Cocks, N., Morgan, G. & Kita, S. (2011). Iconic gesture and speech integration in younger and older adults. Gesture, 11(1), pp. 24-39. doi: 10.1075/gest.11.1.02C0C

# Iconic gesture and speech integration in younger and older adults

Naomi Cocks[a], Gary Morgan[a], and Sotaro Kita[b]
[a]City University London / [b]University of Birmingham

This study investigated the impact of age on iconic gesture and speech integration. The performance of a group of older adults (60–76 years) and a group of younger adults (22–30 years) were compared on a task which required the comprehension of information presented in 3 different conditions: verbal only, gesture only, and verbal and gesture combined. The older adults in the study did not benefit as much from multi-modal input as the younger adults and were more likely to ignore gesture when decoding the multi-modal information.

**Keywords:** gesture, integration, aging

There has been a growing body of research which has investigated how aging impacts on communication (see Thornton & Light, 2006, for a review of this work). Most of the studies in this area however, have focussed on verbal communication with non-verbal aspects, including gesture, receiving very little attention.

Some research on the impact of aging on gesture suggests that production, imitation and comprehension of gesture are all affected. The two studies which investigated gesture production found that older adults produced less co-speech iconic gestures (Cohen & Borsoi, 1996; Feyereisen & Harvard, 1999). Co-speech gesture is the use of body movements (hand, head and face) while a person speaks. Iconic gestures are those which depict visually an action or an object. The only study on gesture imitation also suggests that older adults are less accurate and slower at imitating gesture (Dimeck, Roy, & Hall, 1998). Finally, researchers have found that older adults are worse at naming and categorising pantomime gestures (Ska & Croisile, 1998), at identifying emotional gestures (Montepare, Koff, Zaitchik, & Albert, 1999) and at using gesture as a strategy to support verbal language (Thompson, 1995; Thompson & Guzman, 1999). These studies make an important contribution to our understanding of how aging impacts on gesture comprehension but there are some unanswered questions that remain.

*Gesture* 11:1 (2011), **24–39**. DOI 10.1075/gest.11.1.02coc
ISSN 1568–1475 / E-ISSN 1569–9773 

Ska and Croisile (1998) investigated the impact of age on pantomime gesture in the following five areas: gesture decision, recognition of correct gestures, similarity judgement, action designation and gesture naming. They reported that older adults were significantly worse at the similarity judgement task, the naming task and the gesture decision task. They proposed that these findings were due to two reasons. The first reason they gave was that the older adults had difficulty with the structural attributes of the pantomime gestures. They supported this with evidence that older adults had difficulty with determining whether gestures were similar; however, what they meant by similar was not described. Their second reason was that older adults had difficulties with evaluating and combining visual cues. Older adults tended to make visual or semantic errors when naming gestures and often interpreted meaningless gestures as meaningful.

Montepare, Koff, Zaitchik and Albert (1999) investigated the impact of age on comprehending emotional gestures and found that although both younger and older adults made accurate emotion identifications well above chance levels, older adults made more overall errors. This was particularly true for negative emotions. The older adults' errors were more likely to reflect the misidentification of emotional displays as neutral in content. Again the gestures used in this study were produced in isolation rather than co-occurring with speech.

The task used by Ska and Croisille (1998) and Montepare et al. (1999) required participants to comprehend gestures in isolation. All people, regardless of culture, use gestures alongside their speech as part of communication (McNeill, 2000). Therefore in order to have a true understanding of the impact of age on gesture comprehension, it is necessary to investigate not only comprehension of gestures in isolation but also gestures that occur alongside speech.

The research by Thompson and colleagues (Thompson, 1995; Thompson & Guzman, 1999) did investigate gesture that co-occurred with speech. Thompson (1995) investigated the impact of age on the effectiveness of gesture and visible speech ("visible articulatory movements on the face", p. 215) in aiding recall of sentences and in a second study investigated the impact of age on the benefit of gesture and visible speech in difficult listening conditions (Thompson & Guzman, 1999). In Thompson (1995), while the older adults relied more on visible speech than younger participants, gesture was shown to help younger adults recall sentences but was of little help to older adults. Thompson and Guzman (1999) also found that while young adults benefited from gesture and visible speech in difficult listening conditions, the older adults did not. They suggested that the difference in the impact of gesture on recall may be related to the cognitive changes associated with aging, particularly due to reduced working memory capacity. They wrote that cognitive resources may be "consumed" with language processing and that this would mean there were fewer resources left over for gesture processing.

