Research Article

Task-Specific Iconic Gesturing During Spoken Discourse in Aphasia

Brielle C. Stark[a] iD **and Caroline Cofoid[a]**

Purpose: In persons living with aphasia, we will explore the relationship between iconic gesture production during spontaneous speech and discourse task, spoken language, and demographic information.

Method: Employing the AphasiaBank database, we coded iconic gestures in 75 speakers with aphasia during two spoken discourse tasks: a procedural narrative, which involved participants telling the experimenter how to make a sandwich ("Sandwich"), and a picture sequence narrative, which had participants describe the picture sequence to the experimenter ("Window"). Forty-three produced a gesture during both tasks, and we further evaluate data from this subgroup as a more direct comparison between tasks.

Results: More iconic gestures, at a higher rate, were produced during the procedural narrative. For both tasks, there was a relationship between iconic gesture rate, modeled as iconic gestures per word, and metrics of language dysfluency extracted from the discourse task as

well as a metric of fluency extracted from a standardized battery. Iconic gesture production was correlated with aphasia duration, which was driven by performance during only a single task (Window), but not with other demographic metrics, such as aphasia severity or age. We also provide preliminary evidence for task differences shown through the lens of two types of iconic gestures.

Conclusions: While speech-language pathologists have utilized gesture in therapy for poststroke aphasia, due to its possible facilitatory role in spoken language, there has been considerably less work in understanding how gesture differs across naturalistic tasks and how we can best utilize this information to better assess gesture in aphasia and improve multimodal treatment for aphasia. Furthermore, our results contribute to gesture theory, particularly, about the role of gesture across naturalistic tasks and its relationship with spoken language.

Supplemental Material: https://doi.org/10.23641/asha.14614941

Gesture is a powerful tool that accompanies and, sometimes, replaces speech. Gestures lie along a continuum (Kendon's continuum; McNeill, 1992), ranging from gestures with no linguistic association to gestures associated with speech ("language-like gestures"), to gestures conveying meaning in the absence of speech ("pantomimes"), and to gestures holding independent status as symbolic forms ("emblems"; e.g., the "OK" sign).

Gesturing in Aphasia

In the case of persons living with acquired aphasia (a language disorder), gesturing may serve a particularly

[a]Department of Speech, Language and Hearing Sciences, Indiana University Bloomington

Correspondence to Brielle C. Stark: bcstark@iu.edu

communicative purpose, as a means of compensating for spoken language difficulties or, indeed, as a mechanism to overcome word-finding difficulties—reminiscent of what Luria (1970) termed *intersystemic reorganization* (Dipper et al., 2015; Hadar & Butterworth, 1997; Krauss, 1998; Rose & Douglas, 2001). There is emerging evidence that people with aphasia use gestures to accompany/be redundant with (Kong et al., 2017) or supplement (e.g., disambiguate, add; Dipper et al., 2015) speech. The relationship between language and gesture has also been explored in neurotypical populations, more often finding that gesture is redundant with speech (Dargue et al., 2019; Holle & Gunter, 2007; Hostetter, 2011; Kelly et al., 1999).

In this study, we focus on iconic gestures. Iconic gestures are language-like gestures that are related to the content of speech and have a form (e.g., motion, hand shape, location) that is related to this content; they typically are not meaningful in the absence of speech (Hadar & Butterworth, 1997). Iconic gestures are highly common in aphasia. Indeed, a study in 95 persons with aphasia demonstrated that all subjects employed at least one type of iconic gesture during a fictional story retell (Cinderella story), further emphasizing

reliance on iconic gesture to convey, disambiguate, or add meaning during discourse by persons with otherwise impoverished spoken language (Sekine & Rose, 2013).

Theories of Gesture's Relationship With Language in Aphasia

There are several theories that have been used to describe gesture use in aphasia. The Sketch model suggests that, when gesturing gets harder, speakers will rely relatively more on speech and that, alternatively, when speaking gets harder, speakers will rely relatively more on gestures (de Ruiter, 2006; de Ruiter et al., 2012). This theory shares similarities with the growth point theory, which postulates that gesture and language share prelinguistic conceptual stage resources before diverging—with language taking a linguistic path and gesture taking a motoric one (McNeill, 1992). Broadly, these theories fit with evidence comparing iconic gesture usage in aphasia to gesture usage in matched samples of persons without aphasia, finding almost always that persons with aphasia gesture more often while also producing less speech (Sekine & Rose, 2013). In an alternative hypothesis, the lexical retrieval hypothesis, gestures are thought to directly facilitate lexical retrieval processes (Krauss, 1998; Krauss & Hadar, 2001). Iconic gestures, in particular, are thought to originate in the processes that precede conceptualization/formulation of the preverbal message and, as such, can precede even in cases of blockages or damage to later stages. Importantly, this theory hypothesizes that the imagistic information from iconic gesture may facilitate lexical retrieval by "defining the conceptual input to the semantic lexicon; by maintaining a set of core features while reselecting a lexical entry; and by means of directly activating phonological word-forms" (Hadar & Butterworth, 1997). Indeed, iconic gesturing has been shown to improve object naming (Rose & Douglas, 2001) and occur alongside word-finding problems (Kong et al., 2015; Pritchard et al., 2013). Notably, gesture is also thought to reflect underlying thoughts and knowledge not verbalized in speech (Church & Goldin-Meadow, 1986; Goldin-Meadow & Alibali, 2013), which allow persons with aphasia to communicate and demonstrate competence despite a language production barrier.

The Relationship of Iconic Gesture With Spoken Language Fluency

Not surprisingly, given the theories above, language dysfluency is thought to correlate with greater gesture use in aphasia. Most research evaluating the relationship between language dysfluency and iconic gesture in aphasia use has focused on the differing frequency of iconic gesture use stratified by aphasia type, that is, a comparison of iconic gesturing in persons with nonfluent as compared to fluent aphasia (a dichotomous type of fluency classification) or a comparison of iconic gesturing in persons with aphasia types based on standardized testing batteries (i.e., a classificationist type of fluency classification, e.g., Broca's, conduction, Wernicke's). In general, this research has been mixed, with some studies finding significantly more iconic

gesture use in nonfluent populations, primarily Broca's aphasia (Goldblum, 1978; Hadar, 1991; but see Cicone et al., 1979), while other studies have observed a high incidence of iconic gestures across all aphasia types (Feyereisen, 1983; Sekine & Rose, 2013).

Cicone et al. (1979) found that people with nonfluent aphasia produced fewer gestures (yet these were clear and informative), which contrasted with the frequent gestures produced by people with fluent aphasia, which were vague. In a relatively large sample of people with aphasia completing a story retell task, people with Broca's aphasia (a type of nonfluent aphasia) produced overall more gestures and were more likely to produce iconic gestures (Sekine & Rose, 2013). In a single-case study of a person with conduction aphasia (a type of fluent aphasia), it was found that, when recounting a cartoon, the individual with conduction aphasia produced more iconic gestures than a comparison sample of controls during word searching behavior (Pritchard et al., 2013). Interestingly, this individual produced a similar frequency of iconic gestures compared with control participants when the person with aphasia was producing fluent, errorless language. In a larger sample of persons with aphasia and matched controls, iconic gestures were produced in similar frequencies and forms by both groups, but the aphasia group utilized iconic gestures alongside their otherwise semantically impoverished language (Pritchard et al., 2015). While this evidence seems relatively straightforward (emphasizing the use of more gestures, particularly the iconic type, by persons with nonfluent and/or Broca's aphasia), some studies have provided alternative findings. For example, one study found that persons with Wernicke's aphasia (a fluent type) produce more iconic gestures per unit of time (Carlomagno & Cristilli, 2006).

Overall, it is difficult to draw conclusions based on iconic gesture frequency and rate by evaluations leveraging aphasia type because of the vast intragroup differences in language ability *within* aphasia types (Dalton & Richardson, 2019; Stark & Fukuyama, 2020). For example, iconic gestures likely rely on some intact prelinguistic components, such as prelinguistic conceptual knowledge (Hadar & Butterworth, 1997). Some persons with Wernicke's aphasia can present with impoverished prelinguistic conceptual ability, while other persons with Wernicke's aphasia can present with relatively intact prelinguistic conceptual ability (Kertesz, 2007). This makes drawing overarching conclusions about gesture usage stratified by aphasia type inherently difficult and, perhaps, not meaningful.

Instead, evaluations of gesture use, particularly iconic gesture use, may be more sensitively investigated by evaluating the relationship of gesture with metrics of language derived from the task itself (e.g., dysfluencies) and with some extra-task metric of fluency that provides a category or metric of fluency, as opposed to an aphasia type. Restricting gesture use in neurotypical adults has demonstrated reduced speech fluency, measured by an association of gesture use (or nonuse) with time spent pausing, number of words, number of pauses, mean pause length, and number of hesitations (Graham & Heywood, 1975), with relative frequency

of nonjuncture filled pauses in speech with spatial content (Rauscher et al., 1996) and with a slower speech rate (Morsella & Krauss, 2004). A study by Kong et al. (2015) evaluated 48 Cantonese-speaking persons with aphasia, comparing gesture use with several linguistic measures produced during narrative tasks, including type–token ratio, percentage of simple utterances, percentage of complete utterances, and percentage of dysfluency (defined as repetitions of words or syllables, sound prolongations, pauses, self-corrections, and interjections, as a proportion of total utterances). Those who tended to produce higher numbers of gestures were also those who demonstrated a higher proportion of dysfluencies and a lower number of complete and simple utterances. There was not a significant difference between high/low gesture groups in regard to lexical diversity (type–token ratio; Kong et al., 2015). A more recent study by Kong et al. (2017) further emphasized the relationship between dysfluency and gesture use in Cantonese speakers with aphasia, finding that percentages of complete sentences and dysfluency strongly predicted the gesturing rate in aphasia. Both Kong et al. (2017, 2015) and other studies evaluating language dysfluency's relationship with gesture in aphasia (e.g., Sekine & Rose, 2013) have collapsed across gesture types, so further work evaluating iconic gesture's relationship with language dysfluency in aphasia is warranted.

The Role of Task in Gesture Use

While the literature has made clear that iconic gestures are heavily used in aphasia, there exist critical gaps. The primary gap targeted by our project is on task specificity of iconic gesture use in aphasia. Most studies evaluating iconic gesture in aphasia have focused on gesturing during a single narrative task (e.g., fictional story retell; Kistner et al., 2019; Pritchard et al., 2015; Sekine & Rose, 2013). However, restricting evaluation of iconic gestures to a single narrative task likely lends us an impoverished understanding of how, when, and why iconic gestures are employed in naturalistic contexts. For example, speakers (without aphasia) gesture more when a task is cognitively or linguistically complex (Kita & Davies, 2009). Indeed, evidence from participants with amnesia and healthy comparison participants suggests that both subject groups gesture at lower rates during procedural discourse tasks compared to autobiographical/ episodic narratives, likely because autobiographical/episodic narratives reflect rich episodic details (Hilverman et al., 2016). Persons with aphasia have also been shown to use more iconic gestures during a cartoon narration task compared with a spontaneous conversation (de Beer et al., 2019), supporting the Gesture as Simulated Action framework, which predicts that speakers gesture more when speech is based on imagery (Hostetter & Alibali, 2019).

