Language Proficiency Impacts the Benefits of Co-Speech Gesture for Narrative Understanding Through a Visual Attention Mechanism

Natalia Zielinski (nzielinski1@luc.edu)
Department of Psychology, 1032 W. Sheridan Rd.
Chicago, IL 60607 USA

Elizabeth M. Wakefield (ewakefield1@luc.edu)
Department of Psychology, 1032 W. Sheridan Rd.
Chicago, IL 60607 USA

Abstract

Co-speech gestures can enhance a listener's understanding of a spoken message, yet children show a greater benefit of gesture than adults (Hostetter, 2011). We explore whether this effect may be driven by language proficiency, and in turn, by differences in how children visually attend to gesture when they are processing narratives in their stronger vs. weaker language. Bilingual children were shown narratives with scripted gesture in their stronger and weaker languages, while their visual attention was monitored. Memory for narratives was then assessed. Our findings suggest language proficiency *does* affect the degree to which children benefit from co-speech gesture – children showed a greater boost from co-speech gesture when processing narrative in their weaker language. Results also suggest that greater attention to gesture when processing ones' weaker language may underlie this effect.

Keywords: gesture; visual attention; language proficiency

Introduction

Gestures are spontaneous hand movements that are tied temporally and semantically to speech, and used naturally when we communicate with others (McNeill, 1992). Gestures can represent information through their form and movement trajectory and localize referents. Listeners are better able to comprehend a speaker's message if speech is accompanied by gesture than when it is not, yet the size of this effect depends on multiple factors. One of these factors is the age of a listener – children benefit more from having gesture incorporated in communication than adults (Hostetter, 2011). The purpose of the present study is to address whether children might show this benefit due to their weaker language proficiency, and reliance on gesture for interpreting complex speech (Hostetter, 2011). We also ask whether children visually attend more to co-speech gesture when they are less proficient in a language, and whether attention to gesture then impacts a child's ability to recall a speaker's message. A difference in visual attention, coupled with a relation between looking to gesture and remembering information could provide a visual attention mechanism through which gesture in communication is more powerful for children than adults.

Language Proficiency May Affect Gesture Benefits

In her meta-analysis, Hostetter (2011) posits that one reason gesture may be more crucial to children's understanding of a spoken message than adults, is that children are still learning language: they do not have the command of the language that adults do. In other words, spoken language is inherently more difficult for children to understand, and gesture provides an easier, visual route by which to understand a message, which is independent of a child's spoken language skills. This idea is supported by a study examining simple and complex message comprehension in preschool and kindergarten aged children (McNeil, Alibali, & Evans, 2000). When both groups were asked to interpret a simple message, gestures reinforcing the intended message did not influence comprehension. When the message was complex, reinforcing gestures were found to facilitate speech comprehension for the preschool-aged children only. Assuming that preschool children have poorer language proficiency than kindergarten children, this indicates that gestures are more beneficial for those with weak language skills.

One powerful way to test whether language proficiency plays a role in the utility of gesture would be to consider this question within individuals who have different competencies in two languages. This would control for confounding individual differences that might impact the ability of gesture to support comprehension and memory for a spoken message, along with many other differences between children and adults than their general language competency. Because children do have *basic* language skills, a useful population to consider are bilingual children who have a basic level of competency within a second language but are more competent in their first language. We define bilingual children as those who have learned two languages from birth, with regular exposure to both languages. But, because bilingual children often develop language skills in two languages at an uneven pace (Pearson, Fernandez, & Oller, 1993), they pose a unique opportunity to study the modulating effects of language proficiency on the utility of communicative gesture.

Relative Language Proficiency and Visual Attention to Gesture

If language proficiency predicts the utility of gesture for understanding a spoken message, there remains the question of why. One possible mechanism relates to visual attention; perhaps when an individual is less competent in a spoken language, they are more likely to attend to gesture, thus giving gesture a greater opportunity to impact understanding of a message.

