The Relationship between Co-speech Gesture Production and Macrolinguistic Discourse Abilities in People with Focal Brain Injury

Seda Akbıyık[1], Ayşenur Karaduman[1, 2], Tilbe Göksun[1], & Anjan Chatterjee[3]

[1]Koç University, Istanbul - Turkey

[2]Bilkent University, Ankara -Turkey

[3]University of Pennsylvania, Philadelphia, PA - USA

Abstract

Brain damage is associated with linguistic deficits and might alter co-speech gesture production. Gesture production after focal brain injury has been mainly investigated with respect to intrasentential rather than discourse-level linguistic processing. In this study, we examined 1) spontaneous gesture production patterns of people with left hemisphere damage (LHD) or right hemisphere damage (RHD) in a narrative setting, 2) the neural structures associated with deviations in spontaneous gesture production in these groups, and 3) the relationship between spontaneous gesture production and discourse level linguistic processes (narrative complexity and evaluation competence). Individuals with LHD or RHD (17 people in each group) and neurotypical controls (n = 13) narrated a story from a picture book. Results showed that increase in gesture production for LHD individuals was associated with less complex narratives and lesions of individuals who produced more gestures than neurotypical individuals overlapped in frontal-temporal structures and basal ganglia. Co-speech gesture production of RHD individuals positively correlated with their evaluation competence in narrative. Lesions of RHD individuals who produced more gestures overlapped in the superior temporal gyrus and the inferior parietal lobule. Overall, LHD individuals produced more gestures than neurotypical individuals. The groups did not differ in their use of different gesture forms except that LHD individuals produced more deictic gestures per utterance than RHD individuals and controls. Our findings are consistent with the hypothesis that co-speech gesture production interacts with macro-linguistic levels of discourse and this interaction is affected by the hemispheric lateralization of discourse abilities.

Keywords: gesture, focal brain injury, macrolinguistic abilities, narrative complexity, narrative evaluation

1. Introduction

The integration of gesture and speech when people communicate starts early in development (e.g., Capirci, Iverson, Pizzuto, & Volterra, 1996). Gesture and speech also share overlapping neural substrates (e.g., Bernard, Millman, & Mittal, 2015; Gentilucci & Volta, 2008; Green et al., 2009; Willems & Hagoort, 2007; Xu et al., 2009). After focal brain injury, gesture might compensate for or facilitate speech in cases of verbal impairment (e.g., Ahlsén, 1991; Akhavan, Göksun, & Nozari, in press; Akhavan, Nozari, & Göksun, 2017; Cicone et al., 1979; Cocks, Hird, & Kirsner, 2007; Feyereisen, 1983; Glosser, Wiener, & Kaplan, 1986; Göksun, Lehet, Malykhina, & Chatterjee, 2015; Herrmann et al., 1988; Hogrefe, Ziegler, Weidinger, & Goldenberg, 2012; Lanyon & Rose, 2009; Pritchard, Dipper, Morgan, & Cocks, 2015; Scharp, Tompkins, & Iverson, 2007). Because most studies focus on people with injury to one or the other hemisphere and because they typically use different methods in their approach, it has been difficult to directly compare left and right hemispheres' contributions to gesture-speech interactions (for a review, see Hogrefe, Rein, Skomroch, & Lausberg, 2016).

Few studies have compared the two hemispheres' roles in gesture (e.g. Blonder et al., 1995; Göksun et al., 2015; Göksun, Lehet, Malykhina, & Chatterjee, 2013; Hadar, Burstein, Krauss, & Soroker, 1998; Hogrefe et al., 2016; Rousseaux, Daveluy, & Kozlowski, 2010). These studies have not settled uncertainties about relative importance of each hemisphere in co-speech gesture production. To address this gap in literature, we investigated the relationship between gesture use and discourse of individuals with left hemisphere damage (LHD) and right hemisphere damage (RHD). Gesture and narrative production was elicited via a story-telling task. In different analyses of the same task, first, we compared the three groups for gesture types used in this narrative task. Second, we investigated the neural correlates of deviations in spontaneous gesture production in a narrative setting. Third, we asked if gesture use was related to macro-level discourse competence in spontaneous narratives.

Only a few studies have investigated specific brain areas associated with spontaneous gesture production (e.g., Bernard et al., 2015; Göksun et al., 2013, 2015; Hogrefe et al., 2017). Gesture production after brain damage is frequently investigated in terms of comprehensibility or quality of gestures (Hogrefe et al., 2017) or for a specific linguistic category (e.g., spatial language gestures in Göksun et al., 2013, 2015). Studies on neural correlates of gestures usually rely on settings where people produce pantomimes on command or tool-related gestures (e.g., Buxbaum, Shapiro & Coslett; Häberling, Corballis, & Corballis, 2016). However, co-speech gestures are different than emblems, pantomime gestures, or self-touch and grooming as co-speech gestures rely on the speech for their occurrence (McNeill, 1992). Besides, such artificial settings do not provide an ecologically valid context for studying spontaneous co-speech gesture production (for a discussion, see Bernard et al., 2015.

Subtypes of gestures also differ with respect to the functions they serve during story telling. Yet, the literature on the neural substrates of gestures usually reports content-carrying gestures alone (Bernard et al., 2015; Leonard & Cummins, 2011). In that sense, including gesture subtypes that are less emphasized in the literature (e.g., beat gestures) in a naturalistic setting would extend our understanding of the relationship between spontaneous co-speech gestures and discourse-level competence, as well as its neural substrates.

According to "*Kendon's continuum,*" gestures can be classified based on the degree of their symbolism and the need for accompanying speech. At the two extremes of this continuum lie gesticulations and sign language (McNeill, 1992). Gesticulations are the most frequent type of gestures and are usually produced with speech (Kendon, 1988). McNeill and Levy (1982) divided gesticulations into four categories as *iconic*, *metaphoric*, *deictic,* and *beat*.

Iconic gestures are related to the semantic content of speech such that their form depicts the meaning by similarity. They support story-telling by illustrating a concrete event or object in the narrative (Kendon, 1988). *Metaphoric gestures* also convey meaning, but they do so

indirectly as in a non-literal way. *Beats* are rapid hand movements that are timed to the prosody of speech (McNeill & Levy, 1982). These subtypes of gestures lack semantic content, but they can be used to emphasize semantic content by putting emphasis on the word (Andric & Small, 2012). Finally, *deictic gestures* are produced with an extended index finger to point to an object that does not have to be physically present (Cassell & McNeill, 1991). Some classifications also include *emblems,* which are conventional gestures (e.g., Kong et al., 2015; McNeill, 1992). Iconic and deictic gestures are categorized as *representational gestures* (Kita, 2014; McNeill, 1992) whereas emblems and beats belong to the category of *non-representational gestures* (McNeill, 1992). In this paper, we will not define metaphoric gestures as a separate category because they can be regarded as a form of iconic gesture with abstract content (Kita, 2014). Gesticulations, co-speech gestures that comprise an integral part of the language context, are the focus of the present study. When we use the term gesture without any specification, we refer to co-speech gestures that bear a meaning-related or pragmatic relation to the accompanied speech.

One further subdivision of *deictic gestures* is important particularly for narration. *Concrete deixis* occurs when one points to a real object in space (McNeill, 1992) whereas *abstract deixis* occurs when the narrator assigns meaning to a specific location in the gesture space (Cassell & McNeill, 1991; McNeill, Cassell, & Levy, 1993). These pointing gestures might serve different functions. *Concrete deixis* can replace linguistic forms (Bangerter & Louwerse, 2005) and ease identification of referents in speech (Bangerter, 2004) whereas *abstract deixis* serves a metanarrative function (Kendon, 2004; McNeill, 1992). The interconnected nature of gesture and speech and the idea that gestures might serve as an additional channel to convey meaning or replace linguistic forms after speech impairment (e.g. Hadar & Butterworth, 1997; Lanyon & Rose, 2009) makes the study of gesture speech interaction a valuable area to examine.

1.1.Gesture-Speech Interaction after Speech Impairment

The study of individuals with aphasia forms an important part of our understanding of the relationship between gesture and speech (e.g. Ahlsén, 1991; Cicone et al., 1979; Cocks, Dipper, Pritchard, & Morgan, 2013; Dipper, Cocks, Rowe, & Morgan, 2011; Hadar et al., 1998; Pritchard, Cocks, & Dipper, 2013; Wilkinson, 2013). If co-speech gestures and speech are produced by a common system of processing, then one would expect impairments in one system to be reflected in the other (see Eling & Whitaker, 2009; McNeill, 1985, 1992; McNeill et al., 2008). For example, patients with aphasia use structurally simpler and less diverse gestures compared to patients without aphasia and their severity of aphasia correlates negatively with the efficacy of their use of gesture as a communicative tool (Cicone et al., 1979; Glosser et al., 1986). This evidence is in line with Growth-Point Theory (McNeill, 1985, 1992), which states that gesture and speech are different aspects of a unified process. Growth point is defined as the smallest unit of thought that incorporates speech and gesture. Speech refers to the analytic and hierarchical component of the message that is to be conveyed, whereas gesture uses a synthetic and global mode of representation (McNeill et al., 2008). This model of discourse-level linguistic processing predicts a concomitant decrease in use of gestures with impairment in the discourse structure (Cocks et al., 2007).

Conversely, if gesture and speech depend on different, even if related systems, linguistic abilities and gesture use could dissociate. People with aphasia sometimes rely on meaningful gestures more than healthy controls (Ahlsen, 1991; Akhavan et al., in press; Béland & Ska, 1992; De Ruiter, 2006; Goodwin, 1995; Göksun et al., 2015; Herrmann et al., 1988; May, David, & Thomas, 1988; Sekine & Rose, 2013; Wilkinson, 2013). These studies suggest that gesture can compensate for verbal communication deficits. Furthermore, if gestures are *restorative*, rather than compensatory, they could even facilitate speech production (e.g. Krauss, Chen, & Gottesman, 2001; Rauscher, Krauss, & Chen, 1996).

The Lexical Facilitation Model (Hadar & Butterworth, 1997) and the Interface Model (Kita & Özyürek, 2003) postulate that speech and gesture depend on different but related systems. According to the Lexical Facilitation Model (Hadar et al., 1998; Hadar & Butterworth, 1997), some features of conceptual processing activate visual imagery, which in turn links conceptual processing and iconic gesture production. This model predicts better verbal output as use of representational gestures increase (Butterworth & Hadar, 1989).

The Interface Model (Kita & Özyürek, 2003) also suggests that gesture and speech result from distinct modes of thinking with a common communicative intent. One of the main functions of gestures is to support the organization of thoughts into smaller units such that they can be verbally expressed conveniently (Kita & Özyürek, 2003). Thus, greater conceptual demands could result in more gestures (Alibali, Kita, & Young, 2000; Hostetter, Alibali, & Kita, 2007; Kita, Alibali, & Chu, 2017). This model can accommodate compensation, because spontaneous gestures might provide an alternative means of communication for the existing conceptual message when the individual cannot retrieve the appropriate words (Göksun et al., 2013; Göksun et al., 2015; Hogrefe et al., 2017).

