimental hypotheses

coded 10% of the trials. Inter-rater agreement on stroke identification was 92.3%. Inter-rater agreement on gesture categorization was 97.9%, Cohen's Kappa = .949.

To normalize for differences in speech rate, we computed the gesture rate as the number of gestures per 100 words.

3. Results

Table 1 lists the mean values and standard deviations for the various measures by age group and condition. We first computed an average value per participant and condition, and then computed means and standard deviations based on these averages. Like words and gesture rate, explicit reference to common ground was computed as the average number of explicit references made across trials per condition. For all analyses, we performed a 2 (age: old vs. young) x 2 (common ground: CG vs. no-CG) ANOVA as well as pair-wise comparisons using t-tests or, where applicable, Wilcoxon tests, in combination with Bonferroni corrections. All p-values are two-tailed unless clearly stated otherwise.

Table 1. *Means and standard deviations of various measures per trial for age groups and conditions.*

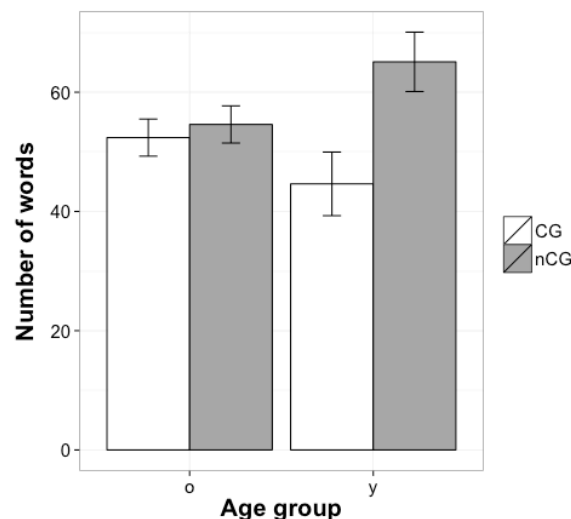|	CG		no-CG	
	Young	**Old**	**Young**	**Old**
Narrative words	44.13	52.39	65.59	54.59
	(21.37)	(12.45)	(20.23)	(12.47)
Representational gesture rate	5.67	5.95	7.62	4.86
	(4.28)	(4.01)	(3.92)	(3.72)
Non-rep. gesture rate	1.91	1.25	1.84	1.08
	(1.93)	(1.12)	(1.74)	(1.02)
CG reference	.72	.11	.03	0
	(.59)	(.23)	(.09)	

3.1 Speech

3.1.1 Narrative words

Figure 2. *Average number of narrative words per age group and common ground condition. Error bars represent the SE.*

Figure 2 shows the average number of narrative words per age group and common ground condition. The results of the ANOVA yielded a significant interaction of age by common ground, $F(1,60) = 5.043, p = .028$. The main effect of common ground was also significant, with participants producing more words in the no-CG condition than in the CG condition, $F(1,60) = 7.621, p = .008$, but the main effect of age was not, $F(1,60) = .102, p = .75$. To explore this

---

[2] Note that "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.

interaction further, we calculated four pairwise comparisons, adopting a Bonferroni-corrected alpha level of .0125. First, we compared the number of narrative words per common ground condition within age groups using one-tailed paired t-tests. As predicted, young adults used significantly fewer words to describe CG content vs. no-CG content, $t(15) = 4.852$, $p < .001$. For old adults, the difference between CG and no-CG was not significant, $t(15) = .746$, $p = .23$. We then compared the average number of narrative words used per common ground condition across age groups. Young and old adults did not differ significantly in the number of words used to describe CG content, $t(24.145) = 1.256$, $p = .221$, or no-CG content, $t(25.186) = 1.786$, $p = .086$.

3.1.2 Explicit reference to common ground

As the data were not normally distributed, we used Wilcoxon rank sum and signed rank tests to explore the differences between age groups and conditions in the explicit reference to common ground. Young adults made significantly more explicit references to common ground in the CG condition than in the no-CG condition, $Z = 3.189$, $p = .001$. Also, young adults used significantly more explicit references than old adults across conditions, $Z = 2.903$, $p = .003$. None of the other pairwise comparisons were significant (all p's > .05).