Evidence that focus on gesture may increase when understanding a speaker's message is difficult comes from a comparison of attention to gesture during a narrative versus instruction on a new concept. In previous studies of adult's visual attention to gesture in narrative, looking to the speaker's hands is rare (Gullberg & Holmqvist, 2006; Gullberg & Kita, 2009). For example, Gullberg and Holmqvist (2006) found that listeners focused predominately on a speaker's face, only looking to the hands less than 1% of the time gesture was present. Gullberg and Kita (2009) further showed that when gesture *is* attended to, this is most often in response to a speaker providing a cue that gesture is important, such as looking to the gesture while producing it. In contrast, work by Wakefield and colleagues (2018), in which children were asked to attend to instruction on the concept of mathematical equivalence, showed that children fixated on gesture 20% of the time gesture was present. This heightened focus on gesture may be due to children's difficulty understanding the meaning of spoken instruction. If this is the case, listeners may modulate their attention to gesture when they are not understanding enough from speech.

Interestingly, some work has demonstrated that gestures may actually harm comprehension and recall for individuals with weaker language skills. Drijvers, Vaitonytė, and Özyürek (2019) presented native and non-native adult listeners of Dutch with clips of a speaker with either normal speech or degraded speech, and with either no gestures or co-speech iconic gestures. This study found that although non-native speakers allocated more visual attention towards gesture, they did not gain as much benefit from gesture as did native listeners. This may be because the challenge presented from disambiguating degraded speech outweighed benefits from gesture. Native speakers, on the other hand, had less difficulty disambiguating speech and thus had more cognitive resources to allocate towards gesture.

Current Study

The main purpose of this study is to further explore the impact of co-speech gesture on communication, focusing on whether the presence of gesture differentially impacts memory for a spoken message based on a listener's proficiency in that language. We also ask whether direct visual attention to gesture plays a role in this phenomenon. To do this, we employ eye tracking methods to monitor visual attention of bilingual children while they are watching narratives in which a speaker uses co-speech gesture in both their stronger and weaker languages. We then ask children to recall these narratives.

Our paradigm is adapted from Cassell and colleagues (1998), in which adults watched a video of a storyteller recount the plot of a Sylvester and Tweety Bird cartoon. The narration was accompanied by choreographed gesture ***matches*** (i.e., gestures that are redundant with speech) and ***mismatches*** (i.e., gesture that provide additional – but not incongruent – information not found in speech). After watching the video, participants were asked to retell the story. In our study, two narrations are presented, one in a child's stronger language and one in their weaker language, with the inclusion of choreographed gesture matches and mismatches.

We expect children will remember more of the narrative presented in their stronger language due to less difficulty in comprehending the spoken message. We also anticipate that gesture will have a greater effect when children are recalling the narrative presented in their weaker language.

Including both gesture matches and mismatches allows us to further consider the nuances of gesture's effects. If gesture is helping listeners process the basic information in speech, it follows that matching gestures would be of more use to children when processing narrative in their weaker language, whereas mismatching gestures may either clarify the spoken message or perhaps not benefit children if integrating this information with speech is too cognitively taxing.

In terms of visual attention, we hypothesize that children will pay greater visual attention to co-speech gestures in their weaker language, as attention to gesture may be used as a compensatory mechanism for understanding speech (Church, Ayman-Nolley, & Mahootian, 2004). In relating visual attention to recall of information, we hypothesize that looking to a matching gesture will result in higher likelihood of recall, owing to the reinforcement of information. On the other hand, looking to a mismatching gesture may either lead to more modifications in participants' retelling of the narrative, or perhaps even lack of recall, as the additional information in gesture may be integrated into participants' understanding of the story or distract from the verbal information.

Methods

Participants

Data from 17 Polish-English bilingual children between the ages of 6 and 8 (10 female, M_{age} = 7.92, SD_{age} = 1.22) were analyzed for the present study[1]. Children were recruited from Polish Saturday schools and were all identified as Non-Hispanic and White, but coming from diverse socio-economic backgrounds. Children were compensated with stickers and a certificate noting their participation in the study. The schools were compensated with a gift card.

Materials

Language questionnaire. Parents completed a questionnaire assessing their child's proficiency in and exposure to Polish and English. The questionnaire asked for the parents to rate their child's ability in speaking, reading, writing, and comprehension in both languages, in addition to questions about the history of the child's formal education in Polish.

[1] Two additional participants were excluded from this sample for having poor eye tracking (<65%). Additional recruitment of participants was halted due to the COVID-19 pandemic.