Evaluating task effects is clinically and theoretically meaningful. It is well acknowledged that spoken language is task specific, that is, that the microlinguistic and functional (macrostructural) aspects of spoken language shift according to task demands (Dalton & Richardson, 2019; Li et al., 1996; Shadden et al., 1990; Stark, 2019; Ulatowska et al., 1981).

For that reason, it is best practice to employ a variety of tasks to most comprehensively evaluate spoken language ability (Brookshire & Nicholas, 1994; Stark, 2019). It follows that employing a variety of tasks to assess a person's co-speech gesturing is important.

Furthermore, evaluating the extent to which task influences gesture can lend valuable information toward planning treatment, for example, provide information to the clinician regarding typical gesture use and atypical or inaccurate gesture use at a task-specific level. Presently, gesture-based or gesture-integrative therapies are few, and these therapies have largely not resulted in significant improvement in spoken language in aphasia (Rose et al., 2013). One such reason may be our lack of understanding about task-specific gesturing. Therefore, directly comparing gesture use between tasks in the same person will allow us to understand gesture use more sensitively and comprehensively, as well as the relationship between task and gesture. Doing so in naturalistic tasks can also lend critical insight about how gesture is used in a spontaneously communicative sense, rather than in a more heavily constrained or structured task. Finally, if we can understand the extent to which gestures are produced across a range of naturalistic tasks in aphasia and how these gestures relate to spoken language competency, we can more accurately formulate predictive hypotheses regarding language recovery. For example, we can answer clinically critical questions such as "To what extent does iconic gesturing during narrative in the acute phase of aphasia predict communicative success in the chronic stage of aphasia (or indeed, predict in which individual's aphasia will resolve)?"

Evaluating task-specific gesturing also has critical importance for theories related to gesture use in typical populations, as well as those with language impairments. Some discourse tasks employ pictures, which in turn may facilitate different patterns of gesturing—for example, a trend toward more concrete/deictic gestures, rather than iconic gestures. In tandem, picture-oriented tasks may elicit more nouns and simpler language structure than other tasks (Stark, 2019). For speakers with more severe aphasia, certain tasks can elicit more spoken output than less structured tasks (Stark & Fukuyama, 2020). It is therefore of interest to compare gesture usage across tasks with varying cognitive and linguistic complexity and to evaluate the interaction between gesture production, task, and the linguistic and cognitive characteristics of speakers. For example, iconic gesturing may be used more often and with a greater success rate when tasks do not involve other visual stimuli (e.g., picture descriptions), as iconic gestures may "stand in" for the absent visual, concrete imagery. However, this may only hold true for individuals who have impoverished postconceptual linguistic processes, but not necessarily those who have impoverished preconceptual knowledge.

Research Questions

In this study, we address some of these research gaps, comparing iconic gestures made by persons with aphasia

during two monologic discourse tasks: a procedural narrative and an expository, picture sequence description. We investigate two research questions:

1. To what extent do frequency and rate of iconic gesture differ between two spoken discourse tasks, that is, is there a main effect of task on gesture?

2. To what extent do frequency and rate relate to spoken language, and is this relationship specific to task?

Methodology and Design

Participants

Participant data were collected from AphasiaBank, a password-protected database for researchers interested in spoken discourse use in aphasia (MacWhinney et al., 2011). Inclusion criteria included members of the Aphasia subset within AphasiaBank, all of whom had acquired brain damage and aphasia (or latent aphasia) as per clinical assessment and standardized testing scores. Included participants spoke English as their primary language and had audiovisual data for both tasks ($N = 303$ persons had data for the picture sequence task and $N = 234$ persons for the procedural task; $N = 233$ people had data for both tasks). In AphasiaBank, some participants have data for multiple time points (tagged as "a" [first time point], "b" [second time point], and so on). For this study, we included data from only the first time point for each participant.

We then excluded participants for whom their video (for either task) did not show the entirety of both upper limbs or where the angle of the video did not allow for complete gesture viewing. This was necessary to ensure accurate gesture coding. We then excluded persons with aphasia who did not produce an iconic gesture during at least one task of interest. With these exclusion criteria, approximately 98 were excluded. Finally, participants were excluded if they were given a picture aid for the "Sandwich" procedural discourse (described in more detail in the Stimuli section). This decision was made because not every individual was given this additional visual support (roughly 20% of the database received a picture during the Sandwich task, $N = \sim 60$). As we wanted to evaluate the difference in gesture usage between discourse genres and only the picture sequence task was meant to include a visual aid, inclusion of those individuals with a visual aid during the procedural Sandwich task was inconsistent. Following the parameters described above, this study included a total of 75 persons with aphasia, who are further described in Table 1.

Stimuli

To evaluate our first research question (evaluate a main effect of task on iconic gesture frequency and rate), gestures were analyzed during two spontaneous speech discourse tasks, drawn from the AphasiaBank protocol (MacWhinney et al., 2011), called the Sandwich and Window tasks. These two tasks were chosen (a) based on prior research out of our lab implicating cognitive and linguistically different

requirements per task (Stark, 2019; Stark & Fukuyama, 2020) and (b) because these tasks are less frequently explored in aphasia gesture research.

The Sandwich task was a procedural narrative, in which participants described how to make a peanut butter and jelly sandwich. As noted earlier, this task did not include any visual aids. The instructions for the Sandwich task were as follows: "Let's move on to something a little different. Tell me how you would make a peanut butter and jelly sandwich." If no response in 10 s was given, the examiner gave a second prompt: "If you were feeling hungry for a peanut butter and jelly sandwich, how would you make it?" If no response was given, the examiner utilized a set of troubleshooting questions (available on aphasia.talkbank.org).

The "Broken Window" task (shortened here to "Window") was a descriptive task (specifically, a picture sequence description) in which participants described a sequence of four pictures: a boy kicking a soccer ball through a picture window, knocking over a lamp and surprising a sitting man (Menn et al., 1998). The instructions for the Window task were as follows: "Now, I'm going to show you these pictures." The examiner then presents a picture series. "Take a little time to look at these pictures. They tell a story. Take a look at all of them, and then I'll ask you to tell me the story with a beginning, a middle, and an end. You can look at the pictures as you tell the story." If no response was given in 10 s, the examiner gave a second prompt: "Take a look at this picture (point to first picture) and tell me what you think is happening." If needed, the examiner pointed to each picture sequentially, giving the prompt: "And what happens here?" For each panel, if no response, the examiner provided the prompt: "Can you tell me anything about this picture?" If no response was given to any of these prompts, the examiner utilized a set of troubleshooting questions (available on aphasia.talkbank.org).

Gesture Coding

Gesture Types

Iconic gestures represent meaning that is closely related to the semantic content of the speech that they accompany—indeed, these forms of gesture can only be interpreted within the context of speech, unlike other gestures that are imitations of motor actions (i.e., pantomimes) or carry culturally specific meaning on their own (i.e., emblems; McNeill, 1992). We drew our definition and coding parameters from Sekine and Rose (2013), classifying viewpoint, referential, and metaphoric gesture types as iconic. Viewpoint gestures are those that depict a concrete action, event, or object as though the speaker is observing it from afar (observer viewpoint) or as though he is the character/ object itself (character viewpoint; McNeill, 1992). Referential gestures are those that place objects, places, or characters in the story into the space in front of a speaker where any concrete object is absent, typically taking the form of a pointing gesture or of the hand "holding" some entity (Gullberg, 2006). Metaphoric gestures convey meaning in a nonliteral, abstract way and are typically regarded as a

Table 1. Demographics of subject group (*N* = 75).

Demographics	*M* (*SD*) or frequency
Age (years)	60.70 (11.22)
Education (years)	15.43 (2.58)
Aphasia severity (WAB AQ)[a]	73.73 (14.37)
Aphasia duration (years)	5.01 (4.31)
Years of speech-language pathology therapy	3.26 (2.44)
Race	8 African American
	1 Asian
	1 Native Hawaiian or Other Pacific Islander
	62 White
Ethnicity	3 Hispanic or Latinx
Gender[b]	34 females
Handedness (premorbid)	3 ambidextrous
	7 left-handed
	64 right-handed
	1 unknown
Language status	6 childhood bilinguals (English plus 2nd language by 6 years old)
	6 late bilinguals (English plus 2nd language after 6 years old)
	62 monolinguals
	1 multilingual (speaks 3 or more languages fluently)
Presence of dysarthria and/or apraxia of speech	43 with apraxia of speech
	8 with dysarthria (3 unknown)
Presence of hemiparesis or hemiplegia	23 no motor impairment
	21 right-sided hemiplegia (i.e., paralysis)
	28 right-sided hemiparesis (i.e., weakness)
	2 left-hemisphere hemiparesis (i.e., weakness)
	1 unknown
Aphasia etiology	73 stroke, 2 other or unknown
Types of aphasia	27 anomic
	19 Broca's
	18 conduction
	0 global
	0 mixed transcortical
	3 transcortical motor
	1 transcortical sensory
	2 Wernicke's
	5 not aphasic by WAB (i.e., scoring > 93.8 on WAB)

Note. WAB AQ = Western Aphasia Battery Aphasia Quotient.

[a]As measured by the Western Aphasia Battery–Revised Aphasia Quotient, where 100 = no aphasia. [b]"Gender" is the terminology used by AphasiaBank as of publication, but only two options were given to respondents: male or female. This therefore may not reflect other genders (e.g., nonbinary, transgender) reflected in the data set.

---

form of iconic gesture with abstract content (Kita, 2014). Note that, during our gesture coding, we did not differentiate between referential and metaphoric gestures. These were scored as a single type (referential + metaphoric type). Examples of iconic gestures demonstrated in this study are shown in Supplemental Material S1.

Note that we also coded three other gesture types, which are not reported in detail here but can be found in Supplemental Material S2. These were concrete deictic (i.e., a concrete referent in the physical environment, such as a picture; McNeill, 1992), emblem (i.e., a culturally specific gesture where form and meaning can usually be understood without speech; Kendon, 1980), and number (i.e., using the speaker's fingers to represent numbers; Cicone et al., 1979).

Gesture Coding

Gestures were coded using the online Browsable Database on AphasiaBank. The Browsable Database time locks each spoken utterance with the video of the participant, allowing the experimenter to watch the video while also reading verbatim transcribed speech. C. C. (second author) was the primary coder. She watched all videos through once prior to coding. Then, she watched each video in 10-s chunks, tallying frequencies by gesture type (viewpoint, referential + metaphoric, concrete deictic, number, and emblem). Each video was coded in this way twice, to check for coding consistency. A unit of gesture was defined as the duration from the start of a movement until the hand(s) returned to its resting position (McNeill, 1992). If the hand(s) did not return to the resting position, gestures were divided by either a pause in the movement or an obvious change in shape or trajectory (Jacobs & Garnham, 2007). Note that many gesture types were not recorded for the purposes of this study, including beats and other nongesture movements (e.g., self-adjusting, touching hair, touching table).