The need for researching how aging affects co-speech gesture comprehension has been identified by a number of researchers (e.g., Goldin-Meadow, 2003). While the Thompson (1995) and Thompson and Guzman (1999) studies make an important first step in investigating the impact of aging on gesture that co-occurs with speech, the gesture and speech messages in these studies were redundant as the meanings portrayed by both modalities overlapped. This meant that the older adult did not need to attend to gesture in order to complete the task. Therefore the reason gestures may not have been of benefit for older adults in both studies may be simply because the task did not require the participants to attend simultaneously to speech and gesture. In order to more fully understand how age impacts on gesture comprehension, we must also explore older adults' ability to integrate related but different information that is portrayed in speech and gesture. This can only be done by including a task in which gesture adds more meaning to the spoken language and the participant is therefore required to integrate the meaning from gesture and speech.

In this research we were interested in gestures that depict the properties of some object or action, sometimes referred to as "iconic gestures". Such gestures are often used to convey information that may not be overtly conveyed verbally, such as object size, object location, manner of movement, spatial relationships and an object's path of movement (Church & Goldin-Meadow, 1986; Kita & Özyürek, 2003). Kendon (2004, p. 160–161) outlines several ways of classifying gestures of this sort including: modelling, enactment (or pantomime) and depiction. Kendon makes the further point that the contribution of these gestures to the meaning of the multi-modal utterance is often 'perfunctory' or 'sketchy' and very reliant on an understanding of the context in which they are employed. Therefore, in order to fully understand a speakers' intention, addressees are required to comprehend and integrate information from both the speech and the gesture channels. This psycholinguistic process is be referred to here as gesture and speech integration (Cocks, Sautin, Kita, Morgan, & Zlotowitz, 2009).

In recent years there has been a growing body of research that has investigated the integration of gesture and speech. This research has mainly focussed on how processing is affected by gesture and speech congruity and synchrony. The main finding is that when a semantically incongruent speech-gesture combination was presented, processing of the gesture was negatively affected by the incongruent speech. Furthermore, that processing of the speech was also negatively affected by the incongruent gesture; namely, concurrent speech and gesture influenced each other's processing (Kelly, Özyürek, & Maris, 2010). This therefore suggests that gesture provides context for speech comprehension and speech provides context for gesture comprehension. Similarly, the automaticity of gesture and speech integration is affected by asynchronous production of gesture and speech (Habets, Kita, Shao, Özyürek, & Hagoort, 2011; Obermeier, Holle, & Gunter, in press). Thus researchers conclude that when gesture and speech are synchronised they are integrated more efficiently (Habets, Kita, Shao, Özyürek, & Hagoort, 2011) and when they are not integration relies more on controlled (non-automatic) memory processes (Obermeier, Holle, & Gunter, in press). While all of these studies make a contribution to the body of research on gesture and speech integration, none of them investigated the impact of aging on this process.

While no research has investigated whether there are age-related changes in gesture and speech binding and/or integration, there is a growing body of research which suggests that there are changes in integration of spatial information in other types of tasks (Anderson & Ni, 2008; Copeland & Radvansky, 2007) and age-related changes to bisensory augmentation of written and audio material (Stine, Wingfield, & Myers, 1990). Similarly, there is a growing body of research which suggests older adults have more difficulty remembering context (see Spencer & Raz, 1995, for review). This therefore sets up the possibility that there will also be age-related changes to the ability to successfully integrate information received from gesture and information received from speech.

The current study investigated whether there were differences in speech and gesture integration abilities between younger and older adults. In addition to a task that required participants to integrate gesture and speech, we also included two uni-modal conditions that required participants to either comprehend just gesture or just speech. Based on the previous findings reviewed earlier, we expected that the older participants would be significantly worse in the integration task than the younger participants, but would perform similarly in the two uni-modal conditions.

## Method

### *Participants*

20 young adult participants (10 female and 10 male, 16 right handed and 4 left handed), aged 22–30 ($M = 24.35$, $SD = 2.43$) and 19 older adult participants (11 female and 8 male, 17 right handed and 2 left handed), aged 60–76 ($M = 68.05$, $SD = 5.3$) participated in this study. All participants were English-speakers and had no history of severe head trauma, stroke or progressive neurological disease. The younger group had received slightly more years of formal education ($M = 14.85$, $SD = 0.93$) than the older group ($M = 12.26$, $SD = 2.6$). The older group was re-

cruited as control participants for a separate study that investigated the impact of aphasia on gesture comprehension (Cocks et al., 2009).