1.2.Neural Underpinnings of Gesture

The idea that gesture and speech are intimately connected is supported by neuroimaging evidence (Bernard et al., 2015; Green et al., 2009; Xu et al., 2009). Broca's area, a core language processing area, is implicated in both production and comprehension of gesture (Gentilucci & Volta, 2008; Häberling, Corballis, & Corballis, 2016). Even though people with aphasia have been a focus in the studies on neural correlates of gesture production (e.g. Ahlsén, 1991; Akhavan et al., 2017; Cicone et al., 1979; Dipper et al., 2011; Hadar et al., 1998; Pritchard et al., 2013; Wartenburger et al., 2010; Wilkinson, 2013), relatively few studies have investigated specific brain areas associated with spontaneous co-speech gesture production (e.g. Göksun et al., 2013, 2015; Hogrefe et al., 2017).

The main reason for this dearth in the literature on the neural correlates of spontaneous gesture production is the motion artifacts that arise in neuroimaging paradigms. Due to this problem, neural underpinnings of gesture are usually studied with respect to gesture-speech integration in comprehension instead of production (Kircher et al, 2009; Straube et al., 2011). These attempts to reduce neuroimaging artifacts create ecologically limited environments that do not capture the crucial components of spontaneous co-speech gesture production (Bernard et al., 2015). Due to these constraints that arise during the neuroimaging, associating gesture production patterns with neuropsychological evidence or structural properties of the brain areas of interest provide a promising line of inquiry (Bernard et al., 2015; Wartenburger et al., 2010).

Although both hemispheres contribute to the gesture production process, we observe left hemisphere dominance in the production of gestures that are tied to the conceptual processes and the communication of spatial information (see Hogrefe et al., 2016). This might be due to the suggestion that the left hemisphere is responsible for the conceptualization of representational gestures (Helmich & Lausberg, 2014). The importance of left hemisphere for tool-related and imitational gestures also provides support for this notion (Goldenberg & Randerath, 2015). Conversely, right hemisphere damage is associated with the deficits of rhythmic gestures (Blonder et al., 1995; Cocks et al., 2007; Hogrefe et al., 2016; Rousseaux et al., 2010) along with other nonverbal and non-motor aspects of communication such as intonation and prosody (Lausberg, Zaidel, Cruz, & Ptito, 2007; Shapiro & Danly, 1985). Right hemisphere damage has even been associated with reduced use of representational gestures (McNeill & Pedelty, 1995).

Previous research also suggests that damage to the inferior frontal lobe, anterior temporal lobe, and supramarginal and angular gyrus of the left hemisphere are associated with deficiencies in the production of gestures that carry semantic content (Göksun et al., 2015; Hogrefe et al., 2017). Damage to the middle frontal gyrus and posterior middle temporal gyrus

of the left hemisphere is associated with deficiencies in tool-related gesture production. On the other hand, imitation of novel and meaningless gestures relies on the integrity of the left inferior parietal and somatosensory cortices, which mostly depend on somatosensory and visuospatial processing (Buxbaum et al., 2014). The deficiencies in gesture production that are associated with the left anterior temporal lobe and the left inferior frontal regions are attributed to semantic retrieval deficiencies and problems in executive control when choosing between alternate modes of communication (Hogrefe et al., 2017). These areas are implicated in structural (Bernard et al., 2015; Wartenburger et al., 2010) and functional (Willems & Hagoort, 2007; Willems, Özyürek, & Hagoort, 2007; Straube et al., 2011) neuroimaging studies of gesture comprehension conducted with neurotypical individuals as well. By analyzing structural morphology in relation to co-speech gesture production, Wartenburger and colleagues (2010) also revealed that the production of conceptual gestures is related to cortical thickness in the left temporal lobe and inferior frontal gyrus.

Göksun and colleagues found neural evidence for the link between gesture and language in the context of impaired spatial preposition and motion event production. In two studies, they reported that while people with damage to the left posterior middle and inferior frontal gyri could not use gesture to compensate for verbal deficiencies, people with left anterior superior temporal gyrus damage produced more gestures than expected, presumably to compensate for their speech problems (Göksun et al., 2013, 2015). These findings support the idea that semantic information conveyed through gestures can be interrupted with damage to the inferior frontal network, but is not affected by superior temporal gyrus damage (Göksun et al., 2015).

Left inferior frontal damage impairs the comprehensibility or quality of gestures (e.g. Hogrefe et al., 2017), or specific linguistic content such as spatial language (e.g. Göksun et al., 2013, 2015). However, the impairments might be specific to the level of analysis or to the types of gestures that are investigated. We need to assess a wider range of gestures accompanied by

a thorough analysis of discourse for a fuller understanding of how damage to certain regions of the brain links to co-speech gesture production.

1.3.Neural Underpinnings of Narrative Production

A narrative is a sequence of temporally related clauses expressed from a particular point of view (Reilly, Bates, & Marchman, 1998). Narratives are organized typically around the actions of agents and can be evaluated at both micro- and macro-linguistic levels (Karaduman, Göksun, & Chatterjee, 2017). Micro-linguistic level refers to the analysis of lexical and grammatical properties of a sentence. Macro-linguistic analysis refers to the relationship between different utterances and the global understanding of the entire narrative (Kintsch, 1994). At the macro-linguistic level, a narrative includes referential or evaluative information (Labov & Waletzky, 1997). The referential function of a narrative informs the audience about temporal features during events. This requires the narrator to order the events linearly along a *horizontal axis* (Bamberg & Marchman, 1991). Informing the audience about the actions of the characters in a successive manner would be an example of referential information (e.g. "*the boy falls over the cliff and the dog joins him there*"). However, events can also be ordered along a *vertical axis,* where different events are integrated around a global theme referring to the evaluation of the narrative (Küntay & Nakamura, 2004; Labov & Waletzky, 1997). In that sense, making inferences regarding the motivations or emotional states of characters can be taken as an example of evaluative information (e.g. "*the owl is mad cause the boy woke him up"*).

People with LHD are usually impaired at the micro-linguistic level, with lower mean length of utterances (e.g. Karaduman et al., 2017), fewer motion sentences (Göksun et al., 2015), and they use more indefinite words (Ulatowska, North, & Macaluso-Haynes, 1981). The results are more complicated regarding their macro-linguistic abilities (see Andreetta, Cantagallo, & Marini, 2012). For instance, individuals with fluent aphasia might exhibit

problems of cohesion and thematic informativeness (Andreetta, 2014) or discourse organization (Kaczmarek, 1984). However, others report that people with LHD maintain macro-linguistic abilities similar to neurotypical individuals (see Pritchard et al., 2015).

The linguistic competence of people with RHD is usually investigated with respect to emotional processing of prosody (e.g. Blonder et al., 1995; Blonder, Bowers, & Heilman, 1991), understanding of pragmatics, and importantly for this study, inter-sentential or global level discourse (e.g. Bartels-Tobin & Hinckley, 2005). Even though people with RHD can bind discrete linguistic units, they usually have problems in *globally connecting* narratives (Ash et al., 2006). Thus, their macro-level deficit is evident when the narrative has to be comprehended coherently (for a review, see Mar, 2004; Ross & Mesulam, 1979; Sherratt & Bryan, 2012). People with right frontotemporal injury also demonstrate organizational problems in discourse (Ash et al., 2006; Joanette & Goulet, 1990; Joanette, Goulet, Ska, & Nespoulous, 1986).

In a recent paper, Karaduman, Göksun, and Chatterjee (2017) investigated micro-linguistic and macro-linguistic abilities of people with LHD and RHD. They found that the LHD group had problems maintaining the story theme and used fewer evaluative devices than the neurotypical control group. The RHD group performed well, but individual cases with lesions to the dorsolateral prefrontal cortex, the anterior and superior temporal gyrus, the middle temporal gyrus and supramarginal gyrus of the right hemisphere produced less complex narratives.

1.4.The Current Study

In sum, gesture production is associated with both micro-linguistic (e.g. Krauss & Hadar, 1999; Mol & Kita, 2012) and macro-linguistic (e.g. Azar & Özyürek, 2016; Kendon, 1994; McNeill, 1992; McNeill & Levy, 1993; Nicoladis, Marentette, & Navarro, 2016; Nicoladis, Pika, Yin, & Marentette, 2007) levels of discourse processing. Any general claim about the relationship between speech and gesture needs to take these different levels of

discourse into consideration. Thus, a comprehensive investigation of narrative and gestures in people with brain injury is needed (for a relevant discussion, see Cocks et al., 2007).

We investigated gesture use of LHD and RHD individuals in relation to both standard neuropsychological measures of verbal impairment and macro-level discourse abilities with a special focus on the neural substrates that are associated with this interaction. Here, we focused on referential (i.e., narrative complexity) and evaluative functions of narrative. Narrative complexity relates to the content of discourse while evaluation competence relates to pragmatic and emotional processing. In line with the Interface Model, we expect that people who produce less complex narratives would use more representational gestures, as a compensatory strategy since representational gestures, especially deictics, can replace linguistic forms (Bangerter & Louwerse, 2005). However, gestures also contribute to non-verbal communication and pragmatic processing (e.g., Colletta et al., 2015; Enfield, Kita, & De Ruiter, 2007; Holler & Beattie, 2003; Wu & Coulson, 2007). Thus, we expect that gesture use in general would be positively related to evaluation competence. That is, regardless of injury, people who use more evaluative verbal devices would produce more gestures.

Second, we investigated how frequently individuals with LHD or RHD use different types of gestures. As mentioned above, Growth-Point Theory predicts parallel deficiencies in gesture production and speech. By contrast, the Lexical Facilitation Model (Hadar & Butterworth, 1997) and the Interface Model (Kita & Özyürek, 2003) predict intact or even increased use of gestures by people with focal brain injury. The Lexical Facilitation Model and Interface Model do not yield different predictions because both allow verbal and nonverbal modalities to dissociate. However, if representational gestures indeed facilitate lexical retrieval as the Lexical Facilitation Model suggests then one would expect a positive relationship between representational gesture production and verbal abilities. The Interface Model, on the other hand, predicts compensatory use of representational gestures with verbal impairment,

thus, we would expect a negative relationship between verbal abilities and representational gesture use.

Regarding non-representational gestures, beats do not share biological substrates with language (Bernard et al., 2015) and are not affected by the content of speech (Kita, 2014). The use of beat gestures should not be affected by the conceptual demands of the task either (Alibali, Heath, & Myers, 2001). However, beats do relate to processing narrative (Krahmer & Swerts, 2007; McNeill, 1992), suggesting that beat gestures might be affected by deficits of discourse-level processing. Since beat gestures serve metanarrative functions such as specifying the transition points in the discourse structure (McNeil, 1992), there could be a positive relationship between beat gesture production and discourse-level competence in narratives.

As to the possibility of compensation, there is evidence that while individuals with severe deficits in nonverbal semantic processing (Hogrefe et al., 2012) and semantic knowledge (Cocks et al., 2013) cannot use nonverbal tools efficiently for interpersonal communication, patients with intact semantic processing can use gesture as a compensatory strategy for communication (Göksun et al., 2015; Hogrefe et al., 2012). Thus, we expected that people with damage to language-related areas of the brain (the left temporal lobe and left inferior frontal cortex) to produce more gestures to compensate for their verbal impairment, provided that they do not have semantic deficiencies.