Raters and Reliability

The primary rater (C. C.) trained two additional undergraduate raters in gesture coding. These raters practiced coding on a random sample of 10 subjects (they trained on a combination of samples from an aphasia group and a neurotypical group from AphasiaBank). C. C. then cross-checked coding to establish interrater reliability of at least 80% agreement on these practice samples. Any disagreements were discussed between all raters, and a consensus was arrived at. Following resolution of any outstanding issues, raters were then assigned approximately 25 total individuals (22% of total sample, not including any samples used to practice) to rate. Though reliability was established on a training set, we further established interrater reliability on 10% of this sample ($N = 3$ subjects). Interrater reliability exceeded a Pearson's correlation of $r > .75$ between primary rater C. C. and each auxiliary rater, for frequency of each gesture type. In summary, all raters trained on a training set prior to rating achieving > 80% agreement across the entire training set, and then interrater agreement was calculated for 10% of the study sample, demonstrating that training set reliability extended to the study sample. Given complexities and idiosyncrasies of gesture rating (e.g., some level of subjectivity, as described by McNeill, 1992), we consider this favorable evidence of rater reliability, while also acknowledging the small n on which reliability was computed.

Spoken Language Data and Fluency

To evaluate our second research question (evaluate the relationship between iconic gesture frequency and rate with spoken language), we acquired information about spoken language produced during the discourse tasks and extracted extra-task information about language fluency from a standardized testing battery. Our primary interest was to measure the relationship between iconic gesture and language dysfluency, by task. When gesture is prohibited in neurotypical speakers, more dysfluencies (e.g., pauses, slower speech rate, false starts, fillers) are produced. In persons living with aphasia, increased frequency of gesture is thought to improve lexical retrieval. That is, gestures are produced in higher rates when language is particularly dysfluent. Here, we postulate that production of iconic gestures will accompany greater dysfluencies, in accordance with this hypothesis. For the purposes of this study, we explore common dysfluencies found in aphasic speech and that have been previously evaluated in studies of gesture use in persons with aphasia and neurotypical participants (Carlomagno & Cristilli, 2006; de Beer et al., 2019; Feyereisen, 1983; Kong et al., 2017, 2015; Rauscher et al., 1996): *hesitations* (composed of filled and unfilled pauses of greater than 3 s, false starts, and fillers) and *words per minute* (a metric of speaking rate). Hesitations were computed as a proportion of total words produced, inclusive of retracings and repetitions. We likewise evaluated metrics of gross language output, including *speaking time* and *total utterances*, to evaluate the relationship of iconic gesture with speech. To obtain this

spoken language data, we extracted variables from each task's language sample using CLAN (MacWhinney, 2018).

Complementary to the linguistic information extracted from the discourse tasks, we also used a metric of language fluency obtained from a neuropsychological test battery as a means of evaluating extra-task spoken language fluency (i.e., a metric of language fluency derived from a different task and scored based on a standardized scale). Specifically, we extracted extra-task information about language fluency from the Fluency, Grammatical Competence, and Paraphasias of Spontaneous Speech score (we will call this the "Fluency" score for short) of the Western Aphasia Battery–Revised (WAB-R). The Fluency score is obtained by first having subjects answer conversational questions (e.g., "Have you been here before?" and "What is your occupation?"), followed by having subjects describe a picture. The test administrator then ranks subjects' fluency from 0 to 10, described below. Using these Fluency scores, we created three fluency groups: a group containing subjects with low fluency ("low fluency" group, scores 0–4); a group containing subjects who are fluent but whose speech is characterized by errors, jargon, and empty speech ("fluency with errors" group, scores 5–7); and a group whose subjects are highly fluent with few errors ("high functional fluency" group, scores 8–10). The fluency descriptions from the WAB-R are briefly described (Kertesz, 2007):

- Low fluency group (scores 0–4): A score of 0 reflects no words or short, meaningless utterances; a score of 1 reflects recurrent, brief, stereotypic utterances with varied intonation, where emphasis or prosody may carry some meaning; a score of 2 reflects single words, usually errors (paraphasias), with speech being effortful and hesitant; a score of 3 reflects longer, recurrent stereotypic or automatic utterances lacking in information; and a score of 4 reflects halting, telegraphic speech containing mostly single words, paraphasias, and some prepositional phrases, with severe word-finding difficulty. A stipulation for a score of 4 is that no more than two complete sentences may be present (excepting automatic sentences).

- Fluent with errors group (scores 5–7): A score of 5 reflects telegraphic but more fluent speech, containing some grammatical organization, though still with marked word-finding difficulty and paraphasias; a score of 6 reflects more propositional sentences with normal syntactic patterns, coupled with paraphasias (optional) and significant word-finding difficulty and/or hesitations; and a score of 7 reflects fluent, phonemic jargon with semblance to English syntax and rhythm, alongside varied phonemes and neologisms.

- High functional fluency group (scores 8–10): A score of 8 reflects circumlocutory, fluent speech, with moderate word-finding difficulty, with or without paraphasias, and may present with semantic jargon; a score of 9 reflects mostly complete, relevant sentences, with occasional hesitations and/or paraphasias and some word-finding difficulty; and a score of 10 reflects

sentences of normal length and complexity, without definite slowing, halting, or paraphasias.

Analyses and Research Questions

All analyses were conducted in SPSS 27. The data were, overall, not normal in distribution, and we therefore employed nonparametric statistics (specifically, Spearman correlation, chi-square tests, Wilcoxon signed-ranks tests for repeated measures, and Kruskal–Wallis H tests of independent samples). Specific tests are described in more detail in each appropriate section. All analyses were corrected for multiple comparisons using Bonferroni correction; specific correction p values are noted in the text. The current study was not preregistered, which we acknowledge as a limitation.

Recall that we had two primary research questions. The first research question is theoretically driven: Do iconic gesture frequency and rate significantly differ between two narrative tasks, that is, is there a main effect of task on gesture, similar to the effect of task demonstrated on the speech of persons living with aphasia (Stark, 2019; Stark & Fukuyama, 2020)? The second research question is also theoretically driven, given prior work examining the relationship of iconic gesture with spoken language in aphasia (Dipper et al., 2011; Pritchard et al., 2015; Sekine & Rose, 2013). Therefore, our second research question is "To what extent are task-specific iconic gesture frequency and rate related to spoken language?" As such, the primary dependent variables of this study were iconic gesture *frequency* (i.e., how many gestures were produced) and *rate* (i.e., how many gestures were produced per spoken word, including repetitions/retracings of words and paraphasias with no known targets; see Table 1).

In addition to evaluating our primary research questions, we explore the relationship of iconic gesture frequency with demographic variables (e.g., age). We additionally provide exploratory analyses that divide iconic gestures into two types ([a] viewpoint and [b] referential + metaphoric), which is motivated by differences in theoretical constructs between gesture types. Given that the study of gesture in aphasia is a blossoming field, it remains unclear the extent to which factors such as gesture type and/or discourse task may be meaningful in delineating the relationship of spoken language and gesture usage in this population. We did not explicitly differentiate between referential and metaphoric gestures for the purposes of this study, and thus, they will be presented as a combined group. The goal of this exploratory analysis is to provide preliminary evidence for future investigations about the relationship of iconic gesture type frequency with task and with language variables, in persons with aphasia.

Results

Below, we present analyses for the whole ($N = 75$) sample. Additionally, we present results for a smaller sample ($N = 43$), who gestured during both Window and Sandwich

tasks; we call this sample the "matched" group. We present both sets of results to provide future researchers with a wealth of data. Throughout the results, we will draw attention to instances where the findings from these two groups diverge.

Relationship Between Iconic Gesture and Demographic Information

We evaluated the relationship between iconic gesture frequency and demographic and neuropsychological information. After multiple comparison correction using Bonferroni correction ($p < .01$, five comparisons), none of the following significantly associated with total iconic gesture frequency, summed across the two tasks: aphasia severity (measured by WAB-R Aphasia Quotient; Kertesz, 2007; $N = 75$, $r_s = .10$, $p = .38$), years of speech-language therapy ($N = 75$, $r_s = .21$, $p = .07$), education ($N = 75$, $r_s = .21$, $p = .08$), or age ($N = 75$, $r_s = -.26$, $p = .02$). Aphasia duration—that is, the amount of time one has lived with aphasia—did significantly correlate with total iconic gesture frequency ($N = 75$, $r_s = .36$, $p = .002$), indicating that persons with more chronic aphasia tended to produce iconic gestures more frequently. Further exploration by discourse task suggested that the relationship between aphasia duration and gesture frequency was mostly related to iconic gesturing on the Window task ($N = 75$, $r_s = .37$, $p = .001$) and not the Sandwich task ($N = 75$, $r_s = .21$, $p = .07$). Hemiparesis could be one reason why gestures are not produced frequently or at a high rate (although this typically is not the case; e.g., Kong et al., 2015). We evaluated the impact of physical status (no motor impairment, unilateral hemiparesis, unilateral hemiplegia) on iconic gesture frequency pooled across tasks, finding no significant difference in iconic gesture frequency across different motor impairments (independent-samples Kruskal–Wallis test, $N = 74$ [1 missing data point], $W = 2.16$, $p = .34$). We also performed an explorative analysis on the impact of gender on iconic gesture frequency, as there is evidence from psychology that women tend to employ more nonverbal communication (Hall & Gunnery, 2013), finding no significant difference in iconic gesture frequency between males and females (Mann–Whitney U test, $N = 75$, $U = 0.63$, $p = .53$).

We likewise evaluated the relationship of gesture rate with demographic and neuropsychological information. After multiple comparison correction using Bonferroni correction ($p < .01$, five comparisons), average iconic gesture rate (i.e., rate averaged across tasks) was found to not significantly associate with aphasia severity ($N = 75$, $r_s = -.19$, $p = .10$), age ($N = 75$, $r_s = -.28$, $p = .01$), or education ($N = 75$, $r_s = -.11$, $p = .36$), but significantly with years of speech-language therapy ($N = 75$, $r_s = .37$, $p = .001$) and aphasia duration ($N = 75$, $r_s = .40$, $p < .001$). Further exploration by task indicated that aphasia duration was positively related to gesture rate for each task (Window, $N = 75$, $r_s = .40$, $p < .001$; Sandwich, $N = 75$, $r_s = .25$, $p = .03$), while the relationship between years of speech-language therapy and gesture rate was driven by performance on the

Sandwich task (Window, $N = 75$, $r_s = .22$, $p = .06$; Sandwich, $N = 75$, $r_s = .31$, $p = .007$). We did find a significant relationship between motor impairment and gesture rate (independent-samples Kruskal–Wallis test, $N = 74$ [1 missing data], $W = 13.85$, $p = .001$). Pairwise comparisons (corrected at Bonferroni $p < .01$ for three comparisons) in $N = 74$ showed that persons with hemiplegia tended to gesture at a higher rate than persons with no motor impairment ($p < .001$), which seems counterintuitive but makes sense in the context that those with hemiplegia also tend to be those who produce fewer words ($W = 7.47$, $p = .02$; where persons with hemiplegia produced significantly fewer words than those with no motor impairment, $p = .006$). We did not find a significant difference in gesture rate between males and females (Mann–Whitney U test, $N = 75$, $U = -1.87$, $p = .06$).