### *Materials and procedure*

An actor, with an opaque face covering to reduce the influence of facial expression on comprehension, was filmed enacting 21 gestures depicting common actions (e.g. dancing, writing, cutting) and saying 21 verbal phrases. Only high frequency, semantically simple verbs were used. Actions were excluded if there was likelihood that they would be more familiar to one of the age groups. For example, we did not use any gestures that related to technology. In an aim to replicate the types of gestures that are produced alongside speech, the gestures produced were vague and less detailed than pantomime gestures. These videos were then edited and three conditions were created. These were as follows: a Verbal-Gesture condition (VG) (verbal phrase was heard and gesture was seen); Gesture only (G) (no verb phrase was heard only gesture was seen); and Verbal only (V) (verbal phrase was heard and video was replaced with a still image of the actor). A total of 63 short videos were produced. Please see Appendix 1 for example of video clips of the gesture for the "I paid" item and Appendix 2 for a list of items.

Participants were asked to watch the five second videos which were embedded in a PowerPoint presentation on a laptop with a 15.4 inch screen size. They were then asked to select from a choice of four colour photographs in a book in front of them with the instruction "point to the photograph that best matched the message portrayed in the video". The photographs were of people carrying out the action. For actions that involved objects, the objects were included in the picture.

Participants watched all 21 speech-gesture combinations in all three conditions. This resulted in a total of 63 trials for each participant. Seven speech-gesture combinations were presented in each of the following orders: V-VG-G, G-V-VG, VG-G-V. Three counterbalancing sets were created, so that each speech-gesture combination was represented in every set in a different condition. For example, the "I lit it" speech-gesture combination, was presented in the VG condition in stimuli set 1, in G condition in set 2, and V condition in set 3, whereas, the "I rode it" speech-gesture combination, was presented in the G condition in stimuli set 1, in V condition in set 2, and VG condition in set 3. Within each set, the order of all items was randomised and items were not "blocked" together according to condition. The three counterbalancing sets were combined to create the 63 trials. Each had an accompanying set of four colour photographs. The same four photographs were used for each gesture-speech combination in each condition but the layout of the four photographs on the A4 page of the response booklet differed for each condition. For example, for "I rode it" in the VG condition the target photograph was in the top left quadrant but in the V condition the target photograph was in the top right quadrant.

The photographs included (e.g. for the stimuli where the speech was 'I paid', and the gesture was a hand moving left to right with repeated circular wrist movements as if to write something with a pen): an integration target (writing a cheque), a verbal only match (handing over cash), a gesture only match (writing a letter), and an unrelated foil (reading a newspaper). The integration target and the verbal only match were both semantically congruent with the speech and were therefore equally likely to be selected in the verbal only condition. The integration target and the gesture only match were both semantically congruent with the gesture and therefore were equally likely to be selected in the gesture only condition. The unrelated distracter was semantically associated to the gesture only match but it was not congruent with the speech or the gesture. It was therefore unlikely to be selected in any of the conditions. The integration target was the only congruent choice when all the information from the speech and the gesture was integrated and therefore was the only possible correct response in the Verbal-Gesture condition. It was the gain in the number of integration target choices between the unimodal tasks and the Verbal-Gesture condition that was of interest.

After each video clip, participants were asked to "point to the photograph that best matched the message portrayed in the video". Participants were instructed that they could only select one picture. The average duration of the total experiment was 45 minutes. All participants had three practice trials. The experimenter recorded which picture the participant pointed to.

### *Data analysis*

In order to derive an index of participants' ability to integrate verbal and gestural information, we used the following procedure. If participants integrate the information from speech and gesture in the VG condition, the integration target is the only possible interpretation. However, the integration target can be selected even when participants focus only on a single modality (speech or gesture) and ignore the other modality (Kelly, Barr, Church, & Lynch, 1999). This is because the integration target is compatible with both the speech and the gesture. Thus, we could not estimate how well participants integrated verbal and gestural information by just looking at how often they successfully selected the integration target in the VG condition. Therefore, we needed to evaluate whether the probability of choosing the integration target in the VG condition was higher than what would be expected from the probabilities of choosing the integration target on the basis of a single modality. To do this we used the same calculation as used in Cocks et al. (2009).