2. Materials and Procedure

2.1.Participants

Individuals with left hemisphere damage (age range: 37-79, N = 17, M_{age} = 64.94), right hemisphere damage (age range: 45–87, N = 17, M_{age} = 63.65), and age-matched healthy elderly controls (age range: 38-77, N = 13, M_{age} = 60.85) participated in the study. All participants were native speakers of English and were right-handed except one LHD individual. The three groups did not differ in age, $F(2, 44) = 0.44$, $p > .05$, however, they differed in years of education, $F(2,$

44) = 3.75, p = .03. Bonferroni t-tests revealed that the control group (M = 16.00, SD = 2.12) had more years of education than LHD individuals (M = 13.41, SD = 1.97) (Bonferroni, p = .03). Demographic information of the individuals with RHD and LHD are shown in Table 1 and lesion overlaps can be seen in Figure 1. Patients were recruited from the Focal Lesion Database at the University of Pennsylvania (Fellows, Stark, Berg, & Chatterjee, 2008); a database which excludes patients with a history of neurological disorders, psychiatric disorders or substance abuse. The RHD and LHD individuals did not differ in lesion size, t(28) = -0.86, p > .05. All participants provided written informed consent and the data analyzed in this paper were collected in accordance with the University of Pennsylvania's Institutional Review Board. Participants were compensated $15/h for their volunteer contribution to the study.

Table 1

Demographic and neuropsychological data of people with LHD and RHD.

Patient	Gender	Age	Education (years)	Lesion Side	Location	Lesion Size (# of voxels)	Cause	Chronicity	WAB (AQ)	OANB (Action)	OANB (Object)
LT_85	F	63	15	L	I	13079	Stroke	177	98.8	100	98.8
KG_215	M	61	14	L	F	17422	Stroke	145	94.4	96	93.8
TO_221	F	77	13	L	O	5886	Stroke	160	100	100	100
BC_236	M	65	18	L	FP	155982	Stroke	210	80.8	88	94
XK_342	F	57	12	L	OT	42144	Stroke	125	91.4	94	93
TD_360	M	58	12	L	T BG	38063	Stroke	118	65.3	52	28
IG_363	M	74	16	L	F	16845	Stroke	117	91.4	96	95
KD_493	M	68	14	L	T	22404	Aneurysm	101	92.1	98	95
DR_529	F	66	12	L	PA F	8969	Stroke	100	94.9	94	90.1
DR_565	F	53	12	L	PA F	14517	Aneurysm	103	99.8	98	97.5
MC_577	F	79	11	L	C	4191	Stroke	50	85.3	82	79
NS_604	F	37	12	L	PO	79231	AVM	113	100	100	98

UD_618	M	77	15	L	F	48743	Stroke	47	89.4	76	85
KM_642	M	77	12	L	P	7996	Stroke	109	96.8	94	98
MR_644	F	74	12	L	C	-	Stroke	-	-	-	-
CC_749	F	71	12	L	P	34266	Stroke	50	88.8	-	-
FC_83	M	70	12	R	FTP	8040	Stroke	169	99.8	96	98
NC_112	F	48	16	R	O	4733	Stroke	178	100	98	-
HX_252	M	77	12	R	MCA	-	Stroke	-	94.6	78	85
RT_309	F	66	21	R	T	79691	Hematoma	128	-	-	-
DF_316	F	87	12	R	P	2981	Stroke	126	97.1	88	93
DC_392	M	56	10	R	PT	39068	Stroke	108	97.6	98	95
DX_444	F	80	12	R	PT	41172	Stroke	106	95.5	94	93
TS_474	F	51	11	R	P	22208	Stroke	100	95.1	98	95
UD_550	F	47	16	R	C	-	Stroke	-	-	-	-
NS_569	F	72	18	R	FT BG	37366	Stroke	77	100	100	99
DG_592	F	45	12	R	PT	130552	Stroke	127	97.8	98	98
KG_593	F	49	12	R	FTP BG	170128	Stroke	58	100	90	95
KS_605	M	63	18	R	C	23217	Stroke	76	98.8	100	100
ND_640	F	70	18	R	PT	64603	Stroke	54	96.8	100	100
CS_657	M	75	18	R	PO	33568	Stroke	43	99.2	98	100
KN_675	M	64	18	R	FT	23779	Stroke	32	-	-	-
MN_738	F	62	16	R	C	32154	Stroke	25	98.4	100	100
DD_755	F	48	16	L	C	-	Stroke	-	-	-	-

Key: F: Frontal; T: Temporal; P: Parietal; O: Occipital; BG: Basal Ganglia; C: Cerebellum; I: Insula; Pe: Perisylvian; PA: Pericallosal artery; ACA: Anterior Cerebral Artery; MCA: Middle Cerebral Artery; AVM: arteriovenous malformations. WAB- AQ indicates a composite language score with a maximum possible score of 100. OANB (action) and OANB (object) demonstrate knowledge of verbs and nouns with a maximum possible score of 100.

2.2.Materials and Procedure

2.2.1. Neuropsychological tasks

All neuropsychological tasks were completed during a different session than the narrative task. Patients were administered Western Aphasia Battery (WAB) (Kertesz, 1982) to investigate their level of speech production and Object and Action Naming Battery (OANB) (Druks, 2000) to ensure that the problems they experienced were not due to deficiencies in naming different classes of words.

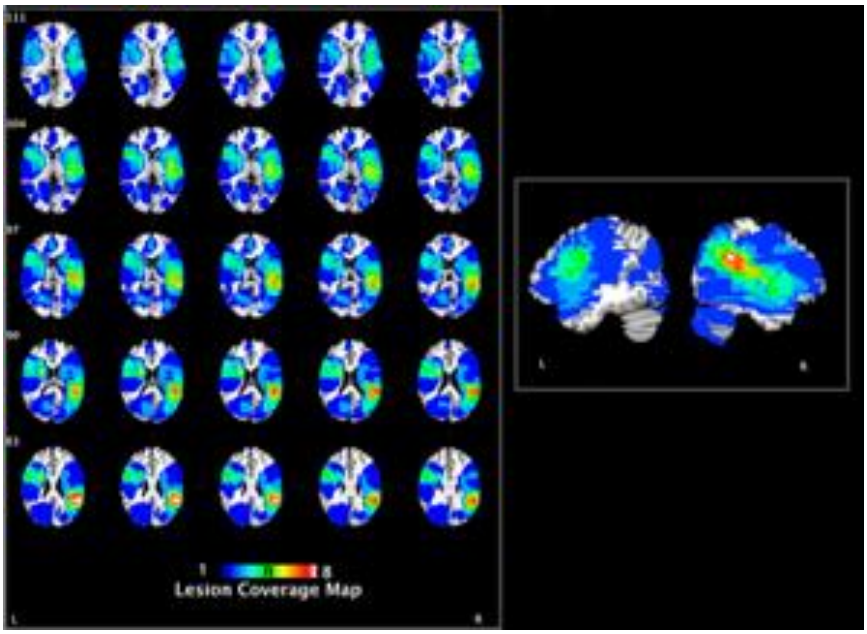


Figure 1. Coverage map indicating the lesion locations for all participants with brain damage.

2.2.2. Narrative Task and Gesture Elicitation

Discourse and gesture production were elicited using "*Frog where are you*?" (Mayer, 1969), a pictured story-book which has been widely used to assess narrative competence in many languages and across a variety of populations (for a review see Reilly et al., 2004). The book consists of the adventures of a boy and his pet dog while they are searching for their frog. The book requires the participants to understand what is going on based on the visual depictions of events and requires the narrator to make inferences about the actions and motivations of the agents. It allows assessing different linguistic abilities of people (Reilly et al., 1998, 2004). Participants were asked to explain what was happening in each page while the book was in front of them. In this way, we restricted memory demands on their performance.

2.2.3. Procedure

All participants were tested individually at their home or in the laboratory. At the beginning of the procedure, participants spent time to familiarize themselves with the book. When they were ready, they were asked to describe what was happening in each page while the experimenter held the book for them. The experimenter did not mention the use of gesture. The

experimenter who listened to the story telling was always the same. Participants were videotaped for later transcriptions of the narratives and gestures as they told the story. Native English speakers transcribed the narratives.

2.2.4. Coding

2.2.4.1. Narrative

The total number of utterances, words, nouns, and verbs as well as the mean length of utterances (MLU) were coded for each participant to examine the general characteristics of their discourse production (for the details of narrative coding, see Karaduman et al., 2017, Appendix A-B). Narrative complexity as formulated by Berman and Slobin (1994) was coded based on the presence of three main plot components; plot onset, plot unfolding and plot resolution with an additional coding for the search theme to assess the extent to which participants understood the theme of the story (see Appendix A).

The narratives of the participants were also coded with respect to evaluation based on a scheme Karaduman et al. (2017) created, which is informed by Reilly et al. (2004), Küntay and Nakamura (2004), and Bamberg and Damrad-Fyre (1991). Subcategories of this scheme were as follows: cognitive inferences, social engagement devices, references to affective states or behaviors, enrichment expressions and hedges. The scores in these subcategories were summed resulting in a total evaluation score for each participant. An evaluative diversity index was also obtained to capture the number of evaluation subcategories used by each subject (see Appendix B for the details of the scheme and examples for the subcategories).

2.2.4.2.Gesture

We transcribed the participants' use of spontaneous gestures during their narration. Gesture forms were created by adopting and combining frameworks suggested by McNeill (1992) and Kong et al. (2015) (for detailed descriptions and examples of gesture categories, see Appendix C). We defined four main forms for gestures: iconic, deictic, emblem, and beat.

Iconic gestures were further coded as static or dynamic. For the deictic gestures, we classified whether the deictic gesture was *concrete* (as occurred as pointing to the book) or *abstract* (see Appendix C for the gesture coding scheme and examples). We also divided the gestures into representational (iconic and deictic) and non-representational gestures (beats and emblems).

All gestures were coded by the first author. A second coder coded 19% of the participants' gestures that were randomly selected for reliability. Pearson's correlation showed a significant association between the judgments of the coders on the frequency of gestures, (r = .98, p < .001). Cohen's κ confirmed high degree of agreement in judgments of the two coders to the different types of gestures, κ = .813, p < .001. Disagreement between coders were resolved after discussions.

Gesture per utterance was calculated for each participant by dividing the total number of gestures to the total number of utterances. This metric was used to indicate the frequency of gesture production with respect to the natural flow of speech. Since the length of speech is positively associated with gesture frequency (Nicoladis et al., 2010), we used this standardization procedure to overcome possible confound of speech duration on the frequency of gestures. Calculating gesture frequency per interval is another conventional method in the studies of co-speech gesture production. However, since brain-damaged patients, especially LHD individuals tend to speak slowly and pose a lot, relying on the duration of gesture could be misleading (see Feyereisen, 1983). Analyses of forms of gestures (e.g., iconic) were done based on their proportion to the total number of gestures each participant used. This measure assessed the importance of a specific type of gesture among other gestures. Additionally, we divided the total number of gestures for each category by the total number of utterances and used this measure to indicate the frequency that gesture category was used during discourse. These two measures were used to test the claims regarding the patterns of gesture production

after brain injury and for further analyses when we investigated the relationship between narrative competence and general or category-specific gesture use.