Note that our matched group ($N = 43$) had a similar correlation matrix. Correlation tables between demographic variables, language variables, and gesture variables are shown in Supplemental Materials S3 ($N = 75$ group) and S4 ($N = 43$ group).

Main Effect of Task on Iconic Gesture

Here, we evaluate our primary research question, which is the main effect of task on iconic gesture frequency and rate. Of our entire sample ($N = 75$), 57.3% of subjects made an iconic gesture during the Window task, while 100% made an iconic gesture during the Sandwich task (see Table 2).

Employing Wilcoxon signed-ranks tests, which is a nonparametric repeated-measures analysis, we identified a significant difference between tasks in gesture frequency ($N = 75$, $Z = 3.45$, $p = .001$; $N = 43$, $Z = 4.98$, $p < .001$) and gesturing rate ($N = 75$, $Z = 7.36$, $p < .001$; $N = 43$, $Z = 5.42$, $p < .001$; Bonferroni, $p < .025$, two comparisons per group; see Table 3). For a distribution of gesture frequency and rate in all subjects ($N = 75$), see Figure 1. These results demonstrate a clear main effect of task, specifically greater iconic gesture frequency and rate during the Sandwich task.

An exploratory analysis to identify the extent to which iconic gesture type frequency differed between tasks demonstrated that referential + metaphoric gestures were produced significantly more often during the Sandwich task ($N = 75$, $Z = -6.48$, $p < .001$; $N = 43$, $Z = -4.73$, $p < .001$), and this was also true for viewpoint gestures, which were made more often during the Sandwich task ($N = 75$, $Z = -6.33$, $p < .001$; $N = 43$, $Z = -3.97$, $p < .001$; Bonferroni, $p < .025$,

two comparisons per group). This finding shows that not a single iconic gesture type drove the difference in gesture frequency between the tasks.

Further exploratory analyses of iconic gesture rate in the whole group demonstrated that referential + metaphoric gestures were produced at a significantly higher rate during the Sandwich task ($N = 75$, $Z = 6.43$, $p < .001$; $N = 43$, $Z = -4.63$, $p < .001$), and this was likewise the case for viewpoint gestures ($N = 75$, $Z = -6.91$, $p < .001$; $N = 43$, $Z = -4.76$, $p < .001$; Bonferroni, $p < .025$, two comparisons per group). This finding again suggests that not a single iconic gesture type drove the difference in gesture rate between the tasks.

Relationship Between Iconic Gesture and Spoken Language

Relationship of Iconic Gesture With Fluency Variables Extracted From Discourse

Our second research prerogative was to identify relationships between spoken language and iconic gesture frequency and rate. Given a main effect of task, we evaluated the relationship between iconic gesture and language variables, stratified by task.

In regard to task-related differences in spoken language, Wilcoxon signed-ranks tests evaluating repeated measures demonstrated that the Window task had a significantly longer speaking time than the Sandwich task ($N = 75$, $Z = -2.65$, $p = .008$; $N = 43$, $Z = -2.88$, $p = .004$), but there were no significant differences between the two tasks for any other language metric: total utterances ($N = 75$, $Z = -2.01$, $p = .045$; $N = 43$, $Z = -2.10$, $p = .036$), words per minute ($N = 75$, $Z = -0.75$, $p = .46$; $N = 43$, $Z = -0.27$, $p = .79$), or proportion hesitations ($N = 75$, $Z = 0.69$, $p = .49$; $N = 43$, $Z = 1.49$, $p = .14$; Bonferroni correction, $p < .0125$, four comparisons per subject group). We show descriptive statistics and statistical information for the whole group ($N = 75$) and the matched group ($N = 43$) in Table 4.

Gesture Frequency's Relationship With Language Fluency Variables

For the Sandwich task, we identified a significant relationship between gesture frequency and speaking time ($N = 75$, $r_s = .62$, $p < .001$; $N = 43$, $r_s = .65$, $p < .001$) and total utterances ($N = 75$, $r_s = .69$, $p < .001$; $N = 43$, $r_s = .63$, $p < .001$), but not words per minute ($N = 75$, $r_s = .10$, $p = .39$; $N = 43$, $r_s = .07$, $p = .67$) or proportion hesitations

Table 2. Descriptive statistics divided by task for full sample ($N = 75$).

Variable	Window task M (SD), range	Sandwich task M (SD), range
N using at least 1 gesture	43 (57.3%)	76 (100%)
Task duration (s)	54.51 (37.33), 10–228	44.96 (29.00), 5–140
Iconic gesture frequency	1.84 (2.59), 0–11	7.11 (4.39), 1–34
Iconic gesture rate	0.03 (0.06), 0–0.27	0.16 (0.09), 0.03–0.41

Table 3. Demographics for subgroup making an iconic gesture on both tasks (*N* = 43).

Demographics	*M* (*SD*) or frequency
Age (years)	57.94 (10.69)
Education (years)	15.62 (2.53)
Aphasia severity (WAB AQ)[a]	73.53 (13.81)
Aphasia duration (years)	5.91 (5.04)
Years of speech-language pathology therapy	3.59 (2.69)
Race	5 African American
	1 Asian
	1 Native Hawaiian or Other Pacific Islander
	36 White
Ethnicity	0 Hispanic or Latinx
Gender[b]	21 females
Handedness (premorbid)	3 ambidextrous
	5 left-handed
	34 right-handed
	1 unknown
Language status	3 childhood bilinguals (English plus 2nd language by 6 years old)
	5 late bilinguals (English plus 2nd language after 6 years old)
	34 monolinguals
	1 multilingual (speaks 3 or more languages fluently)
Presence of dysarthria and/or apraxia of speech	25 with apraxia of speech
	3 with dysarthria (1 unknown)
Presence of hemiparesis or hemiplegia	11 no motor impairment
	11 right-sided hemiplegia (i.e., paralysis)
	19 right-sided hemiparesis (i.e., weakness)
	2 left-hemisphere hemiparesis (i.e., weakness)
Aphasia etiology	42 stroke, 1 other or unknown
Types of aphasia	18 anomic
	12 Broca's
	10 conduction
	0 global
	0 mixed transcortical
	0 transcortical motor
	1 transcortical sensory
	1 Wernicke's
	1 not aphasic by WAB (i.e., scoring > 93.8 on WAB)

*Note.* WAB AQ = Western Aphasia Battery Aphasia Quotient.

[a]As measured by the Western Aphasia Battery–Revised Aphasia Quotient, where 100 = no aphasia. [b]"Gender" is the terminology used by AphasiaBank as of publication, but only two options were given to respondents: male or female. This therefore may not reflect other genders (e.g., nonbinary, transgender) reflected in the data set.

(N = 75, r_s = −.08, p = .49; N = 43, r_s = .06, p = .72; Bonferroni correction, p < .0125, four comparisons per subject group). As an exploratory analysis, we evaluated the correlation between language metrics and frequency of gesture types (viewpoint, referential + metaphoric), finding that viewpoint gesture frequency during the Sandwich task significantly correlated with speaking time (N = 75, r_s = .63, p < .001; N = 43, r_s = .56, p < .001) and total utterances (N = 75, r_s = .59, p < .001; N = 43, r_s = .59, p < .001), but not with our fluency variables: words per minute (N = 75, r_s = −.17, p = .16; N = 43, r_s = −.19, p = .21) or proportion hesitations (N = 75, r_s = .18, p = .13; N = 43, r_s = .32, p = .04; Bonferroni correction, p < .0125, four comparisons per subject group). Referential + metaphoric gesture frequency during Sandwich also significantly correlated with speaking time (N = 75, r_s = .30, p = .008; N = 43, r_s = .42, p = .005) and total utterances (only for the large group; N = 75, r_s = .42, p < .001; N = 43, r_s = .35, p = .02) and, in contrast to viewpoint gesture frequency,

approached a significant correlation with proportion hesitations (N = 75, r_s = −.29, p = .013; N = 43, r_s = −.24, p = .13) and words per minute (N = 75, r_s = .27, p = .019; N = 43, r_s = .28, p = .07; Bonferroni correction, p < .0125, four comparisons per subject group).

For the Window task, we found a significant relationship between gesture frequency and speaking time (only for the large group; N = 75, r_s = .34, p = .003; N = 43, r_s = .29, p = .06) and total utterances (only for the large group; N = 75, r_s = .45, p < .001; N = 43, r_s = .32, p = .03), but not words per minute (N = 75, r_s = −.09, p = .44; N = 43, r_s = −.24, p = .13) or proportion hesitations (N = 75, r_s = −.13, p = .25; N = 43, r_s = .006, p = .97; Bonferroni correction, p < .0125, four comparisons per subject group). As an exploratory analysis, we evaluated the correlation between language metrics and frequency of gesture types (viewpoint, referential + metaphoric), finding that viewpoint gesture frequency during the Window task significantly correlated with total utterances (only for the large group; N = 75, r_s = .36,

Figure 1. Distribution of gesture frequency (A) and gesture rate (B) on both tasks (*N* = 75). Data callouts indicate average, while each data point (and its relative intensity/shadow) indicates individual subject scores.
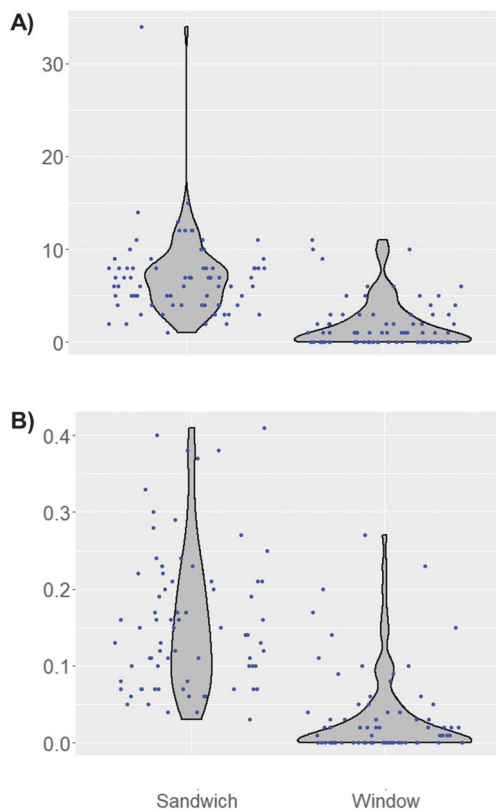
p = .001; *N* = 43, r_s = .21, *p* = .18) but not speaking time (*N* = 75, r_s = .23, *p* = .04; *N* = 43, r_s = .15, *p* = .33), words per minute (*N* = 75, r_s = −.10, *p* = .39; *N* = 43, r_s = −.21, *p* = .19), or proportion hesitations (*N* = 75, r_s = −.01, *p* = .92; *N* = 43, r_s = .19, *p* = .24; Bonferroni correction, *p* < .0125, four comparisons per subject group). Referential + metaphoric gesture frequency significantly correlated with total utterances (*N* = 75, r_s = .45, *p* < .001; *N* = 43, r_s = .40, *p* = .007) and speaking time (*N* = 75, r_s = .39, *p* = .001; *N* = 43, r_s = .40, *p* = .008), but not words per minute (*N* = 75, r_s = −.07, *p* = .57; *N* = 43, r_s = −.14, *p* = .36) or proportion hesitations (*N* = 75, r_s = −.29, *p* = .01; *N* = 43, r_s = −.30, *p* = .06; Bonferroni correction, *p* < .0125, four comparisons per subject group).