The key to this calculation is estimating the probability of the integration target being selected in the VG condition when participants focus on a single modality (without integrating speech and gesture) (henceforth, P(Unimodal)). This probability is estimated as a weighted mean of the following two proportions: the proportion of trials in which the integration target was chosen in the V condition and the equivalent proportion in the G condition, as shown in (1).

(1) P(Unimodal) = WV * Prop. of integration target choice in V + WG * Prop. of integration target choice in G, where WV and WG ( WV + WG = 1) are the weights for the weighted mean.

The two weights, WV and WG, represent how likely the verbal modality and the gestural modality are taken into account. We assume that participants are more likely to use (i.e., give more weight to) the modality that is stronger and provides more accurate information. Therefore, WV and WG are estimated as normalised proportions of trials in which correct choices were made in the V condition (henceforth, PCV) and in the G condition (henceforth, PCG) respectively, as shown in (2) and (3).

(2) WV = PCV / (PCV + PCG)

(3) WG = PCG / (PCV + PCG)

Normalisation (i.e., dividing the proportions by (PCV + PCG) ) ensures that the sum of the two weights equals to one. Note that the correct choices in the V condition are the integration target and the verbal only match, and those in the G condition are the integration target and the gesture only match.

The multimodal gain (henceforth, MMG) is an index of the extent to which verbal and gestural information is integrated in the VG condition. This is estimated as the difference between the proportion of trials in which the integration target was chosen in the VG condition and P(Unimodal), as in (4).

(4) MMG = P(VG) – P(Unimodal)

MMG represents the likelihood of choosing the integration match in the VG condition that cannot be accounted for by a choice made on the basis of a single modality (either verbal or gesture). In other words, MMG represents how well participants integrate verbal and gestural information to form a unified interpretation.

**Table 1.** Mean percentage of three types of targets (Verbal Only Match, Gesture Only Match, Integration Target and Unrelated Foil) chosen by the younger adults and the older adults in each of the conditions (verbal only, gesture only and verbal gesture combined). SDs are shown in parentheses.

| | **Verbal Only Match** | | **Gesture Only Match** | | **Integration Target** | | **Unrelated Foil** | |
|---|---|---|---|---|---|---|---|---|
| | **Young** | **Older** | **Young** | **Older** | **Young** | **Older** | **Young** | **Older** |
| Verbal Only | 58.16 (9.73) | 56.51 (9.89) | 0.26 (1.12) | 2.22 (3.19) | 41.05 (8.97) | 41 (9.44) | 0.53 (1.62) | 0.27 (1.21) |
| Gesture Only | 1.58 (2.47) | 2.49 (4.06) | 38.16 (10.5) | 45.15 (13.16) | 60.26 (10.45) | 51.8 (13.6) | 0 (0) | 0.56 (1.66) |
| Verbal-Gesture Combined | 4.21 (7.36) | 16.34 (10.06) | 2.37 (2.69) | 2.22 (3.6) | 93.16 (7.65) | 80.61 (11.37) | 0.26 (1.18) | 0.83 (1.97) |

Note: In Verbal Only, both verbal matches and integration targets were correct responses. In Gesture Only, both gesture matches and integration targets were correct responses. In Verbal-Gesture Combined, only integration targets were correct responses.

## Results

Our aim was to determine whether people used integrated information to narrow their choices of targets and whether this differed between the younger and older adults. To do this we first examined the percentages of integration targets chosen in VG condition and identified outliers. If both sets of participants selected the target on average more than 2 standard deviations below the mean for both groups, the item was removed from analysis. This was the case for two items 'I arrived' and 'I walked'.

An error analysis was carried out to determine what the younger and older participants selected when they were unable to integrate (see Table 1). The older participants selected the verbal only match significantly more often than the younger participants, $t(37) = 4.31$, $p < .05$ (Older Adults = 16.34%; Younger Adults = 4.21%). We used the alpha level of .05 for statistical significance throughout, except in post-hoc analysis. Thus, when presented with gesture-speech combinations, the older participants used just the verbal information more frequently than the younger participants. There was no significant difference between the older and younger participants on the percentage of trials in which the gesture only match was chosen, $t(37) = 0.15$ (Older Adults = 2.22; Younger Adults = 2.37).

We then conducted Analyses of Variance (ANOVAs) by participants. The scores from all stimulus items were averaged for each participant, and the participant scores were entered into statistics.