Stimuli. Two female actresses were filmed, seated in front of a black curtain, reciting scripted recollections of two cartoons: (1) a Tweety Bird cartoon, in which Sylvester is scheming to capture Tweety Bird, and (2) a Bolek and Lolek cartoon, which follows the two brothers' road trip full of mishaps. One actress recorded the scripts in English; the other actress recorded the scripts in Polish. Actresses spoke and gestured as if they were explaining the story to someone sitting to the right of the camera, thus, they did not look directly at the camera. Gestures were scripted: each narration included 8 matching gestures and 12 mismatching gestures that occurred on isolated pieces of information within a story, what we will refer to as 'story points'. Matching gestures are redundant with speech; for example, a narrator might say "she hears a noise" while cupping her hand behind her ear (see Figure 1). A mismatching gesture provides additional – but not incongruent – information not found in speech; a narrator might say "Sylvester goes up the drainpipe" while cupping her hands in front of her and alternating moving them upwards. Here, the gesture illustrates Sylvester is climbing up a ladder on the drainpipe. The rest of the story points were unaccompanied by gesture, used as a "no gesture" condition in analyses. Actresses filmed each narrative twice, so that different mismatching gestures could be incorporated in each version; these gestures were chosen so that modification to the spoken narrative could be traced back to one version. For presentation, each narration was separated into 6 segments, such that children only had to remember short sections (30-60 sec) of the story at a time. This resulted in 48 stimulus videos (2 stories x 2 versions x 2 languages x 6 segments).

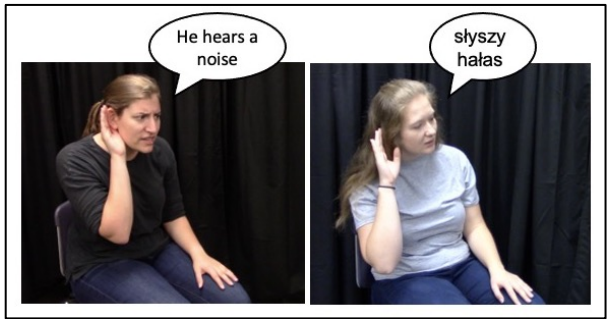

Figure 1: Example of matching gesture in video stimuli.

Eye Tracker. Eye tracking data was collected using a TobiiPro X3-120 remote eye tracker attached to a 15-inch Dell laptop. Tobii eyetracking uses a 9-point calibration procedure using standard animation white dots. Following calibration, the participant was presented with narrative segments during which gaze data was collected and integrated with the video.

Procedure

Children were tested individually in a quiet classroom. Children sat approximately 60 cm in front of the eye tracker mounted to the base of a laptop and were told they would watch videos of women telling two stories and would be asked to recall as much as they could remember about what happened in each story. After calibration, children watched 12 video segments (six belonging to each story) and were asked "what happened?" following each segment. One story was presented in English, and the other was presented in Polish. Children were asked to respond in the language in which the story was presented. Counterbalancing was used to ensure that performance differences were not due to a particular cartoon, the order in which stories were presented, or which language was presented first. Sessions were videotaped on a Logitech camera mounted on the laptop so that speech and gesture could be transcribed in ELAN (2020). Polish responses were then translated into English.

Scoring

Each script was separated into 60 significant story points. Story points are individual pieces of information that could be recalled. Each story thus contained 40 story points with no gesture, 8 story points with matching gesture, and 12 story points with mismatching gesture. Coders, blind to which version of a narration a child saw, determined whether the story points were included in each response (**recall score**). Coders also evaluated whether responses were accurate with respect to the spoken scripts or modified. If the transcript portion was judged to be modified (i.e., words in the participant's response did not explicitly match words in the script), the coders determined if this modification could be traced back to information from a mismatching gesture in Version A or Version B. This **modification score** was used to determine whether information presented uniquely in gesture was integrated in a participant's recollection of the story. Coders also evaluated the participants' gestures, and judged whether the gestures produced during recall could be traced back to either version of the story, and whether these gestures were redundant (matching) or non-redundant (mismatching) with speech.