These results support a relationship between frequency of iconic gestures (regardless of gesture type) and gross speech output, but not necessarily with metrics of speech fluency. This was the case for both tasks.

Gesture Rate's Relationship With Language Fluency Variables

We then evaluated the relationship between gesture rate and language variables, finding that gesture rate during the Sandwich task demonstrated a significant relationship with our fluency variables, words per minute (only for the large group; *N* = 75, r_s = −.47, *p* < .001; *N* = 43, r_s = −.36, *p* = .017) and proportion hesitations (only for the large group; *N* = 75, r_s = .34, *p* = .003; *N* = 43, r_s = .37, *p* = .02), but did not significantly correlate with our gross output variables, speaking time (*N* = 75, r_s = −.06, *p* = .64; *N* = 43, r_s = −.16, *p* = .32) and or total utterances (*N* = 75, r_s = −.21, *p* = .07; *N* = 43, r_s = −.26, *p* = .09; Bonferroni correction, *p* < .0125, four comparisons per subject group). Exploratory analyses demonstrated that viewpoint gesture rate did not significantly associate with speaking duration (*N* = 75, r_s = .04, *p* = .71; *N* = 43, r_s = −.06, *p* = .70) or total utterances (*N* = 75, r_s = −.13, *p* = .25; *N* = 43, r_s = −.10, *p* = .53); approached a significant relationship with words per minute in the matched group; was significantly related to words per minute in the whole group (*N* = 75, r_s = −.49, *p* < .001; *N* = 43, r_s = −.37, *p* = .015); and was significantly, positively related to proportion hesitations (*N* = 75, r_s = .43, *p* < .001; *N* = 43, r_s = .44, *p* = .003; Bonferroni correction, *p* < .0125, four comparisons per subject group). Referential + metaphoric gesture rate was not significantly associated with speaking duration (*N* = 75, r_s = −.007, *p* = .95; *N* = 43, r_s = −.001, *p* > .99), total utterances (*N* = 75, r_s = −.05, *p* = .68; *N* = 43, r_s = −.16, *p* = .32), words per minute (*N* = 75, r_s = −.11, *p* = .33; *N* = 43, r_s = −.08, *p* = .63), or proportion hesitations (*N* = 75, r_s = −.03, *p* = .83; *N* = 43, r_s = .007, *p* = .97; Bonferroni correction, *p* < .0125, four comparisons per subject group). In general, these results support a relationship between language fluency variables and gesture rate during the Sandwich task, which may be in large part driven by viewpoint gesture rate.

We then evaluated the relationship between gesture rate and language variables during the Window task, finding a significant correlation with words per minute (only for the matched group; *N* = 75, r_s = −.19, *p* = .097; *N* = 43, r_s = −.54, *p* < .001) and total utterances (only for the large group; *N* = 75, r_s = .34, *p* = .003; *N* = 43, r_s = −.04, *p* = .82; Bonferroni correction, *p* < .0125, four comparisons per subject group). We did not identify a significant relationship with speaking time (*N* = 75, r_s = .24, *p* = .04; *N* = 43, r_s = −.01, *p* = .94) or proportion hesitations (*N* = 75, r_s = −.03, *p* = .82; *N* = 43, r_s = .32, *p* = .03; Bonferroni correction, *p* < .0125, four comparisons per subject group). Exploratory analyses demonstrated that viewpoint gesture rate during Window did not significantly associate with speaking duration (*N* = 75, r_s = .06, *p* = .62; *N* = 43, r_s = −.04, *p* = .79) or total utterances (*N* = 75, r_s = .21, *p* = .07; *N* = 43, r_s = −.06, *p* = .70) but was significantly related to language fluency in the matched group for words per minute (*N* = 75, r_s = −.07, *p* = .53; *N* = 43, r_s = −.37, *p* = .013) and proportion hesitations (*N* = 75, r_s = .06, *p* = .63; *N* = 43, r_s = .37, *p* = .01; Bonferroni correction, *p* < .0125, four comparisons per subject group). This pattern was similar for referential + metaphoric gesture rate, which did not significantly associate with speaking duration (*N* = 75, r_s = .12, *p* = .29; *N* = 43, r_s = .02, *p* = .91) or words per minute (*N* = 75, r_s = .02, *p* = .84; *N* = 43, r_s = −.19, *p* = .23)

Table 4. Descriptive and comparative statistics for language and gesture variables divided by task (N = 43).

Variable	Window task M (SD), range		Sandwich task M (SD), range		Repeated measures statistics
Group	**N = 75 whole**	**N = 43 matched**	**N = 75 whole**	**N = 43 matched**	**By group (N = 75; N = 43)**
Task duration (s	54.51 (37.33), 10–228	59.67 (31.18), 10–135	44.96 (29.00), 5–140	45.86 (25.49), 5–111	N = 75, Z = −2.65, p = .008 N = 43, Z = −2.88, p = .004*
Language variables					
Utterances	8.81 (4.71), 2–27	10.21 (4.83), 3–27	7.56 (4.25), 2–19	8.37 (4.02), 2–19	N = 75, Z = −2.01, p = .05 N = 43, Z = −2.10, p = .04
Words per minute	65.37 (39.82), 12–194.12	65.06 (40.29), 17.14–194.12	62.90 (39.15), 13.17–213.33	63.24 (31.82	N = 75, Z = −0.75, p = .46 N = 43, Z = −0.27, p = .79
Hesitations (proportion of total words)[a]	0.41 (1.01), 0–7	0.24 (0.46), 0–3	0.23 (0.19), 0–.97	0.22 (0.18), 0–0.68	N = 75, Z = 0.69, p = .49 N = 43, Z = 1.49, p = .14
Iconic gesture variables					
Frequency	1.84 (2.59), 0–11	3.21 (2.71), 1–11	7.11 (439), 1–34	7.35 (3.24), 2–15	N = 75, Z = −7.04, p < .001* N = 43, Z = 4.98, p < .001*
Rate	0.03 (0.06), 0–0.27	0.06 (0.06), 0–0.27	0.16 (0.09), 0.03–0.41	0.15 (0.09), 0.04–0.41	N = 75, Z = 7.36, p < .001* N = 43, Z = 5.42, p < .001*

[a]Hesitations can have a larger proportion than 1 when they occur at a greater rate than total words. Hesitations include filled and unfilled pauses of greater than 3 s, false starts, and fillers.
*Significant at p < .007 (Bonferroni correction across seven repeated measures) per group (whole, matched).

but which approached a significant relationship with total utterances ($N = 75$, $r_s = .27$, $p = .02$; $N = 43$, $r_s = .04$, $p = .79$) and was significantly related to proportion hesitations (whole group only; $N = 75$, $r_s = -.27$, $p = .019$; $N = 43$, $r_s = -.24$, $p = .13$; Bonferroni correction, $p < .0125$, four comparisons per subject group). Therefore, it appears that gesture rate's relationship with language may be influenced by both types of iconic gesture. Globally, these results support a relationship between iconic gesturing rate and speech, with some task specificity: a relationship of gesture rate with speech fluency in the Sandwich task and with both speech fluency and gross speech output in the Window task.

Relationship of Iconic Gesture With Extra-Task Fluency

As a complementary analysis to evaluate language fluency, we evaluated the relationship of iconic gesture with an extra-task metric of spoken language fluency by evaluating iconic gesture use in three fluency groups (from the large group [$N = 75$] and the matched group [$N = 43$]) derived from a standardized aphasia battery: low fluency ($N = 75$ group, 22 subjects; $N = 43$ group, 12 subjects), fluent with errors ($N = 75$ group, 21 subjects; $N = 43$ group, 13 subjects), and high functional fluency ($N = 75$ group, 32 subjects; $N = 43$ group, 18 subjects). To do so, we computed Kruskal–Wallis H tests of independent samples (sometimes also called the "one-way analysis of variance on ranks"), evaluating significant differences in gesture rate and frequency by fluency group per task (corrected at Bonferroni, $p < .025$, two comparisons per subject group). Note that fluency groups did not significantly differ by age ($N = 75$, $H = 0.88$, $p = .65$; $N = 43$, $H = 0.14$, $p = .93$), though did significantly differ by aphasia duration in the large subject group ($N = 75$, $H = 7.59$, $p = .02$; $N = 43$, $H = 5.57$, $p = .06$; corrected at Bonferroni, $p < .025$, two comparisons per subject group).

Gesture Frequency's Relationship With Fluency Group

We did not identify a significant fluency group effect on iconic gesture frequency for either Window ($N = 75$, $H = 0.32$, $p = .85$; $N = 43$, $H = 4.76$, $p = .09$) or Sandwich ($N = 75$, $H = 5.09$, $p = .08$; $N = 43$, $H = 4.64$, $p = .10$; corrected at Bonferroni, $p < .025$, two comparisons per subject group). As an exploratory analysis, we looked at the relationship between fluency groups and iconic gesture type frequency (viewpoint, referential + metaphoric) by task. For the Window task, we did not identify a significant effect of fluency group on referential + metaphoric gesture ($N = 75$, $H = 0.015$, $p = .99$; $N = 43$, $H = 0.33$, $p = .85$) or viewpoint gesture ($N = 75$, $H = 0.46$, $p = .79$; $N = 43$, $H = 2.41$, $p = .30$; corrected at Bonferroni, $p < .025$, two comparisons per subject group). For the Sandwich task, there was no effect (whole group) yet a significant effect (matched group) of fluency group on referential + metaphoric gesture frequency ($N = 75$, $H = 5.43$, $p = .066$; $N = 43$, $H = 7.74$, $p = .021$), but no significant effect of fluency
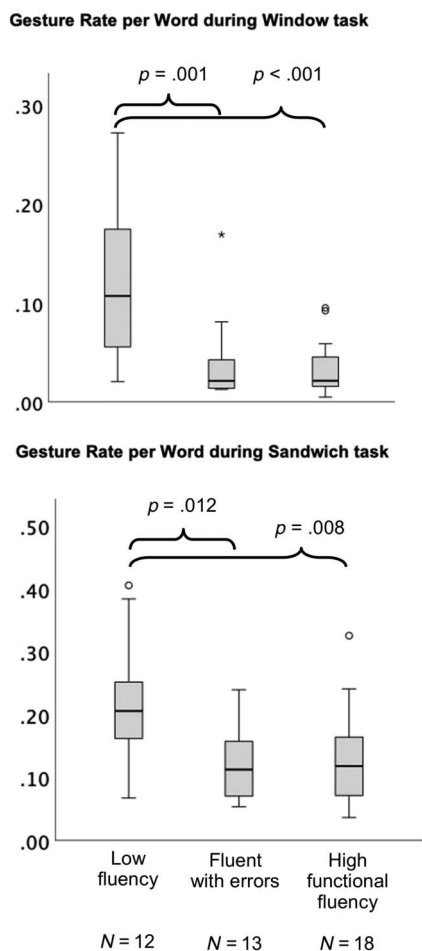
group was identified for viewpoint gesture frequency ($N = 75$, $H = 0.496$, $p = .78$; $N = 43$, $H = 0.895$, $p = .64$; corrected at Bonferroni, $p < .025$, two comparisons per subject group). Pairwise comparisons corrected at Bonferroni $p < .017$ (three comparisons) for the referential + metaphoric gestures made during the Sandwich task by the matched group ($N = 43$) demonstrated that this effect was driven by a difference between low fluency and high functional fluency groups ($p = .029$), where the low fluency group produced significantly more referential + metaphoric gestures. The fluent with errors group and the high functional fluency group did not differ significantly on referential + metaphoric gestures produced during Sandwich ($p = .14$), nor did the low and fluent with errors differ significantly on referential + metaphoric gestures ($p > .99$). Therefore, results suggest that referential + metaphoric gestures, more than viewpoint gesture frequency or total iconic gesture frequency, were produced significantly more often by the low fluency group during the Sandwich task.