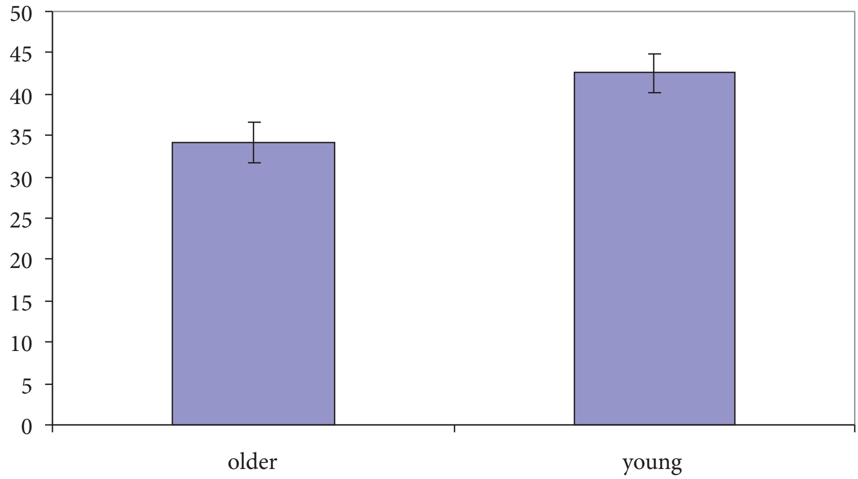


**Figure 1.** Mean MMG (multimodal gain, i.e., the performance gain in the Verbal-Gesture Combined condition that cannot be attributed to the unimodal processes) for younger and older adults. The error bars indicate the standard error of the mean.

First we submitted the percentage of trials in which the integration target was chosen to a mixed design ANOVA for the three conditions (V, G and VG) and age (see Table 1). There was a significant main effect of condition $F(2,74) = 192.61$, $MSE = 21306.4$, $p < .05$, a significant main effect of age $F(1,37) = 13.69$, $MSE = 1441.4$, $p < .05$, and a significant interaction for age and condition $F(2,74) = 3.57$, $MSE = 395.35$, $p < .05$. Post-hoc t-tests with Bonferroni correction revealed the following: For each age group, the percentage of integration targets chosen was significantly higher in VG than in V (younger adults, $t(19) = 15.98$, $p < .0001$; older adults, $t(18) = 12.27$, $p < .0001$) and in G (younger adults, $t(19) = 13.31$, $p < .0001$; older adults, $t(18) = 8.53$, $p < .0001$). This indicates that both the younger and older adults integrated information from gesture and speech in order to select the most appropriate response in the VG condition. Additionally, the percentage of integration targets chosen was significantly higher for G than for V for the younger participants ($t(19) = 5.88$, $p < .0001$). In V both the verbal match and the integration target were appropriate responses and in G both the gesture match and the integration target were appropriate responses. An inspection of the data indicated that there is no clear reason to why younger participants chose the integration target more in the G than the V condition. None of the other post-hoc t-tests with Bonferroni correction were significant.

In order to confirm the above results and thus determine if there were age differences in gesture and speech integration, multimodal gain (MMG) scores were calculated using the formulae described in the methods section. Note that we could not be sure that there was a difference in integration between the two age groups by simply comparing their performance in VG, as the integration targets could have been chosen only on the basis of one of the modalities. Similarly, we could not simply compare VG with one of the unimodal conditions (e.g., V), as the integration targets could have been chosen only on the basis of the single modality that was not compared with VG (e.g., the gesture modality). Thus, we needed to use the MMG score, which took into account, within a single variable, how often the integration targets were chosen in both unimodal conditions. The younger participants had a significantly higher mean MMG score ($M = 42.57$, $SD = 10.7$) than the older participants ($M = 34.13$, $SD = 10.7$), $F(1,37) = 6.05$, $MSE = 695.5$, $p < .05$ (see Figure 1). This suggests that older participants were worse at integrating iconic gesture and speech to obtain meaning.

In order to further understand the impact of participants' age on gesture and speech integration we needed to determine how well participants comprehended gesture and speech in isolation. In the uni-modal verbal only (V) and gesture only (G) conditions, both the older and younger participants were close to ceiling on selecting the matched targets (either the integration target or the verbal/gesture only match) (see Table 1). This indicated that both groups of participants were able to comprehend gesture and speech in isolation.