Across measures, reliability between coders was moderate to strong. **Recall score:** Coders agreed on 1948 of 2040 judgements (94%), K = 0.87. **Modification score**: Coders agreed on 436 of 605 judgements (72%), K = 0.44. In tracing gesture back to the stimulus, coders agreed on 32 of 38 judgements (84%); in judging whether gesture was complimentary or supplementary the coders agreed on all 38 judgements. The kappa values suggest raters agreed substantially for recall and gesture but agreement for modification scores was low. Disagreements were resolved through further discussion.

Results

All analyses were conducted through R Studio (version 1.2.1335), supported by R version 3.6.1 "Action of the Toes" (2019). Analyses relied predominately on the lme4 package, which allows for mixed effects modeling (Bates, Mächler, Bolker, & Walker, 2015). When running mixed effects models through lme4, dummy coding was used, the

default option for coding in this package. To determine whether children showed a difference in their English and Polish proficiency, a Bayesian t-test was conducted, using the Bolstad package.

Behavioral Performance

To address whether the presence of gesture differentially impacts memory for a spoken message based on a listener's proficiency in that language, we sought first to confirm that children showed weaker proficiency in Polish compared to English. Not all parents completed the language questionnaire, but of the 11 who did, children were rated as having stronger ability in English ($M = 4.32$, $SD = 0.83$) than Polish ($M = 3.79$, $SD = 0.85$). A Bayesian t-test favored the alternative hypothesis – that children show differences in language strength ($t(10) = 6.72$, CI: 0.35 to 0.71, $p < .001$). Although this does not account for children whose parents did not submit the questionnaire, it is likely these children also showed weaker Polish language skills, as children were attending Polish school because Polish was their second language, and the goal of Polish school is to improve their language skills. Across the whole sample, we found that proportion of story points recalled from a narrative was much lower when a narrative was presented in Polish ($M = 0.21$, $SD = 0.11$) compared to English narration ($M = 0.39$, $SD = 0.17$). A mixed-effects binomial logistic regression model with whether a story point was remembered as the dependent measure (0,1) and language of the stimulus as the predictor, with participant entered as a random factor revealed that this difference was significant ($\beta = 0.98$, $SE = 0.11$, $t = 9.32$, $p < 0.001$). Together, these results suggest children in our sample were less proficient in Polish compared to English.

The main question, then, is whether the presence of gesture had a different impact on the likelihood children would remember a story point when the narrative was presented in English or Polish (Figure 2). We used a mixed-effect binomial logistic regression model with whether a story point was remembered as the dependent measure (0,1) and language of the stimulus, type of gesture (no gesture, matching gesture, mismatching gesture), and an interaction between these terms as fixed factors, a random intercept for participant, and a random slope for language of stimulus by participant. Our model revealed a significant interaction ($\chi^2(1) = 7.29$, $p = .03$), suggesting that the presence of gesture differently impacted memory for stories in children's stronger and weaker language. Exploring this interaction, children were significantly *more likely* to recall a story point in Polish if it was accompanied by matching gesture compared to no gesture ($\beta = 0.48$, $SE = 0.22$, $t = 2.17$, $p = .03$), but this extra benefit of gesture was not found for English narratives ($\beta = -0.22$, $SE = 0.21$, $t = -1.06$, $p = .29$). Interestingly, children were *not* more likely to recall a story point in Polish if it was accompanied by a mismatching gesture compared to no gesture ($\beta = 0.19$, $SE = 0.20$, $t = 0.94$, $p = .35$) and in the English narrative, there was a trend towards memory being *worse* if story points were accompanied by mismatching gesture, compared to no gesture ($\beta = -0.31$, $SE = 0.18$, $t = -1.78$, $p = .08$).