Gesture Rate's Relationship With Fluency Group

We identified a significant difference in gesture rate by fluency group during the Window task ($N = 75$, $H = 2.30$, $p = .32$; $N = 43$, $H = 15.14$, $p = .001$) and Sandwich task ($N = 75$, $H = 5.98$, $p = .05$; $N = 43$, $H = 8.63$, $p = .013$). Note that, for both tasks, only the $N = 43$ group showed a significant relationship between fluency group and gesture rate, thus post hoc pairwise comparisons were done only for this group. Pairwise comparisons (corrected at Bonferroni $p < .017$, three comparisons) for gesture rate by fluency group on the Window task demonstrated a significantly higher gesture rate in the low fluency group compared with the fluent with errors group ($p = .001$) and high functional fluency group ($p < .001$; see Figure 2, top). There was not a significant difference between the fluent with errors and high functional fluency groups ($p = .84$). Exploration of pairwise comparisons (corrected at Bonferroni $p < .017$, three comparisons) for gesture rate by fluency group on the Sandwich task demonstrated the same pattern, with the low fluency group producing a significantly higher gesture rate compared with the fluent with errors group ($p = .012$) and compared with the high functional fluency group ($p = .008$; see Figure 2, bottom). Again, there was not a significant difference between the fluent with errors and high functional fluency groups ($p = .98$). Altogether, these results demonstrate that the low fluency group produced more gestures per word.

We then conducted exploratory analyses to evaluate the relationship between the fluency groups and gesture rate by gesture type (viewpoint, referential + metaphoric) and task. We did not identify a significant difference between fluency groups in rate of referential + metaphoric gestures, for either the Window ($N = 75$, $H = 0.21$, $p = .90$; $N = 43$, $H = 2.25$, $p = .33$) or the Sandwich ($N = 75$, $H = 1.26$, $p = .53$; $N = 43$, $H = 0.43$, $p = .81$) task. We identified a significant difference between fluency groups in the rate of viewpoint

Figure 2. Pairwise comparisons of gesture rate by fluency group, during the Window task (top) and the Sandwich task (bottom), for subgroup of subjects who made at least one iconic gesture on both tasks (*N* = 43).

Discussion

We evaluated the relationship of iconic gesturing with demographic, language, and task variables in a large group of persons with poststroke, chronic aphasia. Below, we will discuss how these results fit in within overarching, theoretical hypotheses of gesture, as well as prior findings of iconic gesture usage in aphasia. We will end with discussing clinical implications of this work.

Regarding demographic correlates of iconic gesture use, aphasia chronicity was correlated with iconic gesture usage, and this correlation was driven by iconic gesture frequency produced during the Window task. That is, those individuals who were living with aphasia for a longer time were those who tended to use a higher frequency of iconic gestures during the Window task. Generally, this may speak to gestures serving as a compensatory or supportive addition to their spontaneous speech (Dipper et al., 2015). However, more specifically, we identified a significant relationship with aphasia duration only in the Window task and not the Sandwich task, which may suggest an interaction of task demands, aphasia duration, and iconic gesture use. Further evaluation directly contrasting more tasks of varying types (e.g., narrative, procedural, expository, descriptive) and cognitive and linguistic demands (e.g., visual aid, verbal working memory use, semantic memory use, episodic memory use) will enhance our understanding of this possible, clinically relevant interaction. As we do not yet understand the extent to which iconic gesture use changes over time—there is a paucity of longitudinal studies on iconic gesture in aphasia recovery (Braddock, 2007)—cross-sectional designs such as the one suggested may facilitate hypothesis creation for future, longitudinal designs.

Notably, we did not identify a significant relationship between aphasia severity and iconic gesture frequency, which complements prior research that identified a relationship between aphasia severity and concrete deictic gestures but not iconic gestures (Sekine & Rose, 2013). That is, severe aphasia does not preclude iconic gesture use, which has clinical ramifications for assessment of gesture and treatment utilizing gesture as a predominant or complementary communicative modality (Rose et al., 2013).

Innovatively, we evaluated the main effect of task on iconic gesture in aphasia. Task has been shown to influence spoken language in persons with and without aphasia (Dalton & Richardson, 2019; Fergadiotis & Wright, 2011; Fergadiotis et al., 2011; Li et al., 1996; Shadden et al., 1990; Stark, 2019; Ulatowska et al., 1981; Wright & Capilouto, 2009), and it is perhaps not surprising that we identified a main effect of task on iconic gesture frequency and rate in our subject group of persons with aphasia. Specifically, subjects produced statistically more iconic gestures and gestured at a greater rate during the procedural narrative task (Sandwich) than during the picture sequence, expositional task (Window). There is a tradition in the gesture literature to use procedural tasks to evaluate gesture, especially in clinical samples (Cocks et al., 2007; Hilverman et al., 2016; Pritchard et al., 2015), and our current work underscores this trend,

gestures produced during the Window task for the matched group (*N* = 75, *H* = 2.29, *p* = .32; *N* = 43, *H* = 9.04, *p* = .01) but for neither subject group during the Sandwich task (*N* = 75, *H* = 3.84, *p* = .15; *N* = 43, *H* = 5.19, *p* = .08; corrected at Bonferroni, *p* < .025, two comparisons per subject group). Pairwise comparisons evaluating the rate of viewpoint gestures produced during the Window task for the matched group (corrected at Bonferroni *p* < .017, three comparisons) showed significant differences between the low fluency group and the fluent with errors group (*p* = .01) and between the low fluency group and the high functional fluency group (*p* = .006). In both cases, the low fluency group demonstrated a higher viewpoint gesture rate during the Window task. There was not a significant difference between the fluent with errors and the high functional fluency group (*p* = .96). In summary, total iconic gesture rate was higher in persons from the low fluency group, and when looking specifically at rates of gesture types, this seemed to be driven by the viewpoint gesture, particularly in the Window task.

as the procedural task (Sandwich) tended to associate with more iconic gestures and a higher gesture.

One explanation for the difference in iconic gesture production by task is shared knowledge (or common ground). That is, one task—the picture sequence task—provided a visual cue that was available to both the primary speaker and the other interlocutor (i.e., experimenter). Because the picture sequence was available to both persons in the experiment room, it may have been the case that fewer iconic gestures were produced because of shared knowledge (Bottenberg & Lemme, 1991; Campisi & Özyürek, 2013; but see Brenneise-Sarshad et al., 1991). Additionally, gestures tend to be smaller (Hoetjes et al., 2015; Kuhlen et al., 2012) and lower in the visual field (Hilliard & Cook, 2016) when sharing common ground. In the case of a shared visual cue, iconic gestures may not have been favored by either subject group because of a reliance on other types of gestures—specifically, concrete deictic gestures (Sekine & Rose, 2013). While we did not report on other gestures produced during these tasks here, as we wanted to focus on iconic gesture usage, we did collect data on concrete deictic (e.g., pointing) gestures, finding that many gestures produced by the aphasia group included pointing to specific parts of the Window picture (see Supplemental Material S2). We postulated earlier, in the introduction, that iconic gesturing may be used more often and with a greater success rate when tasks do not involve other visual stimuli (e.g., picture descriptions), as they stand in for the concrete imagery that may otherwise facilitate lexical access. It may likewise be that the presence of a visual support systematically varied task demands. For example, the picture may constrain what the participants said and also decrease memory demands. Prior work demonstrates that rich episodic detail gives rise to more gestures (Hilverman et al., 2016). Finally, the Sandwich task may have employed more spatial words and imagery given its procedural nature; iconic gesture frequency has been shown to positively correlate with lexical items related to space and movement (Atit et al., 2013; Hostetter & Alibali, 2019; Kita & Lausberg, 2008; Pritchard et al., 2015). Therefore, it may be the case that these two tasks are not directly comparable because of cognitive demands and task constraint, but results do seem to support the notion that more iconic gestures, at a greater rate, are produced during a narrative where no visual aids are present, which complements prior evidence.

For our second research question, we evaluated the relationship between spoken language and iconic gesture. There are a variety of theories exploring the relationship between spoken language and iconic gesture. These theories hypothesize, respectively, that gestures are used more frequently when spoken language production is made more difficult or otherwise impaired (de Ruiter, 2006; McNeill, 1992) and that using gestures facilitates lexical retrieval (Krauss, 1998). Because of the main effect of task, we evaluated relationships between iconic gesture and language by task. During the Window task, we did not identify a significant relationship between language variables and iconic

gesture frequency but did identify a relationship between iconic gesture rate and speaking rate. Specifically, with increasing speaking rate during the Window task, we observed a reduced gesture rate. This finding is complementary to the result that the low fluency group (recall, a group identified as low fluency from a standardized testing battery) produced significantly higher rates of iconic gestures during the Window task. That is, we identified that iconic gesture rate during the Window task associated with two metrics of dysfluency, one task-specific and one extra-task (standardized battery), together denoting that, at least for this task and its cognitive–linguistic demands, gesture rate was strongly, negatively related to spoken language fluency. During the Sandwich task, we identified a significant relationship between gross language and iconic gesture frequency, such that more iconic gestures were produced alongside longer speaking times and long utterances. The positive correlation between gross language and iconic gesture frequency in the Sandwich task may reflect increased task demands (e.g., no visual material available, limited common ground), culminating in more gestures produced. In regard to the relationship of iconic gesture with dysfluency during the Sandwich task, gesture rate approached a significant relationship with task-specific metrics of fluency (speaking rate, hesitations) and was significantly related to extra-task fluency (from standardized testing battery), complementing the profile shown in the Window results.