3.2 Representational co-speech gesture

Figure 3 shows the average representational gesture rate per age group and common ground condition. The ANOVA revealed no significant main effects of common ground, $F(1,60) = .19$, $p = .66$, or age, $F(1,60) = .1.554$, $p = .21$, and no significant interaction, $F(1,60) = 2.342$, $p = .13$. Since we tested specific hypotheses, we computed four pairwise comparisons, adopting an alpha level of .0125 throughout. Young adults used significantly more representational gestures in the no-CG vs. CG condition, $t(15) = 4.136$, $p < .001$. For old adults, the trend goes in the opposite direction, however, this difference was not significant, $t(15) = 1.981$, $p = .06$. The comparison of age groups within conditions also did not yield significant differences (both p's > .05).
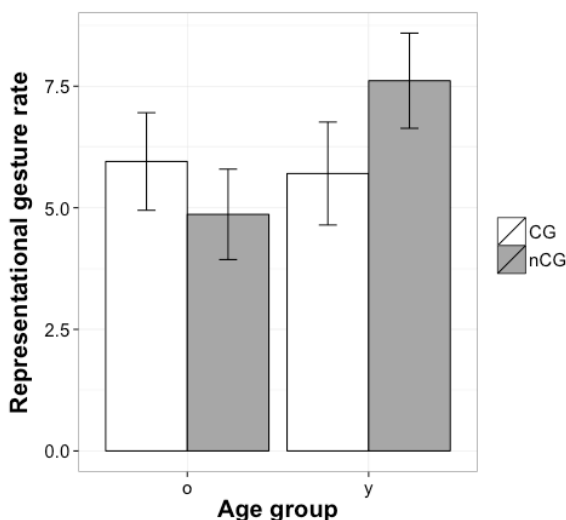

Figure 3. *Average representational gesture rate per age group and common ground condition. Error bars represent the SE.*

4. Discussion

We investigated how younger and older adults adapt their speech and co-speech gestures to an addressee's knowledge state when narrating short comic strips. Younger, but not older

adults produced more words and more representational co-speech gestures when relating content that was novel to the addressee. The individual results will be discussed in more detail in the following.

The expected effect of common ground on speech, i.e. fewer words to narrate known story content and more words to narrate novel content was only significantly present in younger adults. For the younger adults, this is in line with some previous findings on speech and common ground in similar narration tasks (e.g. [3], [28]), supporting the idea that the more knowledge interactants assumed to be mutually shared, the faster and more efficient their communication gets. The fact that younger adults frequently referred to the 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" clearly indicates that they were aware of their addressee's knowledge state. Hence we can safely assume that our manipulation of common ground worked as we intended. Older adults, on the other hand, hardly differed in the number of words they used to narrate familiar vs. novel story content, and made very few explicit references to common ground. Two explanations are conceivable: Older adults may not be able to engage in audience design based on common ground as we induced it here due to cognitive factors, but they may also have different communicative goals than younger adults, as laid out in the following to paragraphs.

From a cognitive perspective, it may be that older adults simply do not remember what does and what does not constitute common ground, i.e., in the present task, which half of the story they had inspected together with the addressee at the beginning of the trial.[3] Alternatively, they may still remember which information is mutually shared between them and the addressee, but then are unable to use this knowledge when designing their utterances, potentially due to a failure to retrieve the relevant information in time in order to plan the utterance accordingly (as suggested by [10]). Remembering which knowledge is in common ground and designing one's utterances accordingly is arguably more challenging than taking global addressee features like addressee age into account, a task which older adults have been shown to be able to do successfully (e.g. [13], [14]). Still, the small difference between the number of words used to narrate familiar vs. novel story content for the older adults is surprising, given that previous research using quite complex manipulations of emerging common ground did find an effect ([10], [11], [12]).

One should therefore also consider the possibility that older adults' communicative goals differ from those of younger adults (see also [13]). Older adults may *choose* to give equal weight to both known and unknown story content in their narrations. Whereas young adults may have the goal to enable their addressee to correctly answer the question he would receive after the narration, focusing on providing information that the addressee does not yet have, older adults may have the primary goal of narrating "a nice story". For example [33] found that older adults are judged to be better at story telling than younger adults. We did not obtain objective ratings of narration quality, but the first author's personal impression was that older adults, more so than younger adults, largely enjoyed the task, putting considerable effort into narrating the stories in an entertaining and animated way. Older adults frequently added additional material to their story, e.g. attributing intentions and feelings to the characters, whereas younger adults were more likely to include information on smaller, visual details of the individual frames which they thought might be relevant to answering the

---

[3] Unfortunately, we did not assess whether participants remembered which part they had inspected together at the end of the task, so this interpretation remains speculative.

question. Obviously, both cognitive *and* pragmatic factors may influence how younger and older adults speak for an addressee.

There were no age differences for representational gesture rate. This is contrary to previous findings by [30] and [31], which suggest that older adults use significantly fewer representational gestures than younger adults. We propose that this is due to the more communicative design we employed here. Whereas participants in the two previous studies either had no addressee at all or an experimenter-addressee, in the present study we used naïve, real addressees. Research with younger adults indicates that the presence of a visible, attentive addressee increases the production of representational gestures (e.g. [21], [34]). This suggests that, given the appropriate context, older adults have sufficient cognitive capacities to produce potentially complex gestural forms concurrently with speech. It should also be noted that the older adults in our sample were a little younger (*M* = 67.69 years) than in [30] and [31], where the mean age was about 70 years.