After classifying each gesture, we also coded whether these gestures were accompanied by an evaluative device (e.g., The participant says, *'he tries to cry'* (i.e., uses reference to affective states) *while moving her hand repetitively* (i.e., a beat gesture). We investigated the proportion of evaluative devices used with gestures. However, we also investigated specific types of gestures that accompany evaluative devices to see which gesture accompanies the evaluative devices the most.

3. Results

3.1.Overall Gesture Production

First, we tested whether overall gesture frequency and rate per utterance differed across the three participant groups. There was a significant negative correlation between the years of education and the total number of gestures (r = -.31, p = .03) and gesture by utterance (r = -.33, p = .03) participants produced. Given that the groups differed with respect to their years of education and their gesture use was associated with education level, we controlled for the education variable in the whole-group analyses. A one-way analysis of covariance (ANCOVA) revealed a main-effect of group for the total number of gestures produced, $F(2, 43) = 3.49$, p = .04, η^2 = 0.12, and the gestures per utterance, $F(2, 43) = 4.39$, p = .02, η^2 = 0.15. Bonferroni t-tests revealed that the control group (M = 10.62, SD = 18.41) produced significantly fewer gestures in total than the individuals with LHD (M = 52.41, SD = 45.37) (Bonferroni, p = .04). There was no difference between controls and individuals with RHD (M = 28.59, SD = 31.55) (Bonferroni, p = .73). The production of gesture by utterance of the control group (M = 0.17, SD = 0.24) was also significantly lower than individuals with LHD (M = 1.10, SD = 0.97) (Bonferroni, p = .02), but there was no difference between controls and individuals with RHD (M = 0.55, SD = 0.60) (Bonferroni, p = .26) (see Figure 2).

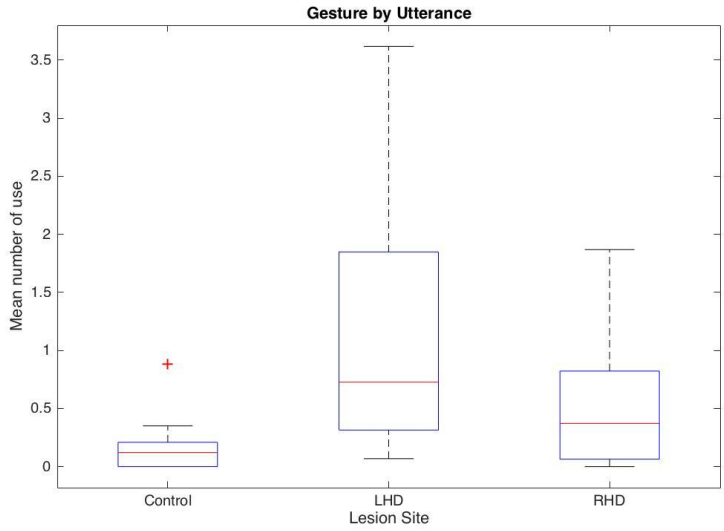


Figure 2. The mean number of gesture by utterance across different participant groups.

The groups did not differ with respect to the proportion of representational/non-representational gestures they produced relative to their overall gesture production, $F(2, 34) = .23$, $p = .80$. However, they differed in the amount of representational gestures they produced relative to their total utterances, $F(2, 43) = 3.70$, $p = .03$, $\eta^2 = .13$. Bonferroni t-tests revealed that LHD group ($M = 0.76$, $SD = 0.78$) produced more representational gestures per utterance than the control group ($M = 0.12$, $SD = 0.18$) (Bonferroni, $p = .01$). There was a main-effect of group in the non-representational gestures per utterance, ($F(2, 44) = 4.51$, $p = .02$), but this effect did not survive after controlling for education, ($F(2, 43) = 2.91$, $p = .06$).

The gesture by utterance and representational gesture production were also tested at the level of individual using Bayesian single-case statistical analysis. This analysis enables the comparison of a case to a comparison group. The parameters of the control group were treated as population parameters and a significant test was conducted accordingly (Crawford & Howell, 1998; for a detailed explanation, see Akhavan et al., 2017). By estimating the percentage of control population obtaining a significantly different score than the case, one can decide whether it is possible for that case to belong to the comparison group. Nine LHD individuals (out of 17) and 6 RHD individuals (out of 17) produced gestures at a higher rate than the

controls, $p < .05$ (see Table 2). Additionally, 9 LHD individuals (out of 17) and 4 RHD individuals (out of 17) produced more representational gestures than the controls, $p < .05$. The same group of people (except one individual) were implicated in increases in gesture by utterance in general and representational gesture production in specific, resulting in the same pattern for lesion overlaps. Lesions of these individuals maximally overlapped in the left inferior frontal gyrus, superior temporal gyrus, putamen-caudate and insula and the right superior temporal and supramarginal (see Figure 3).

Table 2

Single case statistics profile of people with LHD and RHD for gesture by utterance.
(Control sample: n = 13, *M* = 0.17, *SD* = 0.24)

Lesion Site	**ID**	**Gesture by Utterance**	**t**	**Significance test (p-value)**	**Estimated % of control population obtaining higher score than case**
Left	UD_618	1.25	4.261	0.001	0.00
	DR_529	1.00	3.274	0.007	0.00
	MC_577	0.73	2.208	0.047	0.02
	XK_342	2.39	8.763	0.000	0.00
	CC_749	1.84	6.591	0.000	0.00
	LT_85	1.88	6.749	0.000	0.00
	IG_363	2.06	7.460	0.000	0.00
	TD_360	3.62	13.621	0.000	0.00
	KG_215	0.91	2.919	0.013	0.01
Right	DX_444	1.17	3.945	0.002	0.00
	TS_474	1.02	3.353	0.006	0.00
	DF_316	1.73	6.157	0.000	0.00
	KN_675	1.87	6.710	0.000	0.00
	NC_112	0.76	2.326	0.038	0.02
	UD_550	0.73	2.208	0.047	0.02

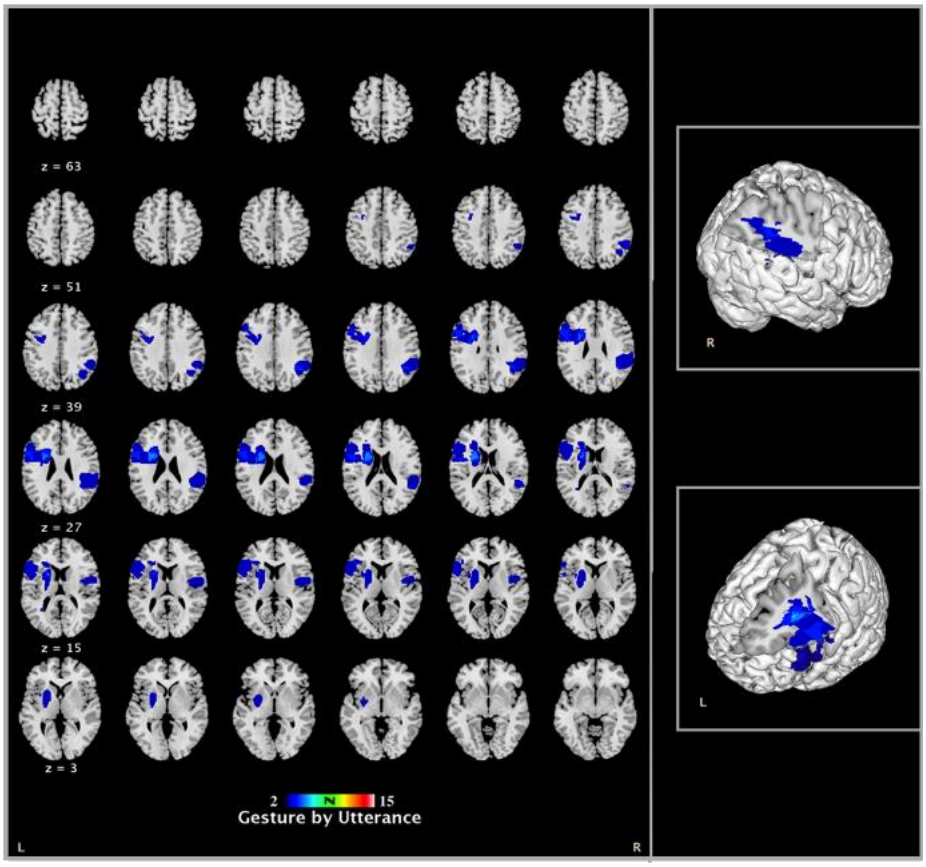


Figure 3. Representative slices based on single-case statistics for gesture by utterance. The maps show the lesion overlaps for 9 LHD and 6 RHD individuals who produced significantly more gestures by utterance than the control comparison group (minimally two participants have lesions on a specific area).

3.2. Production of Different Gesture Forms

To compare whether production of different gesture forms interacted with hemispheric damage, we compared the production of different gesture types across the three participant groups. No main-effect of group was found for the proportion of iconic ($F(2, 33) = 2.29$, $p = .12$), deictic ($F(2, 34) = 0.23$, $p = .79$), emblem ($F(2, 34) = 0.45$, $p = .64$) or beat gestures ($F(2, 34) = 0.34$, $p = .71$) in the total gesture use of the participants (see Table 3 for the raw scores of different forms of gestures, see Figure 4 for the proportion of different gesture subtypes across participant groups). Additionally, the groups did not differ with respect to the proportion of static ($F(2, 34) = 1.29$, $p = .29$) or dynamic ($F(2, 34) = 1.39$, $p = .26$) subcategories of iconic

gestures. However, the groups differed regarding the amount of deictic gestures they produced relative to their total utterances, $F(2, 43) = 4.81$, $p = .01$, $\eta^2 = 0.75$. Bonferroni t-tests revealed that the LHD group ($M = 0.67$, $SD = 0.71$) produced deictic gestures at a higher rate than both RHD individuals ($M = 0.23$, $SD = 0.25$) (Bonferroni, $p = .02$) and control group ($M = 0.08$, $SD = 0.11$) (Bonferroni, $p < .01$). The groups did not differ in the proportion of abstract or concrete deictic gestures ($ps > .05$). There was also no significant effect of group on the number of abstract or concrete deictic gestures participants used per utterance ($ps > .05$).