Together, these results support an overall relationship between iconic gesture rate and spoken language, especially fluency, which lends support for hypotheses related to shared preconceptual space for gesture and speech, as well as for a role of iconic gesture in enhancing lexical retrieval. It should be noted that most metrics of fluency are influenced by measures of grammatical competence, lexical retrieval, and speech production or a combination of these linguistic processes (Gordon & Clough, 2020), which speaks to the complexity of identifying the extent to which iconic gesture (for example) influences each linguistic process (e.g., grammatical competence, lexical retrieval, speech production). This work enhances prior work evaluating fluency based on aphasia types (e.g., nonfluent vs. fluent, Broca's vs. Wernicke's) by evaluating task-specific linguistic information and a standardized metric of spoken language fluency. Sekine and Rose (2013) correlated the Spontaneous Speech Fluency score (a composite score of the Fluency score described here, in addition to an Informational Content score) with iconic gesture frequency in their sample of persons with aphasia during a story retell narrative, finding a small–medium, significant, negative correlation between gesture frequency and Spontaneous Speech Fluency score. Their sample was larger, which may have contributed to improved power and identification of a small–medium relationship between an extra-task metric of fluency and iconic gesture frequency. Recall that we did not identify a relationship between gesture frequency and Fluency score. Note, too, that Sekine and Rose did not evaluate the relationship between gesture rate and their metric of fluency, however. It

should be noted that there is some difference in these comparisons. We chose the Fluency score—rather than the Spontaneous Speech Fluency score—because the former did not incorporate Informational Content. Informational Content score speaks more to the communication/functional value of the speech than the fluency, and the Informational Content score is typically highly related to the overall aphasia severity (Kertesz, 2007). Despite nuances, there is a clear relationship between iconic gesture (be it frequency or rate) and spoken language fluency.

Therefore, there is mounting evidence that discourse task affects spontaneous iconic gesture usage in persons with and without aphasia. We provide one of the few studies directly contrasting iconic gesture usage across two tasks in the same group of subjects, with hope that more work follows our own in exploring these task-specific gesture relationships. Exploring task-specific gesture relationships across all gesture types (i.e., along Kendon's continuum) gets at a more naturalistic evaluation of gesture, which allows us to jump from highly constrained experimental investigations of gesture to more naturalistic evaluation of gesture. While the main purpose of the current study was to look at task-related differences in iconic gesturing, Supplemental Material S2 makes clear that there are likely other task-related differences in other types of gestures (e. g., concrete deictic), and a future, multivariate investigation of the relationship between task and gesture type would be highly valuable in determining typical spontaneous gesture patterns in persons with aphasia.

We did not evaluate the informational relationship of each of iconic gestures to their respective speech (i.e., to disambiguate, to add information, to be redundant), but further evaluation of the informational relationship between speech and iconic gesture type is necessary to better understand the possible task-specific facilitatory effects of each iconic gesture type. The informational relationship of gesture with speech may also differ by aphasia type. In one study, people with nonfluent aphasia primarily used gesture to replace verbal communication, while those with fluent aphasia used gesture redundantly with verbal communication (Behrmann & Penn, 1984). In another study, different patterns of gesture use in a Broca's aphasia group, Wernicke's aphasia group, and neurotypical control group showed that gesture played a dominant role in enhancing communication, rather than facilitating lexical retrieval, among the speakers with aphasia (Kong et al., 2017). Note that these studies did not constrain their evaluation of gesture to iconic gesture. Furthermore, evaluation of other important metrics describing the relationship between gesture and speech, such as temporal ordering (Hadar & Butterworth, 1997; Hadar et al., 1998), will elucidate task effects on iconic gesturing.

Future work should focus on how task-specific and extra-task linguistic ability predict iconic gesture usage and how that differs according to task demands. Doing so can provide evidence about the level of breakdown in the language system in aphasia. Rose and Douglas (2001) explored this in reference to gesture's facilitatory effects in object naming, finding that individuals with phonological access, storage, or encoding difficulties showed improved naming abilities when iconic gestures were present versus those with a semantic impairment or a motor speech disorder (e.g., apraxia of speech). Complementary work evaluating a wordless cartoon retelling in a single case study of a person with conduction aphasia, who had difficulties with phonological encoding, demonstrated production of meaningful iconic gestures, suggesting an intact prelinguistic conceptual and semantic system (Cocks et al., 2011). Furthermore, it may be damaging to the very earliest part of speech production—prelinguistic conceptualization—that predicts the use, informativeness, and correctness of iconic gestures. Persons with aphasia and concomitant prelinguistic conceptual impairments (e.g., nonverbal design copying, spatial rotation mentalization) have been shown to produce gestures less frequently than those with relatively intact conceptual abilities (Hadar et al., 1998), further suggesting that it is the prelinguistic conceptualization that is critical for gesturing (especially iconic). Unfortunately, we could not stratify iconic gesture use by prelinguistic conceptual ability in the current data set, as the AphasiaBank corpus does not provide these data. There are additional future avenues that would be fruitful: As Clough and Duff (2020) note, no experimental studies in aphasia have examined how encouraging or constraining gesture affects the fluency of verbal output or, indeed, whether listener perceptions of fluency are influenced by gesture use. Creating a comprehensive picture in regard to the impact of gesturing on fluency in aphasia, by also taking into account different tasks and the use of encouraging or restraining gesture, will be a valuable addition to the current literature.

Interestingly, while gesture has a considerable role in multimodal communication and countless studies have demonstrated increased gesture use by persons living with aphasia, gesture has received comparatively little focus in the speech-language therapy literature, with mixed evidence for its use as a part of (or stand-alone) therapy (Rose et al., 2013). One such reason for the relatively limited scope of studies evaluating gesture therapy is a school of thought that co-speech gesturing in persons living with aphasia may inhibit linguistic productivity (e.g., constraint-induced language theory; Cherney et al., 2008). Nonetheless, there remains a decisive issue in the aphasia therapy literature: While evidence supports the use of behavioral therapy to improve aphasia, the mechanisms of therapy that promote change remain unclear (Brady et al., 2016). The review by Rose et al. (2013) emphasizes that gestures should be included in clinical assessments in aphasia, and a recent review by Clough and Duff (2020) further emphasizes that a potent means of better understanding the mechanisms of successful behavioral therapy is to comprehensively evaluate multimodal communication (i.e., gesture alongside written and spoken language) in persons with neurogenic communication disorders.

In summary, we need to better understand how and when spontaneous gestures are used, language system characteristics that associate with specific types of gestures (i.e.,

intact prelinguistic semantic ability and iconic gesture usage), how gestures differ across tasks and communication scenarios, and the informational and temporal relationship between gesture and speech. A lot of work has been done to clarify spontaneous gesture frequency in aphasia (de Ruiter, 2006; Dipper et al., 2011, 2015; Hogrefe et al., 2016; Kistner et al., 2019; Krauss & Hadar, 2001; Pritchard et al., 2013; Rose et al., 2017; Scharp et al., 2007; Sekine & Rose, 2013), with Clough and Duff's (2020) recent review detailing this, but an enhanced investigation of the role of gesture in spontaneous speech—especially across more naturalistic elicitation methods—is both under-researched and necessary to comprehensively understand the role of gesture in a variety of naturalistic situations. A comprehensive evaluation of gesture should ideally comprise tasks spanning the cognitive and linguistic spectra: constrained tasks (e.g., ask a person to demonstrate, with gesture, how to use a picture object or perform a pictured action), semiconstrained tasks (e.g., structured discourse tasks provided here, where there are clear, structured instructions and, in one case, imagery), and unstructured tasks (e.g., observation of gesture use in conversation with a trusted conversation partner).

Limitations

We acknowledge limitations in the current study. Limb apraxia (and other forms of apraxia, i.e., ideational, ideomotor) data are not available in AphasiaBank, and we therefore could not evaluate the extent to which apraxia influenced iconic gesture frequency or rate during each task (Borod et al., 1989). This study was not preregistered, but future studies' mining databases such as AphasiaBank should consider preregistration to improve replicability and reliability of proposed analyses. The AphasiaBank sample, while large, is not representative across aphasia (e.g., no global or isolation aphasia in this sample). The AphasiaBank sample includes only those with chronic aphasia (in some cases, many years poststroke), and therefore, the conclusions drawn in this study may not apply to those with more acute or subacute aphasia. Finally, AphasiaBank is a rich resource for studying discourse in aphasia and has been leading the way in advocating for increased rigor in spoken language discourse analysis in aphasia—however, we did have to exclude participants because of video camera angle, and in almost all scenarios (because of the experimental tasks), there was a table present directly in front of the speaker, which has been shown to greatly constrain gesture usage because of the availability of the table on which to rest the hands and make gesture coding (e.g., differentiation into unique strokes) difficult (Jacobs & Garnham, 2007). Future work in gesture and multimodal communication work should apply the same methodological rigor to the study of gesture that we do to speech, taking into consideration some of these critical components (e.g., room and video setup). A final limitation was that we chose to code referential and metaphoric gestures as a single category (referential + metaphoric), which did not allow us to evaluate the differences in referential and metaphoric gestures by task. Future work should

evaluate differences between metaphoric and referential gesture usage in persons with aphasia across task types.

Acknowledgments

References

Atit, K., Göksun, T., Ormand, C. J., Manduca, C. A., Resnick, I., Shipley, T. F., & Tikoff, B. (2013). Spatial gestures point the way: A broader understanding of the gestural referent. *Proceedings of the Annual Meeting of the Cognitive Science Society, 35,* 1786–1791.

Behrmann, M., & Penn, C. (1984). Non-verbal communication of aphasic patients. *The British Journal of Disorders of Communication, 19*(2), 155–168. https://www.researchgate.net/publication/16718129

Borod, J. C., Fitzpatrick, P. M., Helm-Estabrooks, N., & Goodglass, H. (1989). The relationship between limb apraxia and the spontaneous use of communicative gesture in aphasia. *Brain and Cognition, 10*(1), 121–131. https://doi.org/10.1016/0278-2626(89)90079-1

Bottenberg, D., & Lemme, M. L. (1991). Effect of shared and unshared listener knowledge on narratives of normal and aphasic adults. *Clinical Aphasiology, 19,* 109–116.

Braddock, B. A. (2007). *Links between language, gesture and motor skill: A longitudinal study of communication recovery in Broca's aphasia* [Doctoral dissertation, University of Missouri–Columbia].