## Discussion

The results indicated that older and younger adults were both able to accurately comprehend gestures and speech in isolation. However the results of the multi-modal task revealed that the older adults were significantly worse than younger adults at integrating information from both gesture and speech. Furthermore, an error analysis for the multi-modal condition indicated that the older participants ignored the gesture information more often than the younger participants.

This is the first study to show that aging impacts on speech and gesture integration because in this study participants were required to integrate related but different information from gesture and speech. The findings go beyond those of Thompson (1995) and Thompson and Guzman (1999), which investigated the impact of age on the benefit of gesture on the recall of sentences and in difficult listening conditions. The current results add to the growing evidence that suggests that older people are less accurate at tasks that require them to integrate information (Anderson & Ni, 2008; Copeland & Radvansky, 2007).

What is the nature of the speech-gesture integration process relevant to the current study? In order to choose the integration-target photograph when presented with both speech and gesture, the participants needed to use speech to constrain the interpretation of gesture and use gesture to constrain the interpretation of speech. In other words, speech and gesture served as a context for each other (Kelly, Barr, Church, & Lynch, 1999). Speech and gesture were likely to have been processed in parallel, narrowing down each other's interpretations as each modality is being processed, (rather than the processing of one modality waiting for the processing of the modality to be completed) because a recent ERP study showed that speech and gesture are integrated into context in parallel (Özyürek, Willems, Kita, & Hagoort, 2007). This integration is likely to have happened during the encoding phase when the stimuli were initially presented (rather than in the memory retention (response) phase in which the participants selected a photograph based on the speech and gesture they had just seen). This is because ERP studies have shown that speech-gesture integration starts as soon as the visual and/or auditory information becomes available (e.g., Kelly, Kravitz, & Hopkin, 2004; Özyürek, Willems, Kita, & Hagoort, 2007). However, it should be noted that the current methodology does not allow us to rule out the alternative possibilities that speech and gestures were processed one by one (rather than in parallel) and the integration took place during the memory retention (response) phase of the task. It would be an interesting topic for future studies to investigate the time course of the integration process in the task used in the current study.

Why was this integration process affected by aging? One possibility is that the integration process requires working memory capacity in order to retain and update intermediate results of the interpretation processes for speech and gesture. Although like Thompson (1995) we did not assess working memory in this study, poor working memory scores associated with aging has been used to account for previous sets of results linked to poor integration of information (Copeland & Radvansky, 2007). One possibility is that aging reduced the working memory capacity, made it more difficult to take information in two modalities into account and lead to heavier reliance on one modality. The error analysis for the verbal and gesture condition found that the older participants were more likely to use just the verbal information and disregard the gesture information. Furthermore, this is not because the older participants found gestures to be difficult to interpret as the older participants performed as well as the younger participants in the gesture condition. The results therefore support the suggestion by Thompson (1995) and Thompson and Guzman (1999) that reduced working memory capacity in older adults may result in resources being consumed by speech processing operations leaving very little left over for gesture comprehension and therefore subsequent integration. A question then arises as to why speech processing is prioritized over gesture processing when processing demands exceed processing capacity. This may be because speech can sometimes be interpreted without gestures (e.g., conversation on the telephone); in contrast, gestures are often highly ambiguous in isolation (Krauss, Morrel-Samuels, & Colasante, 1991) and when the ambiguity is not resolved by accompanying speech in a timely fashion, speech and gesture are not integrated effectively (Habets, Kita, Shao, Özyürek, & Hagoort, 2011). To summarise, the current results are compatible with the idea that reduced working memory capacity of the older participants made them integrate speech and gesture less and rely more on the speech modality, which can be interpreted unambiguously without gestures in some situations (though not in the current experiment).

The effects of aging on gesture-speech integration mirror the changes during childhood. When three-year olds, five-year olds and young adults performed essentially the same task as in the current study, three-year olds were less able to integrate gesture and speech than five-year olds and adults (Sekine, Sowden, Kita, under review). The authors attributed this age difference in the integration ability to three-year olds' limited capacity to take into account contextual information in communication. This may stem from limited working memory capacity in young children, in line with the findings that young children have limited capacity for language processing as compared to adults (e.g., Bloom, 1970, 1990; Freudenthal, Pine, & Gobet, 2007). Furthermore, Sekine and colleagues also found that both three-year olds and five-year olds relied more on the speech modality than the gesture modality when they failed to choose the integration match in VG. This is similar to the older adults in the current study. The inherent ambiguity of gestures (Krauss, Morrel-Samuels, & Colasante, 1991) may lead to reliance on the speech modality in populations who have integration difficulty such as young children and older adults. The comparison of aging and child development may be a fruitful venue for future research, which may help us triangulate the factors contributing to the ability to integrate gesture and speech.