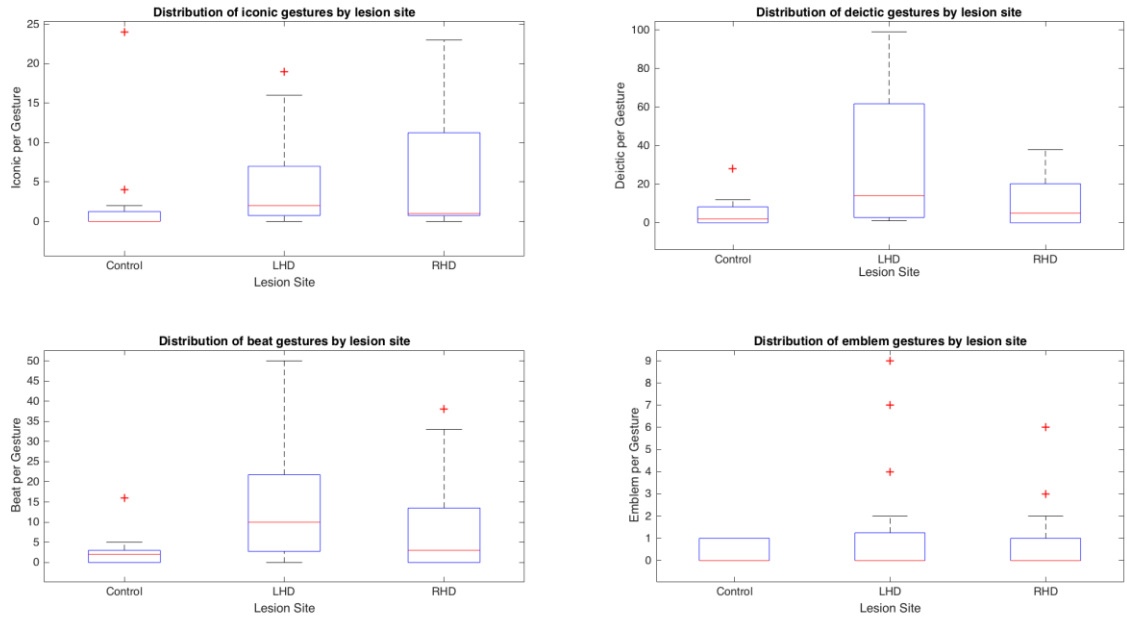


Figure 4. The proportion of gesture subtypes in participants' overall gesture production across participant groups. None of the differences were significant.

Table 3

Raw scores for the specific gesture types.

	Group	Mean	SD	Range
Iconic	Control	2.54	6.55	0-24
	LHD	4.76	6.17	0-19
	RHD	6.53	8.14	0-23
Static	Control	0.54	1.33	0-4

	LHD	0.59	1.06	0-3
	RHD	0.82	1.33	0-5
	Control	2.00	5.45	0-20
Dynamic	LHD	4.18	5.34	0-16
	RHD	5.71	7.13	0-20
	Control	5.23	7.91	0-28
Deictic	LHD	32.29	32.92	1-98
	RHD	12.30	14.06	0-38
	Control	5.08	7.43	0-26
Concrete Deixis	LHD	31.06	32.42	1-98
	RHD	10.81	11.23	0-33
	Control	0.15	0.55	0-2
Abstract Deixis	LHD	1.24	2.41	0-8
	RHD	2.12	3.84	0-13
	Control	2.54	4.35	0-16
Beat	LHD	13.82	13.84	0-50
	RHD	8.94	12.22	0-38
	Control	0.31	0.48	0-1
Emblem	LHD	1.53	2.67	0-9
	RHD	0.82	1.59	0-6

3.3.The Relationship between Gesture Production and Neuropsychological Measurements

Six LHD individuals were diagnosed with anomic aphasia and one LHD participant with Wernicke's aphasia. The LHD group had lower scores on WAB than the RHD group, $F(1, 27) = 7.24$, $p = .01$, $\eta^2 = .20$. There was no significant difference between the groups in their scores on OANB action, $F(1, 26) = 1.56$, $p = .22$, $\eta^2 = .05$, and object subtests, $F(1, 25) = 1.93$, $p = .18$, $\eta^2 = .07$.

To investigate whether spontaneous gesture production was affected by overall verbal impairment, we conducted group-level analyses measuring the relationship between standardized neuropsychological measurements and frequency and type of gestures. Individuals with brain damage who had higher WAB scores produced fewer gestures by utterance ($r = -.63$, $p < .001$). Additionally, the participants' production of gesture by utterance negatively correlated with their scores on Object and Action Naming Battery's action ($r = -.57$, $p < .01$) and object subtests ($r = -.67$, $p < .001$) (see Table 4). When we analyzed the groups separately, for RHD individuals, there were no significant correlations between gesture by utterance and any of the neuropsychological measures ($p < .05$). For the LHD individuals, there was a

significant negative correlation between gesture by utterance and WAB scores (r = -.60, p = .02), OANB action (r = -.61, p = .02) and OANB object (r = -.70, p < .001) scores.

Table 4

Correlations between neuropsychological measures and gesture use for the patient groups.

			2	3	4
1.	WAB	r	.836**	.874**	-.634**
		p	.000	.000	.000
		N	28	27	29
2.	OANB_action	r		.922**	-.566**
		p		.000	.002
		N		27	28
3.	OANB_object	r			-.673**
		p			.000
		N			27
4.	Gesture by Utterance	r			1.000
		p			.
		N			34

** Correlation is significant at the 0.01 level (2-tailed).

When we analyzed the specific gesture types separately, we observed that RHD individuals who produced higher proportions of deictic gestures had also higher scores on WAB (r = .69, p = .02), OANB action (r = .75, p = .007), and OANB object scales (r = .76, p = .01). However, their use of beat gestures was negatively associated with their WAB (r = -.64, p = .03), OANB action (r = -.83, p < .01) and OANB object scores (r = -.75, p = .01). For LHD individuals, the rate of deictic gestures negatively correlated with WAB (r = -.63, p = .01), OANB action (r = -.70, p < .01), and object subtests (r = -.73, p < .01), but no significant association was observed with any other measure.

3.4.The Relationship between Discourse Competence and Neuropsychological Measurements

In this paper, we mainly investigated how overall discourse competence relates to gesture production. We also report some correlations between neuropsychological measures of

verbal impairment and macro-level discourse abilities. There was a positive correlation between narrative complexity and WAB (r = .58, p = .001), OANB action (r = .77, p < .001) and object naming (r = .65, p < .001) scores. Evaluation competence was also positively correlated with WAB (r = .37, p = .05) but no such relationship was observed with OANB action (r = .23, p = .25) and object (r = .28, p = .16) scores. For a comprehensive analysis of the macro-level discourse abilities of the same participants, we refer the reader to Karaduman et al. (2017).

3.5.The Relationship between Gesture Production and Narrative Measures

For the purposes of this study, we focused on the participants' MLU scores and their general performance on narrative complexity and evaluation as indices of macro-level discourse abilities. After controlling for education, the groups differed in their MLU, $F(2, 43) = 4.49$, p = .02, η^2 =.15. LHD individuals (M = 8.63, SD = 0.89) had lower MLU scores than controls (M = 10.42, SD = 1.55; Bonferroni corrected, p = .01). A univariate ANOVA indicated a main effect of group on narrative complexity with the LHD individuals performing worse than controls, $F(2, 44) = 4.11$, p = .02, η^2= .16. However, after controlling for education, the groups did not differ in their narrative complexity, $F(2, 43) = 2.440$, p = .10 , η^2= .09. The groups differed in their scores of evaluative function ($F(2, 43) = 6.27$, p < .01, η^2= .22) and Bonferroni corrected t-tests revealed that LHD individuals (M = 19.41, SD = 9.82) produced fewer evaluative devices than RHD individuals (M = 32.53, SD = 18.85; Bonferroni, p =.03) and controls (M = 37.85, SD = 13.78; Bonferroni, p = .01). However, the groups did not differ with respect to evaluative diversity, $F(2, 43) = 1.93$, p = .16 (for a detailed analysis and interpretation of the narrative measures, see Karaduman et al., 2017).

3.5.1. Gesture Production and Narrative Complexity

For LHD individuals, there was a significant negative correlation between narrative complexity and gesture by utterance (r = -.55, p = .02). When we analyzed representational and non-representational gestures separately, narrative complexity negatively correlated with the

proportion of representational gestures in the overall gesture production (r =-.48, p = .05). The same pattern held for the use of representational gestures per utterance (r = -.68, p < .01), but not for the non-representational gestures (p > .05). When we analyzed subcategories of gestures separately, narrative complexity negatively correlated with the proportion (r = -.63, p = .01) and rate (r = -73, p = .001) of deictic gestures. As to the differentiation between abstract and concrete deictic gestures, the rate of concrete deictic gestures in discourse negatively correlated with narrative complexity for LHD groups alone (r = -56, p = .02). There were no significant correlations between gesture use and narrative complexity for the control group (p > .05). For RHD individuals, no significant association was seen between narrative complexity and gesture use (p > .05).

Having identified a relationship between narrative measures and gesture use as described above, we conducted Bayesian single-case analyses to identify the number of participants displaying similar patterns with each other (Crawford & Garthwaite, 2007; for the details of analyses see Akhavan et al., in press). As there was a negative relationship between narrative complexity and gesture use for LHD individuals, we sought participants who produced less complex narratives, but produced more gestures when compared to the control group. Our analysis revealed that 3 LHD (out of 17) and 1 RHD (out of 17) individuals showed this pattern (see Table 5). Their lesions were in the basal ganglia (putamen, caudate, insula), inferior frontal gyrus, middle frontal gyrus and middle temporal gyrus of the left hemisphere and inferior parietal lobule and middle temporal gyrus of the right hemisphere. No participant from either patient group produced less complex narratives and fewer gestures when compared to controls (p > .05).

Table 5

LHD and RHD individuals who produced less complex narratives but used more gestures than controls.

Control Sample Narrative Complexity			Control Sample Gesture by Utterance					Narrative Complexity	Significance Test	Gesture by Utterance	Significance Test
n	**Mean**	**SD**	**n**	**Mean**	**SD**	**Lesion Site**	**ID**	**Mean**	**p-value**	**Mean**	**p-value**
13	16.31	1.25	13	0.17	0.24	Left	TD_360	3.00	< .001	3.62	< .001
							IG_363	9.00	< .001	2.06	< .001
							CC_749	6.00	< .001	1.84	< .001
						Right	DX_444	11.00	< .001	1.17	.002

3.5.2. Gesture Production and Evaluative Competence

For the control group, gesture by utterance positively correlated with evaluation (r = .56, p = .04). Evaluative diversity in the control group positively correlated with the use of deictic gestures (r = .83, p = .01) and negatively with the use of beat gestures (r = -.90, p < .01). No such relationships were observed for gesture use in general and evaluation competence for the LHD group. For RHD individuals, the total number of gestures, not gesture by utterance, positively correlated with the evaluation score (r = .48, p = .04). There was a positive correlation between evaluative diversity and the rate of representational gestures in general (r = .48, p = .04) and deictic gestures specifically (r = .48, p = .04).

For the total number of evaluative devices that accompanied gestures, there was no main-effect of group, $F(2, 43) = 1.16$, p = .32, but the groups differed with respect to the proportion of evaluative devices that accompany gestures, $F(2, 42) = 4.86$, p = .01, η^2 = .19. Furthermore, the proportion of deictic (M = 43.88, SD = 34.93) and beat (M = 39.73, SD = 32.04) gestures that accompanied evaluative devices were higher than iconic (M = 10.49, SD = 14.08) and emblem gestures (M = 5.38, SD = 13.17) (p > .05). Bonferroni t-tests revealed that the control group (M = 3.91, SD = 6.94) produced a lower proportion of evaluative devices with gestures than the individuals with LHD (M = 35.17, SD = 26.30) (Bonferroni, p = .001). As to the gestures with evaluative devices, single-case statistical analysis revealed only two people, one LHD (out of 17) and one RHD (out of 17) individuals, used more gestures with evaluative devices than the controls (see Table 6).