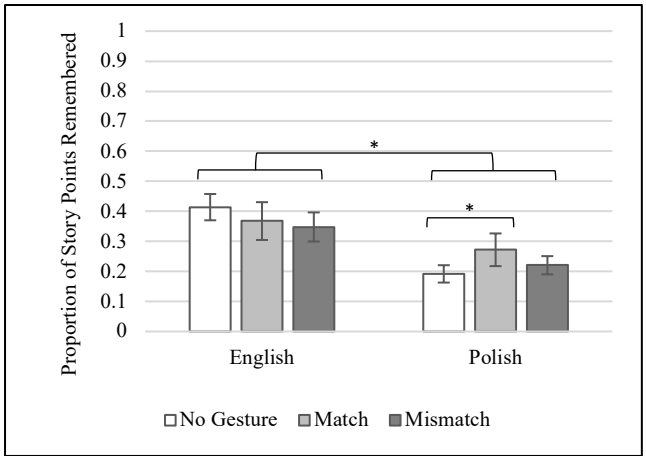

Figure 2: Recall of story points by gesture type in English and Polish

As a second measure of whether gesture impacted memory for narrative differently, based on which language it was presented in, we asked whether participants used almost identical language to the original narration, or modified their language. For example, the narrator said "Bolek presses a button and the car starts to fly" while making either an "airplane" gesture (hands out at both sides) or a "helicopter" gesture (finger pointing up and twirling). An accurate response that one participant recalled was "they hit press a button that makes the car fly". Another participant said "Bolek pressed a button and the car had wings so it could fly"; this would qualify as a modified response because 'wings' were not explicitly said in the original script. We considered whether proportion of modifications were predicted by language and gesture type. There were significantly more modifications to recall following the Polish narration ($M = 0.50$, $SD = 0.50$) than the English narration ($M = 0.27$, $SD = 0.45$; $\beta = 0.99$, $SE = 0.18$, $z = 5.47$, $p < 0.001$). Mismatching gestures resulted in the most modifications to speech in both languages (Eng: $M = 0.56$, $SD = 0.50$; Pol: $M = 0.64$, $SD = 0.49$), followed by matching gestures (Eng: $M = 0.22$, $SD = 0.42$; Pol: $M = 0.62$, $SD = 0.49$), and no gesture accompanying speech (Eng: $M = 0.21$, $SD = 0.41$; Pol: $M = 0.42$, $SD = 0.50$). Linear regression models revealed modification numbers differed significantly ($p < .05$).

We further examined gesture production by participants to consider whether information from gesture is primarily reproduced in gesture. A total of 38 gestures were produced during recall; 21 were produced in English recall, 17 were produced in Polish recall. A majority of gestures were coded as matching with speech (Eng;15, Pol:11), and the rest were considered mismatching with speech (Eng:6, Pol:6). Of the gestures that could be traced back to a gesture from the stimulus, 12 matching gestures were reproduced (Eng:7, Pol:5), and 7 mismatching gestures were reproduced (Eng:4,

Pol:3). Similar to tracing modifications of speech back to stimulus version, we ask whether coders could reliably predict which story version the participant watched based on their reproduction of mismatching gestures. All 7 mismatching gestures were correctly predicted to trace back to the respective version the participant watched.

Visual Attention

Having established that gesture boosts memory for story points more in a child's weaker language compared to stronger language, we next turned to whether visual attention could explain this phenomenon. First, we ask if there are differences in allocation of visual attention to the narrative videos, based on the language of the stimulus. During the study, the two narratives were separated into segments to ease recall but considered at the level of the entire narrative for the purpose of analysis. Dynamic areas of interest (AOIs) were generated for the screen, the narrator's face, body, hands at rest, and gesture space. The body AOI was the rest of the torso space unoccupied by gesture or hands at rest AOIs. When gesture overlapped with torso space, it was considered part of the gesture AOI. Proportion of time spent looking to each AOI was calculated using the total time looking at each AOI divided by the total time looking at the screen.

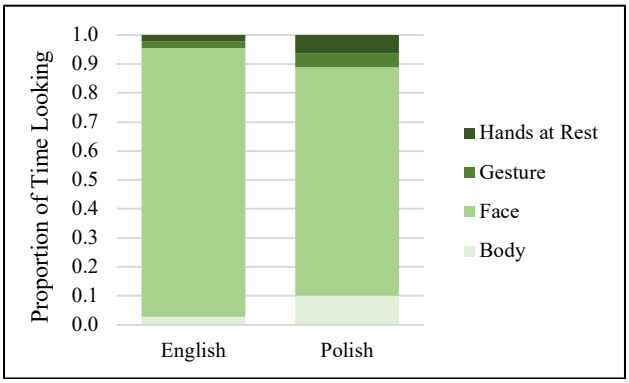

Figure 3: Visual allocation of attention to AOIs in English and Polish

Figure 3 shows participants' allocation of visual attention during narratives to each AOI, separated by language. For the sake of brevity, we consider just attention to gesture: In line with our hypothesis, a linear mixed effects regression with proportion looking to the gesture as the dependent measure, a fixed effect of language of the stimulus (English, Polish), as well as a random intercept for participant, revealed that children allocated significantly more of their attention to gesture when viewing a narrative in Polish ($M = 0.05$, $SD = 0.02$) versus English ($M = 0.02$, $SD = 0.02$; $\beta = 0.02$, $SE = 0.005$, $t = 5.08$, $p < 0.001$).