The current study also found no age difference in comprehension of gestures in isolation. This is at odds with Montepare et al. (1999), who found that older adults were not as accurate as younger adults at interpreting emotion gestures in isolation and also differs to the results of Ska and Croisile (1998) who found that older adults were not as accurate as younger adults at naming and categorising pantomime gestures, that is gestures without accompanying speech. There are several possible reasons for the differences in results. Firstly, the gestures in the current study differed from those used in Montepare et al. (1999) in that they were iconic rather than emotional and the results were at ceiling. Secondly, the tasks differed in the three studies. Ska and Croisille (1998) reported that the older participants often made visual errors and often named a related activity, thus in their task, there was only one right answer. This was not possible in the gesture only (G) condition of this experiment as both the gesture only match and the integration target were classed as correct, it was only when the verbal message was added (in the verbal-gesture combined (VG) condition) that the integration target could be distinguished from the gesture only match. Finally, this difference could have been due to the method of response. In Ska and Croisille (1998) participants were required to verbally name a gesture whereas in the current research, participants were required to point to a picture, which added context and also reduced the number of choices of possible answers, meaning the task was easier. It is also possible, that different types of gestures are affected differently by aging or that a more difficult test may have picked up differences between the two age groups.

While no hearing acuity measures were taken, we assume that all participants were able to hear the speech in the tasks as they were able to accurately decode the speech signal in the verbal only (V) condition and were at ceiling for this task. However, it is possible that the older adults required extra "effort" in order to process the verbal stimuli and thus there were reduced resources available to decode the gesture in the integration task. Rabbit (1991) found that older adults even with very mild hearing loss had more difficulty remembering information presented to them verbally. Rabbit (1991) argued that this was due to the extra effort required by the older adults with mild hearing loss to decode the verbal message and thus leaving very little resources left over for encoding the information in memory. Future research, should investigate the impact of extra effort required for unimodal processing (e.g., speech) on ability to integrate gesture and speech.

In conclusion this study showed that gesture and speech integration is less accurate in older adults than younger adults, and that older adults predominately use verbal information when they do not integrate gesture and speech. These new findings open up new avenues for future research into language, multi-modal communication and aging.

## Acknowledgements

The authors would like to thank Claire Lilley for her help in data collection and analysis. The authors would also like to thank the participants for taking part in the project. This research was funded by a City University London Pump-Priming Grant awarded to the authors.

## References

Anderson, George J. & Rui Ni (2008). Aging and visual processing: Declines in spatial not temporal integration. *Vision Research*, 48, 109–118.

Bloom, Lois (1970). Language development: form and function in emerging grammars. Cambridge, MA: MIT Press.

Bloom, Paul (1990). Subjectless sentences in child language. *Linguistic Inquiry*, 21, 491–504.

Church, R. Breckinridge & Susan Goldin-Meadow (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, 23, 43–71.

Cocks, Naomi, Laetitia Sautin, Sotaro Kita, Gary Morgan, & Sally Zlotowitz (2009). Gesture and speech integration: An exploratory study of a case of a man with aphasia. *International Journal of Language and Communication Disorders*, 44 (5), 795–804.

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

Copeland, David, E. & Gabriel. A. Radvansky (2007). Aging and integrating spatial mental models. *Psychology and Aging*, 22 (3), 569–579.

Dimeck, Paula T., Eric A. Roy, & Craig R. Hall (1998). Aging and working memory in gesture imitation. *Brain and Cognition*, 37 (1), 124–127.

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

Freudenthal, Daniel, Julian M. Pine, & Fernard Gobet (2007). Understanding the developmental dynamics of subject omission: The role of processing limitations in learning. Journal of Child Language, 34, 83–110.

Goldin-Meadow, Susan (2003). *Hearing gestures*. Cambridge, MA: Harvard University Press.

Habets, Boukje, Sotaro Kita, Zeshu Shao, Aslı Özyürek, & Peter Hagoort (2011). The role of synchrony and ambiguity in speech–gesture integration during comprehension. *Journal of Cognitive Neuroscience*, 23 (8), 1845–1854.

Kelly, Spencer D., Dale J. Barr, R. Breckinridge Church, & Katheryn 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.