Table 6

Single case statistics profile of people with LHD and RHD for the proportion of gestures with evaluative devices per gesture.
(Control sample: n = 8, *M* = 0.14, *SD* = 0.13)

Lesion Site	**ID**	**% Gestures with evaluative devices per gesture**	**t**	**Significance test (p-value)**	**Estimated % of control population obtaining lower score than case**
Left	DD_755	0.50	2.594	0.035	0.02
Right	DF_316	0.54	2.883	0.023	0.01

4. Discussion

The aim of our study was to investigate the behavioral relationship and the neural underpinnings of spontaneous gesture production and macro-level discourse processing. We focused on two aspects of discourse; narrative complexity and evaluation competence. As a group, LHD produced more gestures and less complex narratives. Their gesture use was not related to evaluation competence. By contrast, the RHD group's frequency of gestures positively correlated with their evaluation competence, but their production of gestures was unrelated to narrative complexity.

LHD individuals produced more gestures than neurotypical individuals. Although not statistically significant, RHD individuals whose gesture production was comparable to neurotypical individuals, tended to use fewer gestures when compared to LHD individuals. Increased gesture production was linked to lesions in parts of the frontal-temporal-parietal network and basal ganglia that are implicated in language processing. In particular, the inferior frontal gyrus, superior temporal gyrus, putamen, caudate, and insula of the left hemisphere and the superior temporal gyrus and supramarginal gyrus of the right hemisphere were associated with increased rates of gesture production.

4.1. How does spontaneous gesture production during story-telling relate to the informativeness and complexity of a narrative?

Narrative complexity refers to an individuals' understanding of the content and ability to produce a coherent story (Berman & Slobin, 1994). This ability is closely tied to the micro-linguistic abilities of an individual (Karaduman et al., 2017). Following the evidence that people rely on gestures to compensate for their deficiencies at the micro-linguistic level (e.g. Göksun et al., 2015; Sekine & Rose, 2013), we expected gesture production in general and meaningful gestures specifically to be negatively associated with narrative complexity.

We found this expected link between narrative complexity and gesture by utterance only for people with LHD. At the group-level performance LHD individuals generally produced less complex narratives. RHD individuals as a group were comparable to controls. However, individual people with RHD did produce less complex narratives and their lesions maximally overlapped in the superior temporal gyrus, supramarginal gyrus and angular gyrus (Karaduman et al., 2017). Damage to these areas were associated with increased gesture production in the current study. The overlap of brain areas for increased gesture production and impaired narrative complexity in specific RHD individuals are consistent with the general pattern of an inverse relationship between narrative complexity and gesture production.

In addition to a relationship of general gesture use and narrative complexity, LHD individuals with less complex narratives made more *concrete deictic* gestures. Since concrete pointing can replace verbal language (Bangerter & Louwerse, 2005), we expected that *concrete-deixis*, but not *abstract-deixis* production would accompany lower performance in discourse. Conversely, since *abstract-deixis* serves a metanarrative function we expected its production be related to better discourse (McNeill et al., 1993). We did not find this expected link between abstract pointing and macro-level discourse. This absence of evidence might be related to the limited occurrence of abstract pointing in our sample. Most gestures, especially for LHD individuals, were concrete-deictic in our sample since the book was in front of the participants during story-telling to reduce the effects of memory skills on their performance.

This induced a frequent use of deictic gestures and most of the time, it eliminated the need for iconicity (we turn to this point under the implications for theoretical models of gesture section).

LHD individuals with increased use of gesture by utterance and more representational gestures (mainly deictics) had lesions that maximally overlapped in the inferior frontal gyrus (IFG), superior temporal gyrus, and basal ganglia. The left IFG is related to syntactic and phonological processing (Devlin, Matthews, & Rushworth, 2003; Paulesu, Frith, & Frackowiak, 1993) and recruiting lexical information (Hagoort, 2005). The left superior temporal gyrus (see Graves, Grabowski, Mehta, & Gupta, 2008) and the basal ganglia (for a review, see Ullman, 2004) are also implicated in lexical and phonological processing. Furthermore, the basal ganglia are strongly connected to the frontal cortex (see Ullman, 2004), to the extent that lesions in these structures may resemble deficiencies from frontal cortex damage (Leisman, Braun-Benjamin, & Melillo, 2014). Our results regarding the neural correlates of increased gesture production support the idea that people use more gestures to compensate for their language impairment.

This claim of increased use of gestures after the left IFG damage seems to contradict reports that people with frontal damage do not use gestures to compensate for speech deficiencies (e.g. Buxbaum et al., 2014; Göksun et al., 2015; Hogrefe et al., 2017). Yet, these contradictions might arise because of different levels of analyses in the studies. For instance, Göksun et al. (2015) investigated specific aspects of gesture production (*dynamic iconic gestures*) within a specific linguistic context (*spatial preposition production*). Spatial prepositions can be crucial for grammar because they establish a syntactical and semantic relationship between different linguistic entities. Grammatical impairments of language might manifest itself in both verbal and nonverbal modalities, making it impossible for one to compensate for the other (Damasio et al., 1986). While Göksun et al.'s (2015) findings in a

narrow aspect of gesture production are consistent with this view, our current findings support the idea that increased general gesture production can compensate for linguistic impairment.

In Buxbaum et al. (2014) and Hogrefe et al. (2017), people with left IFG lesion did not use gesture as a compensatory strategy either. But these studies investigated imitation. The left IFG activation is implicated in "*observation-execution matching*" for the actions (Nishitani, Schürmann, Amunts, & Hari, 2005), making it a plausible candidate for influencing motor imitation. Our study explored a range of spontaneous co-speech gestures which, by their nature, differ from action-imitation (McNeill, 1992).

4.2.How does gesture production during story-telling relate to evaluation competence in a narrative?

Evaluation competence refers to an individual's ability to invoke meaning of events around the global theme of the story (Küntay & Nakamura, 2004; Reilly et al., 2004). Given that gestures signal important parts of discourse (McNeill, 1992) and are related to pragmatic aspects of a narrative (Colletta et al., 2015), we expected evaluation competence to be positively associated with gesture production. We found this association in the RHD, but not in the LHD group.

The RHD group performance was comparable to the neurotypical controls in evaluative function (Karaduman et al., 2017). However, lesions of specific individuals with lower performance compared to neurotypical individuals clustered around frontotemporal areas that are assumed to be central in executive functioning (see Alvarez & Emory, 2006) and might explain our patients' deficiencies in globally connecting events (Karaduman et al., 2017). Extraction of a mental model from a visually depicted narrative requires the individual keep separate parts of the story in mind, which may increase the cognitive load (Marini et al., 2005).

Performance of RHD people in the macro-linguistic level of processing is usually associated with task demands and attentional deficiencies (Bartels-Tobin & Hinckley, 2005;

Marini, 2012; Sherratt & Bryan, 2012). Gesture might lighten the cognitive load of expression (Iverson & Goldin-Meadow, 1998). In that sense, high frequency of beat and deictic gestures that accompany evaluative utterances is notable given these gestures help segment discourse into meaningful parts (McNeill, 1992) and help individuals to focus their attention (Bangerter & Louwerse, 2005). Future research could investigate the way in which working memory and attention interact with gesture production and discourse competence.

4.3. Implications for Models of Gesture Production

Our results add to the evidence that LHD individuals produce more gestures during speech impairment (e.g., Akhavan et al., 2017; Beland & Ska, 1992; De Ruiter, 2006; Hadar et al., 1998; Herrman et al., 1988; Hogrefe et al., 2016; Lanyon & Rose, 2009; Sekine & Rose, 2013) and support the claim that gesture and speech systems can operate with relative independence (Butterworth & Hadar, 1997; De Ruiter, 2000; Feyereisen, 1983; Kita & Özyürek, 2003; Lausberg et al., 2007). These findings are consistent with *compensation* models (e.g. Hermann et al., 1988) and support the Information Packaging Hypothesis (Hostetter et al., 2007; Kita & Davies, 2009), which was further refined in the Interface Model (Kita & Özyürek, 2003). According to these models, people rely on nonverbal tools of communication when faced with verbal problems. Additionally, the Interface Model predicts increased representational gesture production as conceptual demands increase, because gestures help to structure thoughts into units that are conveniently packaged for verbal expression (Hostetter et al., 2007). In our sample, as the narrative complexity of the LHD individuals decreased (i.e., they had difficulty producing a coherent and cohesive narrative), they used more gestures.

At this point, we note that the majority of the representational gestures in this sample, especially for the LHD individuals, were deictic gestures. The relationship between narrative complexity and representational gesture production was also mainly based on deictics since the same relationship held when we analyzed deictic gestures alone. Even though deictic and iconic

gestures share common features in terms of denoting objects or events by virtue of their content-carrying characteristic, they differ with respect to the mode of representation. While deictic gestures denote a visually present referent, iconic gestures illustrate mental images in an imagistic and holistic manner (Hostetter, Alibali, & Kita, 2007). Thus, it might be possible that this relationship does not generalize to iconic representational gestures. Yet, evidence suggests that deictic gestures also reduce cognitive load and ease the explanation of information (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001), providing support for the idea that a similar relationship could be expected for iconic gestures (Hostetter, Alibali, & Kita, 2007). Future research should clarify the relationship between representational gestures and discourse competence in paradigms where people tell the narrative from their memory, which would decrease the frequency of deictic gestures.

The Lexical Facilitation Model (Hadar et al., 1998) also allows for preserved gesture production after linguistic impairment, because according to this model, gesture and speech depend on different systems of processing. The Lexical Facilitation Model suggests that representational gesture production facilitates lexical retrieval (Rauscher et al., 1996). If representational gestures indeed serve such a function, then one would expect a positive rather than a negative association between narrative complexity and gesture production. The negative association between gestures by utterances and narrative complexity for LHD individuals also contradicts the Growth-Point Theory (McNeill, 1992), which suggest that gesture and speech are different manifestations of a unified process. According to Growth-Point Theory, gesture and speech breakdown together in a way makes it impossible for gestures to compensate for deficiencies in speech (McNeill, 1992).

For the RHD group, we report a positive association between gesture production and evaluative competence. We suggest that the domain-general functions of gestures such as decreasing cognitive load and focusing attention could have served for their evaluative

functioning in narrative. The idea that gestures lighten cognitive load would be in line with the Image Activation Hypothesis (De Ruiter, 1998), which suggests that gestures help keep information that is encoded in speech active in mental imagery. If true, then people who use more gestures could evaluate a narrative better because of the information to be integrated is better maintained (see Alibali, 2005).

The findings with respect to a positive association between evaluation competence and gesture use are in line with the predictions of the Growth Point Theory according to which gesture relies on a holistic and global form of thinking while speech uses an analytic form. The global level of discourse is related to thinking along a *vertical axis*; keeping different parts of the discourse together and adopting a higher order thinking of content in relation to the context (Bamberg & Marchman, 1991). Following this logic, gesture with its global and holistic form of organization would be positively related to a global level of discourse. However, the positive link between evaluation competence and gesture production in RHD does not necessarily contradict the predictions of the Interface Model, because nonverbal tools of communication compensate for the deficiencies in the verbal modality (Kita & Özyürek, 2003) such that the predictions of the model are more relevant for the referential function of narrative.

To conclude, even though gesture and speech are highly coordinated in discourse, they can operate relatively independently of each other. Compensatory use of gesture in cases of linguistic impairment occurs. Additionally, people use more gestures when they have a problem communicating rich information, as evidenced by negative correlation between LHD patients' narrative complexity and gesture production. The evaluation of a story might be related to domain-general processes such as attention and working memory. As a result, individuals with RHD perform better when they produce more gestures that reduce their cognitive load. Our results highlight the idea that the interaction between gesture and speech depends on the level of analyses for both gesture and speech.