Relating Visual Attention to Recall

Having established that children attended more to gesture when processing a narrative presented in Polish, our final question was whether this attention predicted recall. To do this, we considered just the story points that were accompanied by matching or mismatching gesture. Figure 4 shows the recall of story points separated by whether they were fixated to, the language of the stimuli, and gesture type. We conducted a mixed effects logistic regression, predicting memory for story points (0, 1), with an interaction between whether gesture was fixated (0,1) and the type of gesture (match, mismatch) as a fixed effect, participant included as a random intercept and a random slope of language of stimulus by participant. This model was a significantly better fit than a simple model without the interaction term ($\chi^2 = 1.95$, $p < .001$). Our model revealed a marginal trend that looking to matching gesture in both languages aids children in recalling story events ($\beta = 0.51$, $SE = 0.30$, $t = 1.74$, $p = 0.08$), but that attention to mismatching gesture does not affect recall ($\beta = 0.12$, $SE = 0.24$, $z = 0.51$, $p = 0.61$). While finding a marginal effect of looking to matching gesture on recall must be treated with caution and further substantiated by increasing the sample size when data collection is allowable, this *may* suggest the following: Matching gesture that reinforces information in speech may boost memory for a story point. However, because listeners are *more likely* to focus on gesture when they are not as proficient in the spoken language, gesture has more of an opportunity to affect memory for narrative when a listener is processing a language in which they are less proficient – this would explain our behavioral result that recall is significantly more likely when matching gesture accompanied story points in the Polish narrative, but not the English narrative.

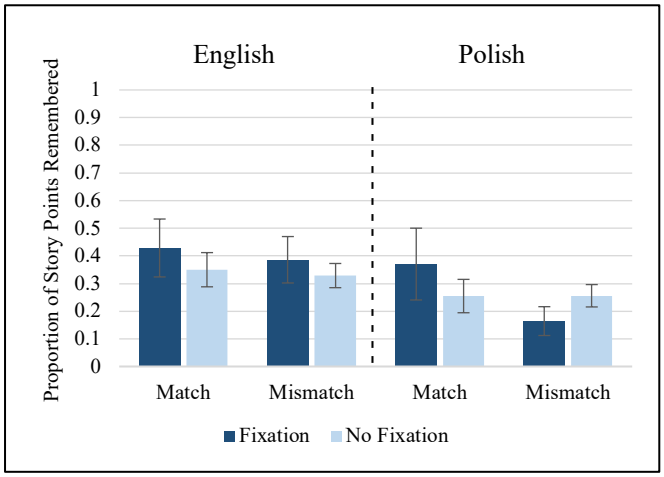

Figure 4: Recall of story points by allocation of visual attention to matching and mismatching gesture

Discussion

Our goals in the present study were to ask (1) whether the utility of co-speech gesture for understanding a spoken message is affected by a listener's language proficiency, and (2) whether this effect is driven by greater visual attention to gesture when a spoken message is more difficult to understand. Our data do suggest that gesture boosts children's likelihood of recalling information from a story when the story is presented in a child's relatively weaker

language, but only if that gesture is redundant with the spoken message (i.e., matching gesture). This boost was not found when narratives were presented in English – the child's stronger language. Nor was this boost found when gesture expressed information *not* found in speech: It appears that only gestures that reinforce a spoken message are helpful for children in their weaker language. Additionally, we suggest this effect may be driven by greater visual attention to gesture when the spoken message was difficult to understand. Supporting this idea, children allocated significantly more of their visual attention to gesture when the narrative was in their weaker language compared to their stronger language. Further, although marginal, the pattern of the data suggests that visually attending to matching gestures in either language facilitated higher recall of narrative than when gesture was not fixated. This should be substantiated with additional data when data collection is possible. If this pattern persists, this would suggest that the reason we see a boost in recall of story points when children are processing a narrative in their weaker language is because they look more to gesture, and this gives gesture an opportunity to boost memory for the spoken message it accompanies.