Crucially, with respect to the hypotheses we set out to test, older adults' representational gesture production was not sensitive to their addressee's knowledge state. Whereas younger adults produced representational gestures at a higher rate when relating novel as compared to mutually shared content, this was not the case for older adults. Our finding for younger adults replicates some of the earlier findings on common ground and gesture production in studies using comparable tasks ([4], [21]). The fact that representational gesture rate is influenced by contextual factors such as mutually shared knowledge between a speaker and an addressee lends support to views of gestures as communicatively motivated [35]. Additionally, it is in line with accounts of gesture production claiming that speech and gesture are part of a single, integrated system [36] in which both modalities tightly interact with each other and information conveyed in gesture is semantically coordinated with information conveyed in speech (Interface Hypothesis, [37]). In the current study, we found that an increase in information conveyed in the spoken modality is coupled with an increase in information conveyed in the gestural modality. The idea that "more speech goes with more gesture, less speech with less gesture" ([26], p. 243) is expressed in the "hand-in-hand" hypothesis of gesture production as formulated by [38] who propose that speech and gesture behave in a parallel fashion. Although in our case, more speech goes with *even* more gesture, our findings support the notion of a parallel increase in both modalities.

Thus considered, it is also not surprising that older adults' representational gesture rate is not influenced by the presence of common ground. As they show no sign of audience design in their speech, why should they do so in gesture? The same cognitive and/or pragmatic factors that influence older adults' verbal behavior may also influence their gestural behavior, again underlining the tight parallel between the two modes of communication.

5. Conclusions

The results of the present study suggest that there is an age-related difference in how speakers adapt their speech and co-speech gestures based on mutually shared knowledge with an addressee. Younger, but not older adults, convey more information both in their speech *and* in their gestures when there is common ground as opposed to when there is not. Whether these differences in verbal and gestural behavior have an impact on how older adults are comprehended by others, and on the overall quality of their interactions remains to be investigated.

7. References

[1]   Thornton, R. and Light, L.L. (2006). Language comprehension and production in normal aging. In J. E. Birren and K. W. Schaie (Eds.). *Handbook of the Psychology of Aging* (6th Ed.) (pp. 261-287). San Diego, CA: Elsevier Academic Press.

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

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

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

[5]   Kelly, S. D., D. J. Barr, R. B. Church, & K. Lynch (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, 40: 577-592.

[6]   Kelly, S.D., Özyürek, A. & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, 21(2), 260-267.

[7]   Clark, H.H. (1996). *Using Language*. Cambridge University Press, Cambridge.

[8]   Fussel, S.R. and 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.

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

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

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

[12]   Lysander, K. and 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.

[13]   Adams, C., Smith, M.C., Pasupathi, M., and Vitolo, L. (2002). Social context effects on story recall in older and younger women: Does the listener make a difference? *Journal of Gerontology: Psychological Sciences*, 57B(1), 28-40.

[14]   Keller-Cohen, D. (2014). Audience design and social relations in aging. *Research on Aging*, 1-22.

[15]   Gould, O.N. and Shaleen, L. (1999). Collaboration with divers partners: How older women adapt their speech. *Journal of Language and Social Psychology*, 18, 395-418.

[16]   Holler, J. & G. Beattie (2003). Pragmatic aspects of representational gestures: Do speakers use them clarify verbal ambiguity for the listener? *Gesture*, 3(2), 127-154.

[17]   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.

[18] 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.

[19] Bavelas, J., Gerwing, J., Sutton, C., and Provost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495-520.

[20] 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.

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

[22] Gerwing, J. and Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture*, 4(2), 157-195.

[23] Holler, J. & R. Stevens (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.

[24] Parrill, F. (2010). The hands are part of the package: Gesture, Common ground and information packaging. In S. Rice and J. Newman (Eds.), *Empirical and Experimental Methods in Cognitive/Functional Research*, pp. 285-302.

[25] Campisi, E., & Ozyurek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. Journal of Pragmatics, 47, 14-27.

[26] 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.

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

[28] Holler, J. and 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.

[29] Holler, J., Tutton, M. and Wilkin, K. (2011). Co-speech gestures in the process of meaning coordination. In *Paper presented at the 2nd GESPIN – Gesture and Speech in Interaction Conference*, Bielefeld.

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

[31] Feyereisen, P. and 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. [32] Wittenburg & al., 2006

[33] James, L.E., Burke, D.M., Austin, A., and Hulme, E. (1999). Production and perception of "verbosity" in younger and older adults. *Psychology and Aging*, 13, 355-367.

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

[35] Melinger, A. & W. J. M. Levelt (2004). Gesture and the communicative intention of the speaker. *Gesture*, 4(2): 119-141.

[36] McNeill, D. (1992). *Hand and mind*. Chicago, London: The Chicago University Press.

[37] 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: 16-32.

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