References

Ahlsén, E. (1991). Body communication as compensation for speech in a Wernicke's aphasic: a longitudinal study. *Journal of Communication Disorders*, *24*(1), 1-12.

Akhavan, N., Göksun, T., & Nozari, N. (in press). Integrity and functions of gestures in aphasia. *Neuropsychologia*.

Akhavan, N., Nozari, N., & Göksun, T. (2017). Expression of motion events in Farsi. *Language, Cognition and Neuroscience*, *32*(6), 792-804.

Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation*, *5*(4), 307-331.

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production. *Journal of Memory and Language*, *44*(2), 169–188.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: we think, therefore we gesture. *Language and Cognitive Processes*, *15*(6), 593-613.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: a meta-analytic review. *Neuropsychology Review, 16*(1), 17-42.

Andreetta, S. (2014). Features of narrative language in fluent aphasia (Doctoral thesis). Università degli Studi di Udine, Italy.

Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, *50*(8), 1787–1793.

Andric, M., & Small, S. L. (2012). Gesture's neural language. *Frontiers in Psychology*, *3*, 99.

Ash, S., Moore, P., Antani, S., McCawley, G., Work, M., Grossman, M. (2006). Trying to tell a tale: discourse impairments in progressive aphasia and frontotemporal dementia. *Neurology*, *66*(9), 1405–1413.

Azar, Z., & Özyürek, A. (2016). Discourse management. *Dutch Journal of Applied Linguistics*, *4*(2), 222-240.

Bamberg, M., & Damrad-Frye, R. (1991). On the ability to provide evaluative comments: Further explorations of children's narrative competencies. *Journal of Child Language*, *18*(03), 689-710.

Bamberg, M., & Marchman, V. (1991). Binding and unfolding: towards the linguistic construction of narrative discourse. *Discourse Processes*, *14*(3), 277-305.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, *15*(6), 415-419.

Bangerter, A., & Louwerse, M. M. (2005). Focusing attention with deictic gestures and linguistic expressions. In *Proceedings of the Cognitive Science Society*, 27.

Bartels-Tobin, L. R., & Hinckley, J. J. (2005). Cognition and discourse production in right hemisphere disorder. *Journal of Neurolinguistics*, *18*(6), 461-477.

Béland, R., & Ska, B. (1992). Interaction between verbal and gestural language in progressive aphasia: a longitudinal case study. *Brain and Language*, *43*(3), 355-385.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A cross linguistic developmental study*. Hillsdale, NJ: Erlbaum, p, 20-21, 41-42, 46.

Bernard, J. A., Millman, Z. B., & Mittal, V. A. (2015). Beat and metaphoric gestures are differentially associated with regional cerebellar and cortical volumes. *Human Brain Mapping*, *36*(10), 4016-4030.

Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain*, *114*(3), 1115-1127.

Blonder, L. X., Burns, A. F., Bowers, D., Moore, R. W., & Heilman, K. M. (1995). Spontaneous gestures following right hemisphere infarct. *Neuropsychologia*, *33*(2), 203-213.

Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to McNeill. *Psychological Review*, *96*(1), 168-174.

Buxbaum, L. J., Shapiro, A. D., & Coslett, H. B. (2014). Critical brain regions for tool-related and imitative actions: a componential analysis. *Brain, 137*(7), 1971-1985.

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language, 23*(3), 645-673.

Cassell, J., & McNeill, D. (1991). Gesture and the poetics of prose. *Poetics Today, 12*(3), 375-404.

Cicone, M., Wapner, W., Foldi, N., Zurif, E., & Gardner, H. (1979). The relation between gesture and language in aphasic communication. *Brain and Language, 8*(3), 324-349.

Cocks, N., Dipper, L., Pritchard, M., & Morgan, G. (2013). The impact of impaired semantic knowledge on spontaneous iconic gesture production. *Aphasiology, 27*(9), 1050-1069.

Cocks, N., Hird, K., & Kirsner, K. (2007). The relationship between right hemisphere damage and gesture in spontaneous discourse. *Aphasiology, 21*(3-4), 299-319.

Colletta, J. M., Guidetti, M., Capirci, O., Cristilli, C., Demir, O. E., Kunene-Nicolas, R. N., & Levine, S. (2015). Effects of age and language on co-speech gesture production: an investigation of French, American, and Italian children's narratives. *Journal of Child Language, 42*(1), 122-145.

Crawford, J. R., & Garthwaite, P. H. (2007). Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology, 21*(5), 611.

Damasio, A., Bellugi, U., Damasio, H., Poizner, H., & Van Gilder, J. (1986). Sign language aphasia during left-hemisphere Amytal injection. *Nature, 322*(6077), 363-365.

De Ruiter, J. P. (1998). Gesture and speech production. Unpublished dissertation, Katholieke Universiteit Nijmegen.

De Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture,* 284-311. Cambridge: Cambridge University Press.

De Ruiter, J. P. (2006). Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech Language Pathology*, *8*(2), 124-127.

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. S. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of Cognitive Neuroscience*, *15*(1), 71-84.

Dipper, L., Cocks, N., Rowe, M., & Morgan, G. (2011). What can co-speech gestures in aphasia tell us about the relationship between language and gesture?: a single case study of a participant with conduction aphasia. *Gesture*, *11*(2), 123–147.

Druks, J. (2000). Object and action naming battery. Psychology Press.

Eling, P., & Whitaker, H. (2009). History of aphasia: from brain to language. *Handbook of Clinical Neurology*, *95*, 571-582.

Enfield, N. J., Kita, S., & De Ruiter, J. P. (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, *39*(10), 1722-1741.

Feyereisen, P. (1983). Manual activity during speaking in aphasic subjects. *International Journal of Psychology*, *18*(1-4), 545–556.

Gentilucci, M., & Volta, R. D. (2008). Spoken language and arm gestures are controlled by the same motor control system. *The Quarterly Journal of Experimental Psychology*, *61*(6), 944-957.

Glosser, G., Wiener, M., & Kaplan, E. (1986). Communicative gestures in aphasia. *Brain and Language*, *27*(2), 345–359.

Goldenberg, G., & Randerath, J. (2015). Shared neural substrates of aphasia and apraxia. *Neuropsychologia*, *75*, 40–49.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516-522.

Goodwin, C. (1995). Co-constructing meaning in conversations with an aphasie man. *Research on Language and Social Interaction*, *28*(3), 233-260.

Göksun, T., Lehet, M., Malykhina, K., & Chatterjee, A. (2013). Naming and gesturing spatial relations: evidence from focal brain-injured individuals. *Neuropsychologia, 51*(8), 1518- 1527.

Göksun, T., Lehet, M., Malykhina, K., & Chatterjee, A. (2015). Spontaneous gesture and spatial language: evidence from focal brain injury. *Brain and Language, 150*, 1-13.

Graves, W. W., Grabowski, T. J., Mehta, S., & Gupta, P. (2008). The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. *Journal of Cognitive Neuroscience*, *20*(9), 1698-1710.

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., & Kircher, T. (2009). Neural integration of iconic and unrelated coverbal gestures: a functional MRI study. *Human Brain Mapping, 30*(10), 3309-3324.

Häberling, I. S., Corballis, P. M., & Corballis, M. C. (2016). Language, gesture, and handedness: Evidence for independent lateralized networks. *Cortex, 82*, 72-85.

Hadar, U., Burstein, A., Krauss, R., & Soroker, N. (1998). Ideational gestures and speech in brain-damaged subjects. *Language and Cognitive Processes*, *13*(1), 59-76.

Hadar, U., & Butterworth, B. (1997). Iconic gestures, imagery, and word retrieval in speech. *Semiotica*, *115*(1-2), 147-172.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences, 9*(9), 416-423.

Helmich, I., & Lausberg, H. (2014). Hand movements with a phase structure and gestures that depict action stem from a left hemispheric system of conceptualization. *Experimental Brain Research, 232*, 3159–3173.

Herrmann, M., Reichle, T., Lucius-Hoene, G., Wallesch, C. W., & Johannsen-Horbach, H. (1988). Nonverbal communication as a compensative strategy for severely nonfluent aphasics? A quantitative approach. *Brain and Language*, *33*(1), 41–54.

Hogrefe, K., Rein, R., Skomroch, H., & Lausberg, H. (2016). Co-speech hand movements during narrations: what is the impact of right vs. left hemisphere brain damage? *Neuropsychologia*, *93*, 176-188.

Hogrefe, K., Ziegler, W., Weidinger, N., & Goldenberg, G. (2012). Non-verbal communication in severe aphasia: influence of aphasia, apraxia, or semantic processing? *Cortex*, *48*(8), 952–962.

Hogrefe, K., Ziegler, W., Weidinger, N., & Goldenberg, G. (2017). Comprehensibility and neural substrate of communicative gestures in severe aphasia. *Brain and Language*, *171*, 62-71.

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

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

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396*(6708), 228.

Joanette, Y., & Goulet, P. (1990). Narrative discourse in right brain-damaged right-handers. In Y. Joanette & H. H. Brownell (Eds.), *Discourse ability and brain damage*. New York: Springer-Verlag.

Joanette, Y., Goulet, P., Ska, B., & Nespoulous, J. L. (1986). Informative content of narrative discourse in right-brain-damaged right-handers. *Brain and Language*, *29*(1), 81-105.

Kaczmarek, B. L. J. (1984). Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes. *Brain and Language*, *21*(1), 52-58.

Karaduman, A., Göksun, T., & Chatterjee, A. (2017). Narratives of focal brain injured individuals: a macro-level analysis. *Neuropsychologia*, *99*, 314-325.

Kendon, A. (1988). How gestures can become like words. In F. Poyatos (Ed.), *Cross cultural perspectives in nonverbal communication*. Toronto: Hogrefe Publishers.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, *27*(3), 175-200.

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

Kertesz, A. (1982). Western aphasia battery test manual. Psychological Corp.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, *49*(4), 294-303.

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., ... & Green, A. (2009). Neural interaction of speech and gesture: differential activations of metaphoric co-verbal gestures. N*europsychologia*, *47*(1), 169-179.

Kita, S. (2014) Production of speech-accompanying gesture. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *Oxford Handbook of Language Production* (451-459). Oxford: Oxford University Press

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, *124*(3), 245-266.

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

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

Kong, A. P. H., Law, S. P., Kwan, C. C. Y., Lai, C., & Lam, V. (2015). A coding system with independent annotations of gesture forms and functions during verbal communication: development of a database of speech and gesture (DoSaGE). *Journal of Nonverbal Behavior, 39*(1), 93-111.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language, 57*(3), 396-414.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2001). Lexical gestures and lexical access: a processing model. In D. McNeill, (Ed.), *Language and gesture* (261-283). Cambridge: Cambridge University Press.

Krauss, R. M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. In L. S. Messing & R. Campbell (Eds.), *Gesture, speech, and sign* (93–116). New York: Oxford University Press.