Brady, M., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews, 6.* https://doi.org/10.1002/14651858.CD000425.pub4

Brenneise-Sarshad, R., Nicholas, L. E., & Brookshire, R. H. (1991). Effects of apparent listener knowledge and picture stimuli on aphasic and non-brain-damaged speakers' narrative discourse. *Journal of Speech and Hearing Research, 34*(1), 168–176. https://doi.org/10.1044/jshr.3401.168

Brookshire, R., & Nicholas, L. (1994). Speech sample-size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*(2), 399–407. https://doi.org/10.1044/jshr.3702.399

Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics, 47*(1), 14–27. https://doi.org/10.1016/j.pragma.2012.12.007

Carlomagno, S., & Cristilli, C. (2006). Semantic attributes of iconic gestures in fluent and non-fluent aphasic adults. *Brain and Language, 99*(1–2), 104–105. https://doi.org/10.1016/j.bandl.2006.06.061

Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research, 51*(5), 1282–1299. https://doi.org/10.1044/1092-4388(2008/07-0206)

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition, 23*(1), 43–71. https://doi.org/10.1016/0010-0277(86)90053-3

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. https://doi.org/10.1016/0093-934X(79)90060-9

Clough, S., & Duff, M. C. (2020). The role of gesture in communication and cognition: Implications for understanding and treating neurogenic communication disorders. *Frontiers in Human Neuroscience, 14,* 23. https://doi.org/10.3389/fnhum.2020.00323

Cocks, N., Dipper, L., Middleton, R., & Morgan, G. (2011). What can iconic gestures tell us about the language system? A case of conduction aphasia. *International Journal of Language & Communication Disorders, 46*(4), 423–436. https://doi.org/10.3109/13682822.2010.520813

Cocks, N., Hird, K., & Kirsner, K. (2007). The relationship between right hemisphere damage and gesture in spontaneous discourse. *Aphasiology, 21*(3–4), 299–319. https://doi.org/10.1080/02687030600911393

Dalton, S. G. H., & Richardson, J. D. (2019). A large-scale comparison of main concept production between persons with aphasia and persons without brain injury. *American Journal of Speech-Language Pathology, 28*(1S), 293–320. https://doi.org/10.1044/2018_AJSLP-17-0166

Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin, 145*(8), 765–784. https://doi.org/10.1037/bul0000202

de Beer, C., de Ruiter, J. P., Hielscher-Fastabend, M., & Hogrefe, K. (2019). The production of gesture and speech by people with aphasia: Influence of communicative constraints. *Journal of Speech, Language, and Hearing Research, 62*(12), 4417–4432. https://doi.org/10.1044/2019_JSLHR-L-19-0020

de Ruiter, J. P. (2006). Can gesticulation help aphasic people speak, or rather, communicate? *Advances in Speech Language Pathology, 8*(2), 124–127. https://doi.org/10.1080/14417040600667285

de Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis. *Topics in Cognitive Science, 4*(2), 232–248. https://doi.org/10.1111/j.1756-8765.2012.01183.x

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? *Gesture, 11*(2), 123–147. https://doi.org/10.1075/gest.11.2.02dip

Dipper, L., Pritchard, M., Morgan, G., & Cocks, N. (2015). The language–gesture connection: Evidence from aphasia. *Clinical Linguistics & Phonetics, 29*(8–10), 748–763. https://doi.org/10.3109/02699206.2015.1036462

Fergadiotis, G., & Wright, H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology, 25*(11), 1414–1430. https://doi.org/10.1080/02687038.2011.603898

Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology, 25*(10), 1261–1278. https://doi.org/10.1080/02687038.2011.606974

Feyereisen, P. (1983). Manual activity during speaking in aphasic subjects. *International Journal of Psychology, 18*(1–4), 545–556. https://doi.org/10.1080/00207598308247500

Goldblum, M. C. (1978). Les troubles des gestes d'accompagnement du langage au cours des lesions corticales unilaterales. In H. Hecaen & M. Jeannerod (Eds.), *Du contrôle moteur à l'organisation du geste* (pp. 383–395). Masson et Cie.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology, 64*(1), 257–283. https://doi.org/10.1146/annurev-psych-113011-143802

Gordon, J. K., & Clough, S. (2020). How fluent? Part B. Underlying contributors to continuous measures of fluency in aphasia.

Aphasiology, 34(5), 643–663. https://doi.org/10.1080/02687038.2020.1712586

Graham, J. A., & Heywood, S. (1975). The effects of elimination of hand gestures and of verbal codability on speech performance. *European Journal of Social Psychology, 5*(2), 189–195. https://doi.org/10.1002/ejsp.2420050204

Gullberg, M. (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning, 56*(1), 155–196. https://doi.org/10.1111/j.0023-8333.2006.00344.x

Hadar, U. (1991). Speech-related body movement in aphasia: Period analysis of upper arms and head movement. *Brain and Language, 41*(3), 339–366. https://doi.org/10.1016/0093-934X(91)90160-3

Hadar, U., & Butterworth, B. (1997). Iconic gestures, imagery, and word retrieval in speech. *Semiotica, 115*(1–2), 147–172. https://doi.org/10.1515/semi.1997.115.1-2.147

Hadar, U., Wenkert-Olenik, D., Krauss, R., & Soroker, N. (1998). Gesture and the processing of speech: Neuropsychological evidence. *Brain and Language, 62*(1), 107–126. https://doi.org/10.1006/brln.1997.1890

Hall, J. A., & Gunnery, S. D. (2013). Gender differences in nonverbal communication. In J. A. Hall & M. Knapp (Eds.), *Nonverbal communication* (pp. 639–670). De Gruyter. https://doi.org/10.1515/9783110238150.639

Hilliard, C., & Cook, S. W. (2016). Bridging gaps in common ground: Speakers design their gestures for their listeners. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 91–103. https://doi.org/10.1037/xlm0000154

Hilverman, C., Cook, S. W., & Duff, M. C. (2016). Hippocampal declarative memory supports gesture production: Evidence from amnesia. *Cortex, 85,* 25–36. https://doi.org/10.1016/j.cortex.2016.09.015

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E., & Swerts, M. (2015). Reduction in gesture during the production of repeated references. *Journal of Memory and Language, 79–80,* 1–17. https://doi.org/10.1016/j.jml.2014.10.004

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. https://doi.org/10.1016/j.neuropsychologia.2016.10.015

Holle, H., & Gunter, T. C. (2007). The role of iconic gestures in speech disambiguation: ERP evidence. *Journal of Cognitive Neuroscience, 19*(7), 1175–1192. https://doi.org/10.1162/jocn.2007.19.7.1175

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin, 137*(2), 297–315. https://doi.org/10.1037/a0022128

Hostetter, A. B., & Alibali, M. W. (2019). Gesture as simulated action: Revisiting the framework. *Psychonomic Bulletin & Review, 26,* 721–752. https://doi.org/10.3758/s13423-018-1548-0

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56*(2), 291–303. https://doi.org/10.1016/j.jml.2006.07.011

Kelly, S. D., Barr, D. J., Church, R. B., & Lynch, K. (1999). Offering a hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language, 40*(4), 577–592. https://doi.org/10.1006/jmla.1999.2634

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In A. Kendon (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207–227). De Gruyter. https://doi.org/10.1515/9783110813098.207

Kertesz, A. (2007). *Western Aphasia Battery–Revised*. The Psychological Corporation.

Kistner, J., Dipper, L. T., & Marshall, J. (2019). The use and function of gestures in word-finding difficulties in aphasia. *Aphasiology, 33*(11), 1372–1392. https://doi.org/10.1080/02687038.2018.1541343

Kita, S. (2014). Production of speech-accompanying gesture. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 451–459). Oxford University Press. https://psycnet.apa.org/record/2014-14212-027

Kita, S., & Davies, T. S. (2009). Competing conceptual representations trigger co-speech representational gestures. *Language and Cognitive Processes, 24*(5), 761–775. https://doi.org/10.1080/01690960802327971

Kita, S., & Lausberg, H. (2008). Generation of co-speech gestures based on spatial imagery from the right-hemisphere: Evidence from split-brain patients. *Cortex, 44*(2), 131–139. https://doi.org/10.1016/j.cortex.2006.04.001

Kong, A. P.-H., Law, S.-P., & Chak, G. W.-C. (2017). A comparison of coverbal gesture use in oral discourse among speakers with fluent and nonfluent aphasia. *Journal of Speech, Language, and Hearing Research, 60*(7), 2031–2046. https://doi.org/10.1044/2017_JSLHR-L-16-0093

Kong, A. P.-H., Law, S.-P., Wat, W. K.-C., & Lai, C. (2015). Coverbal gestures among speakers with aphasia: Influence of aphasia severity, linguistic and semantic skills, and hemiplegia on gesture employment in oral discourse. *Journal of Communication Disorders, 56*, 88–102. https://doi.org/10.1016/j.jcomdis.2015.06.007

Krauss, R. M. (1998). Why do we gesture when we speak? *Current Directions in Psychological Science, 7*(2), 54. https://doi.org/10.1111/1467-8721.ep13175642

Krauss, R. M., & Hadar, U. (2001). The role of speech-related arm/hand gestures in word retrieval. In R. Campbell & L. Messing (Eds.), *Gesture, speech, and sign* (pp. 93–116). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524519.003.0006

Kuhlen, A. K., Galati, A., & Brennano, S. E. (2012). Gesturing integrates top-down and bottom-up information: Joint effects of speakers' expectations and addressees' fee2dback. *Language and Cognition, 4*(1), 17–41. https://doi.org/10.1515/langcog-2012-0002

Li, E. C. D., della Volpe, A., Ritterman, S., & Williams, S. E. (1996). Variation in grammatic complexity across three types of discourse. *Journal of Speech-Language Pathology and Audiology, 20*(3), 180–186.

Luria, A. (1970). *Traumatic aphasia: Its syndromes, psychology and treatment* (Vol. 5). De Gruyter. https://doi.org/10.1515/9783110816297

MacWhinney, B. (2018). *Tools for analyzing talk, Part 2: The CLAN program (Issue 2000)*. https://doi.org/10.21415/T5G10R

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*(11), 1286–1307. https://doi.org/10.1080/02687038.2011.589893

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

Menn, L., Reilly, K. F., Hayashi, M., Kamio, A., Fujita, I., & Sasanuma, S. (1998). The interaction of preserved pragmatics and impaired syntax in Japanese and English aphasic speech. *Brain and Language, 61*(2), 183–225. https://doi.org/10.1006/BRLN.1997.1838

Morsella, E., & Krauss, R. M. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology, 117*(3), 411–424. https://doi.org/10.2307/4149008

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. https://doi.org/10.3109/17549507.2012.712157

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. https://doi.org/10.1080/02687038.2014.993912

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. https://doi.org/10.1111/j.1467-9280.1996.tb00364.x

Rose, M., & Douglas, J. (2001). The differential facilitatory effects of gesture and visualisation processes on object naming in aphasia. *Aphasiology, 15*(10–11), 977–990. https://doi.org/10.1080/02687040143000339

Rose, M., Mok, Z., & Sekine, K. (2017). Communicative effectiveness of pantomime gesture in people with aphasia. *International Journal of Language & Communication Disorders, 52*(2), 227–237. https://doi.org/10.1111/1460-6984.12268

Rose, M., Raymer, A., Lanyon, L., & Attard, M. (2013). A systematic review of gesture treatments for post-stroke aphasia. *Aphasiology, 27*(9), 1090–1127. https://doi.org/10.1080/02687038.2013.805726

Scharp, V. L., Tompkins, C. A., & Iverson, J. M. (2007). Gesture and aphasia: Helping hands? *Aphasiology, 21*(6–8), 717–725. https://doi.org/10.1080/02687030701192273

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. https://doi.org/10.1044/1058-0360

Shadden, B. B., Burnette, R. B., Eikenberry, B. R., & DiBrezzo, R. (1990). All discourse tasks are not created equal. *Clinical Aphasiology. 20*, 327–342.

Stark, B. C. (2019). A comparison of three discourse elicitation methods in aphasia and age-matched adults: Implications for language assessment and outcome. *American Journal of Speech-Language Pathology, 28*(3), 1067–1083. https://doi.org/10.1044/2019_AJSLP-18-0265

Stark, B. C., & Fukuyama, J. (2020). Leveraging big data to understand the interaction of task and language during monologic spoken discourse in speakers with and without aphasia. *Language, Cognition and Neuroscience*. Advance online publication. https://doi.org/10.1080/23273798.2020.1862258

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. https://doi.org/10.1016/0093-934X(81)90100-0

Wright, H. H., & Capilouto, G. J. (2009). Manipulating task instructions to change narrative discourse performance. *Aphasiology, 23*(10), 1295–1308. https://doi.org/10.1080/02687030902826844