Our results begin to elucidate *why* Hostetter's (2011) meta-analysis showed that gesture boosts children's ability to comprehend a spoken message more than it boosts adults' understanding: language proficiency seems to play a role, as does visual attention. Focusing on visual attention during narrative, our finding that children allocate more attention to gesture in their weaker language provides direct support for previous literature suggesting that attention to gesture may be used as a compensatory mechanism to understand speech, particularly when the message is complex or the individual has weak verbal fluency (McNeil, Alibali, & Evans, 2000; Hostetter, 2011). As another mode of communicating a target message, gesture may act as a visual tool to help reinforce information in speech. Our finding also aligns with and extends previous eye tracking work: in the few studies that have considered attention to gesture in narrative compared to gesture in instruction (Gullberg & Holmqvist, 2006; Gullberg & Kita, 2009; Wakefield et al., 2018), listeners attend more to gesture during instruction presumably because they do not understand the spoken message and need the support of gesture. We extend this finding by showing that within the same individuals, visual attention to gesture can be modulated based on how difficult the spoken message is to understand on its own.

Furthermore, our results suggest that visual attention to gesture has consequences: we found a marginal trend that story points with matching gesture were better remembered if participants fixated on the gesture. It follows that attending more to matching gesture in the Polish narrative led to better recall of associated story points compared to story points accompanied by no gesture. But this behavioral boost was *not* found when narratives were in English, likely because children did not look to the gesture as often, and therefore the gesture did not have the opportunity to affect recall to the same extent. So, it may be that matching gesture can always lead to better memory for spoken information, but that unless spoken language is difficult to understand – based on either the content, or in this case, a relatively low language proficiency – listeners tend not to look at the gesture and gain the full extent of gestures' potential to aid in comprehension.

The beneficial effect of matching gesture corroborates previous work, which shows that matching gesture aids recall and comprehension due to redundancy of information (Holle & Gunter, 2007). But what about mismatching gesture? Why were children no more likely to recall story points accompanied by mismatching gesture compared to no gesture in the Polish narrative, and marginally *less* likely to remember story points accompanied by mismatching gesture compared to no gesture in the English narrative? Previous literature suggests that mismatching gesture should aid comprehension due to the addition of unique information to a mental model (Hostetter, 2011). But recent work has shown that this might not always be the case. Dargue and Sweller (2020) showed that both children and adults are better able to remember story points from a narrative when that narrative is accompanied by typical gestures, but not atypical gestures. They define typical and atypical gesture based on the degree of sematic-relatedness between the gesture and the speech it accompanies. Although this does not *directly* map onto matching and mismatching gestures, by their nature, mismatching gestures are less semantically related to speech than matching gestures. Thus, our findings may be due to the same phenomenon. Our results suggest that mismatching gesture accompanying a language in which a child is proficient may be difficult for the child to integrate with speech and harm the formation of a cohesive story model by distracting attention away from critical verbal information. Children's recall was not harmed by mismatching gesture in their weaker language because they did not attend to mismatching gesture in that language as much as they attended to mismatching gesture in their stronger language, thus not giving mismatching gesture the opportunity to confuse their recollection of information.

These findings may help elucidate the contradictory findings that looking to gesture harms recall in listeners with weaker language skills (Drijvers, Vaitonytė, & Özyürek, 2019; Kartalkanat & Göksun, 2020). It may be that matching gestures (i.e. gestures that give the *exact* same information as in speech) clarify speech for individuals with weak language skills. However, using mismatching gestures (i.e. gestures that give *any additional* information not found in speech, which would be any gesture when speech is degraded) may increase cognitive load of an individual with weak language skills, resulting in the effects seen in previous studies.

In further examining the effects of matching and mismatching gesture on the quality of recall, we find that most modifications to recall occurred for story points accompanied by mismatching gestures across both languages, conceptually replicating Cassell et al. (1998) and showing the influence of gesture on memory of narrative. In other words, gesture impacted *whether* a story point was remembered and *how* it was remembered.