Küntay, A., & Nakamura, K. (2004). Linguistic strategies serving evaluative functions: a comparison between Japanese and Turkish narratives. In S. Strömqvist, & L. Verhoeven (Eds.*), Relating events in narrative, Volume 2: Typological and contextual perspectives* (329-358). Mahwah: Lawrence Erlbaum Associates.

Labov, W., & Waletzky, J. (1997). Narrative analysis: oral versions of personal experience. *Journal of Narrative & Life History*, *7*(1-4), 3-38.

Lanyon, L., & Rose, M. (2009). Do the hands have it? The facilitation effects of arm and hand gesture on word retrieval in aphasia. *Aphasiology*, *23*(7-8), 809–822.

Lausberg, H., Zaidel, E., Cruz, R. F., & Ptito, A. (2007). Speech-independent production of communicative gestures: evidence from patients with complete callosal disconnection. *Neuropsychologia*, *45*(13), 3092-3104.

Leisman, G., Braun-Benjamin, O., & Melillo, R. (2014). Cognitive-motor interactions of the basal ganglia in development. *Frontiers in Systems Neuroscience*, *8*, 16.

Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, *26*(10), 1457-1471.

Mar, R. A. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. *Neuropsychologia*, *42*(10), 1414-1434.

Marini, A. (2012). Characteristics of narrative discourse processing after damage to the right hemisphere. *Seminars in Speech and Language*, *33*(01), 68-78.

Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language*, *93*(1), 46-54.

May, A. L., David, R., & Thomas, A. P. (1988). The use of spontaneous gesture by aphasic patients. *Aphasiology*, *2*(2), 137-145.

Mayer, M., 1969. Frog, where are you? Dial Press, New York.

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review*, *92*(3), 350-371.

McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D., Cassell, J., & Levy, E. T. (1993). Abstract deixis. *Semiotica*, *95*(1-2), 5-20.

McNeill, D., Duncan, S. D., Cole, J., Gallagher, S., & Bertenthal, B. (2008). Growth points from the very beginning. *Interaction Studies*, *9*(1), 117-132.

McNeill, D., & Levy, E. (1982). Conceptual representations in language activity and gesture, In R. J. Jarvella & W. Klein (Eds.), *Speech, Place, and Action* (271-95). Chichester: John Wiley and Son.

McNeill, D., & Levy, E. T. (1993). Cohesion and gesture. *Discourse processes, 16*(4), 363-386.

McNeill, D., & Pedelty, L. (1995). Right brain and gesture. In K. Emmorey & J. Reilly (Eds), *Language, gesture and space* (63–86). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Mol, L., & Kita, S. (2012). Gesture structure affects syntactic structure in speech. In *Proceedings of the Cognitive Science Society, 34*.

Nicoladis, E., Marentette, P., & Navarro, S. (2016). Gesture frequency linked primarily to story length in 4–10-year old children's stories. *Journal of Psycholinguistic Research*, *45*(2), 189-204.

Nicoladis, E., Pika, S., Yin, H., & Marentette, P. (2007). Gesture use in story recall by Chinese–English bilinguals. *Applied Psycholinguistics*, *28*(4), 721-735.

Nishitani, N., Schürmann, M., Amunts, K., & Hari, R. (2005). Broca's region: from action to language. *Physiology, 20*(1), 60-69.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*, 342-345.

Pritchard, M., Cocks, N. & Dipper, L. (2013). Iconic gesture in normal language and word searching conditions: a case of conduction aphasia. *International Journal of Speech-Language Pathology*, *15*(5), 524-534.

Pritchard, M., Dipper, L., Morgan, G., & Cocks, N. (2015). Language and iconic gesture use in procedural discourse by speakers with aphasia. *Aphasiology*, *29*(7), 826-844.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: the role of lexical movements in speech production. *Psychological Science*, *7*(4), 226-231.

Reilly, J.S., Bates, E.A., Marchman, V.A., 1998. Narrative discourse in children with early focal brain injury. *Brain and Language*, *61*(3), 335–375.

Reilly, J., Losh, M., Bellugi, U., & Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language*, *88*(2), 229–247.

Ross, E. D., & Mesulam, M. M. (1979). Dominant language functions of the right hemisphere. *Archives of Neurology*, *36*(3), 144–148.

Rousseaux, M., Daveluy, W., & Kozlowski, O. (2010). Communication in conversation in stroke patients. *Journal of Neurology*, *257*(7), 1099-1107.

Scharp, V. L., Tompkins, C. A., & Iverson, J. M. (2007). Gesture and aphasia: helping hands? *Aphasiology*, *21*(6-8), 717–725.

Sekine, K., & Rose, M. L. (2013). The relationship of aphasia type and gesture production in people with aphasia. *American Journal of Speech-Language Pathology*, *22*(4), 662-672.

Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and Language*, *25*(1), 19-36.

Sherratt, S., & Bryan, K. (2012). Discourse production after right brain damage: gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics*, *25*(4), 213-239.

Straube, B., Green, A., Bromberger, B., & Kircher, T. (2011). The differentiation of iconic and metaphoric gestures: Common and unique integration processes. *Human Brain Mapping*, *32*(4), 520-533.

Ulatowska, H. K., North, A. J., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain and Language*, *13*(2), 345-371.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231-270.

Wartenburger, I., Kühn, E., Sassenberg, U., Foth, M., Franz, E. A., & Van der Meer, E. (2010). On the relationship between fluid intelligence, gesture production, and brain structure. *Intelligence, 38*(1), 193-201.

Wilkinson, R. (2013). Gestural depiction in acquired language disorders: on the form and use of iconic gestures in aphasic talk-in-interaction. *Augmentative and Alternative Communication, 29*(1), 68-82.

Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language, 101*(3), 278-289.

Willems, R.M., Ozyurek, A., & Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cerebral Cortex, 17*(10), 2322–2333.

Wu, Y. C., & Coulson, S. (2007). How iconic gestures enhance communication: an ERP study. *Brain and Language, 101*(3), 234-245.

Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. In *Proceedings of the National Academy of Sciences, 106*(49), 20664-20669.

Acknowledgements

This research was supported in part by NIH RO1DC012511 and grants to the Spatial Intelligence and Learning Center, funded by the National Science Foundation (subcontracts under SBE- 0541957 and SBE-1041707). We would like to thank Marianna Stark and Eileen Cardillo for their help in recruiting people with brain injury. We also thank Language and Cognition Lab members at Koç University for discussions about the project, and Hazal Kartalkanat for helping with reliability coding.

APPENDIX A

The definitions and examples of components of the coding template for narrative complexity (adapted from Berman & Slobin, 1994 and Köksal, 2011) (see Karaduman et al., 2017, Appendix A)

Plot components	Plot Sub-Components	Examples and Explanations
Plot onset	Precedent event	-The boy and the dog wakes up
	Temporal location	-In the morning/evening
	Characters	The boy/child/kid, the dog, the frog Scoring ranges between 0-3: Only one character = 1 Two of the characters = 2 Three characters = 3
	The main characters learn something	-The boy and the dog noticed that the frog had gone missing -The boy and the dog wake up and see that the frog is gone
	Inference about the frog's disappearance	-The frog escaped from the jar -The frog is gone -The jar is empty
	The response of protagonist	-They are fascinated by the frog's disappearance -The boy was shocked
Plot unfolding	Searching for the lost frog in the home	-The boy is looking in his boots -The dog is looking in the jar
	Encountering the bees	-The dog is looking into the beehive -The bees are chasing the dog
	Encountering the gopher	-The boy is looking down a hole and a gopher comes out -The boy gets bitten by a gopher
	Encountering the owl	-The boy falls down because an owl comes out of the tree -The boy disturbed an owl in the tree
	Encountering the deer	-The child climbs on the deer -The deer tosses the boy over the cliff
	Falling down	-The boy and the dog falls in the water/pond/lake
Resolution	Protagonist finds the lost frog	-The kid finds his frog
Search theme	Explicit mention of the lost frog	-Whether the subject explicitly mentions that the frog is missing and the boy is searching for him (range: 0-2). 1 point for mentioning each aspect of initiating the search theme: -The frog is missing/ gone. -The boy is looking for the frog. * Only mentioning that the frog leaves its jar will not get point
	Reiteration of search theme	-Whether the search theme was reiterated later. (range: 0-2). No additional mention = 0 1 or 2 additional mentions = 1 Multiple additional mentions = 2.

APPENDIX B

The coding scheme for evaluative function and examples of subcategories. (The titles were taken from Karaduman et al., 2017, Appendix B.)

1. *Cognitive inferences*: Frames of mind in the formulations of Bamberg and Damrad-Frye (1991) and Küntay and Nakamura (2004). The inferences narrator makes regarding the mental states and motivations of the agents in the story. Examples: “Frog *wanted* to live with his frog family*”* or “the dog *decides* that maybe he should get a closer look to the beehive”.
2. *Social engagement devices*: The phrases that are used to catch the attention of the listener. Onomatopoeic words, character speech, exclamations are all included in this category. This category does not have a direct counterpart in Küntay & Nakamura’s (2004) formulation. Exclamations as in “*Ohp*! A deer got him!” or the sound symbolism as in here “The dog and the boy are looking at his pet frog, "*ribbet ribbet*"?” can be given as examples.
3. *References to affective states or behaviors*: When the narrator makes reference to the emotional states of the agents in the story. Examples: “he's *annoyed* now” or “was unhappy that the dog got into such mischief”.
4. *Enrichment expressions:* Adverbial phrases which indicate that the event was not anticipated or inferred and repetitions which are used to emphasize an event. The adverbial phrase in “The bees *suddenly* started to swarm” or the connective in “*but* they're safe” can be taken as examples.
5. *Hedges:* The phrases which indicate that the narrator is not certain about the truth value of what she says. The uncertainty expression in “he's starting to put on his clothes *I think*” or this epistemic-distancing device in “he's *probably* looking for the frog” can be taken as examples.

6. *Evaluative remarks:* When the narrator makes a subjective judgment or comment about an event in the narrative. Examples: "*Oh, poor thing*, he done fell in the water!" or "down into a ravine that's filled with water *sure that's gotta hurt*".

APPENDIX C

The definitions and examples for gesture scheme. The definitions for the forms are adopted from McNeill (1992) and Kong et al. (2015).

1. *Iconic*: The movements of hand which depict the meaning carried through the speech by similarity. They can be *static* (e.g. using the hand with the palm looking upward while saying jar) or *dynamic* (e.g. both hands moving towards each other to depict cuddling up).
2. *Deictic*: In other words, *pointing*; they do not represent the meaning but they refer things in space by virtue of their proximity to them. We classified the gestures in this category into two as pointing to the book and others. Pointing to the locations of objects and whole hand pointing where the participant uses her hand to locate something in space were included in the abstract pointing category. Pointing to the book and pointing to concrete objects were classified as concrete pointing.
3. *Emblem*: These are gestures that have conventional meaning and they can convey meaning without the need of accompanying speech (e.g. waving the hands to say good bye).
4. *Beat*: Gestures that do not represent meaning related to the content of speech; they are rapid and rhythmic movements of the hand which usually occur synchronous to the prosody of the speech.

Gestures with evaluative devices: Gestures that were defined as belonging to one of the four categories were further classified as to whether they were accompanied with evaluative devices. Representational gestures: Iconic and deictic gestures were classified as representational gestures. Non-representational gestures: Beats and emblems were classified as non-representational gestures.