Kelly, Spencer D., Corrine Kravitz, & Michael Hopkins (2004). Neural correlates of bimodal speech and gesture comprehension. *Brain and Language*, 89 (1), 253–260.

Kelly, Spencer D., Aslı Özyürek, & Eric Maris (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, 21, 260–267.

Kendon, Adam (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.

Kita, Sotaro & Aslı Özyürek (2003). What does cross-linguistic variation in semantic co-ordination of speech and gesture reveal? Evidence of an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, 48, 16–32.

Krauss, Robert K., Palmer Morrel-Samuels, & Christina Colasante (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology*, 61 (5), 743–754.

McNeill, David (Ed.) (2000). *Language and gesture*. Cambridge: Cambridge University Press.

Montepare, Joann, Elissa Koff, Deborah Zaitchik, & Marilyn L. Albert (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior*, 23 (2), 133–152.

Obermeier, Christian, Henning Holle, & Thomas C. Gunter (in press). What iconic gesture fragments reveal about gesture–speech integration: When synchrony is lost, memory can help. *Journal of Cognitive Neuroscience*, 1–16.

Özyürek, Aslı, Roel M. Willems, Sotaro Kita, & Peter Hagoort (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, 19 (4), 605–616.

Rabbit, Patrick M. A. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce IQ. *Acta Otolaryngologica Supplementum*, 476, 167–176.

Sekine, Kazuki, Hannah Sowden, Sotaro Kita (under review). Five-year-olds, but not three-year-olds, integrate information in speech and iconic gesture in comprehension.

Ska, Bernadette & Bernard Croisile (1998). Effects of normal aging on the recognition of gestures. *Brain and Cognition*, 37 (1), 136–138.
Spencer, Wesley D. & Naftali Raz (1995). Differential effects of aging on memory for content and context: A meta-analysis. *Psychology and Aging*, 10 (4), 527–539.
Stine, Elizabeth A. L., Arthur Wingfield, & Sheldon D. Myers (1990). Age differences in processing information from television news: the effects of bisensory augmentations. *Journal of Gerontology: Psychological Sciences*, 45 (1), 1–8.
Thompson, Laura A. (1995). Encoding and memory for visible speech and gestures: A comparison between young and older adults. *Psychology and Aging*, 10 (2), 215–228.
Thompson, Laura A. & Felipe A. Guzman (1999). Some limits on encoding visible speech and gestures using a dichotic shadowing task. *Journal of Gerontology: Psychological Sciences*, 54 (6), 347–349.
Thornton, Robert & Leah L. Light (2006). Language comprehension and production in normal aging. In James. E. Birren, K. Warner Schaie, Ronald P. Abeles, Margaret Gatz, & Timothy A. Salthouse (Eds.), *Handbook of the psychology of aging* (6th ed., pp. 261–287). San Diego: Elsevier.

## Appendix 1: The gesture for "I paid"

## Appendix 2: List of Items

I danced
I walked
I left
I paid
I arrived
I cleaned
I held
I read
I rode
I ate
I sent
I rubbed

I lit
I measured
I flipped
I cut
I lifted
I killed
I raced
I opened
I played

### *Author's address*

Naomi Cocks
Senior Lecturer
Language and Communication Sciences
City University London
Northampton Square
London EC1VOHB

naomi.cocks.1@city.ac.uk

### *About the authors*

**Naomi Cocks** is a qualified speech pathologist and researcher. She received her PhD from Curtin University, Australia. She is currently a Senior Lecturer in the Department of Language and Communication Sciences at City University London. Her research interests include the impact of communication disorder and aging on gesture production and comprehension.

**Gary Morgan** is a psychologist with an interest in gesture and spoken and signed languages. He received a PhD from the University of Bristol and is now professor in the Department of Language and Communication Sciences at City University London. He is also deputy director of the ESRC research centre DCAL at University College London. He is currently working on visual language acquisition, processing, breakdown and evolution in children and adults.

**Sotaro Kita** received a PhD in psychology and linguistics from the University of Chicago. He has worked at the Max Planck Institute for Psycholinguistics and at the University of Bristol, and he is currently a Reader in Psychology at the University of Birmingham in the UK. His research interests include gestures that accompany speaking and thinking processes, word learning in children, and lexical semantics. His research on gesture spans both production and comprehension, and includes studies of patients and children. He has used various research methods: experimental studies with conventional behavioural measures and neurophysiological measures, and field-work based observational studies with qualitative and quantitative analyses.