Finally, the finding that our coders were able to successfully predict which version a participant watched based on their gestures, but not based on their verbal responses, suggests that children may not have developed the capacity to fully express all their ideas in speech and instead express their thoughts through their hands. This is supported by previous literature; when presented with a math equivalence task, children who are not verbally able to explain the correct strategy were able to do so through their gestures (Perry, Church, & Goldin-Meadow, 1988). Moreover, there was no difference in non-redundant gesture production between the two languages, suggesting that bilinguals do not use gesture during recall to compensate for weak verbal skills. Rather, they use gesture to help reiterate or emphasize their recall.

Results from the present study address important questions about how gesture can benefit communication, but also raise questions for future research. Our choice to study bilingual children was well-motivated, as this allowed us to study the effects of language proficiency on gesture's ability to influence recall of narrative within the same individuals. Yet, there is literature that suggests bilinguals may better utilize multiple sources of information in communication (Genesee, Tucket, & Lambert, 1975), and attend differently to gesture than monolinguals (Yow, 2013). For example, Yow (2013) conducted a study with 3-4 year-old monolingual and bilingual children and monolingual adults. In this case, English was the native language of all participants, and monolingual and bilingual children showed similar language proficiency. In the study, an experimenter gestured to one of two characters or did not gesture at all while using an ambiguous pronoun (e.g. "This is Miss Owl and this is Miss Ducky. She wants the ball"). Gesture had a greater impact on how bilingual children interpreted the ambiguous pronouns than both groups of monolinguals, suggesting that bilinguals exhibit a heightened sensitivity to gestures produced by a speaker. Thus, future work should be conducted to compare how monolingual and bilingual children allocate visual attention to gesture, to more fully understand the impact of individual differences on the utility of gesture in communication.

Despite the need for additional research to more fully understand the role of language background in shaping how listeners use gesture to make sense of narrative, the present work provides a novel contribution to our understanding of what individual factors may affect the beneficial effects of gesture in communication: Language proficiency impacts the utility of gesture for communication, and visual attention may be the key to understanding why this occurs.

References

Cassell, J., McNeill, D., & McCullough, K.-E. (1998). Speech-gesture mismatches: Evidence for one underlying representation of linguistic and nonlingustic information. *Pragmatics & Cognition*, *6*.

Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism*, *7*, 303–319.

Dargue, N., & Sweller, N. (2020). Learning stories through gesture: Gesture's effects on child and adult narrative comprehension. *Educational Psychology Review*, *32*, 249-276.

Drijvers, L., Vaitonytė, J., & Özyürek, A. (2019). Degree of language experience modulates visual attention to visible speech and iconic gestures during clear and degraded speech comprehension. *Cognitive Science*, *43*, e12789.

ELAN (Version 6.0) [Computer software]. (2020). Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. Retrieved from https://archive.mpi.nl/tla/elan

Genesee, F., Tucker G. R., & Lambert, W.E. (1975). Communication skills of bilingual children. *Child Development, 46*, 1010-1014.

Gullberg, M., & Holmqvist, K. (2006). What speakers do and what addressees look at. *Pragmatics and Cognition Pragmatics & Cognition, 14*, 53-82.

Gullberg, M., & Kita, S. (2009). Attention to Speech-Accompanying Gestures: Eye Movements and Information Uptake. *Journal of Nonverbal Behavior, 33*, 251-277.

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*, 1175–1192.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*, 297–315.

Kartalkanat, H., & Göksun, T. (2020). The effects of observing different gestures during storytelling on the recall of path and event information in 5-year-olds and adults. *Journal of Experimental Child Psychology*, *189*, 104725.

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: now they need it, now they don't. *Journal of Nonverbal Behavior*, *24*, 131–150.

McNeill, D. (1992). Hand and mind. In *The University of Chicago Press*.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: comparison to monolingual norms. *Language Learning, 43*, 93–120.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development, 3*, 359–400.

Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science, 21*.

Yow, W. Q. (2013). Monolingual and bilingual children's use of gestures and grammatical agreement in pronoun interpretation. In S. Baiz, N. Goldman, & R. Hawkes (Eds.), *BUCLD 37 Proceedings* (pp. 482-94). Somerville, MA: Cascadilla Press.