J Nonverbal Behav (2015) 39:93–111
DOI 10.1007/s10919-014-0200-6

ORIGINAL PAPER

A Coding System with Independent Annotations of Gesture Forms and Functions During Verbal Communication: Development of a Database of Speech and GEsture (DoSaGE)

Anthony Pak-Hin Kong · Sam-Po Law · Connie Ching-Yin Kwan · Christy Lai · Vivian Lam

Published online: 25 September 2014

Abstract Gestures are commonly used together with spoken language in human communication. One major limitation of gesture investigations in the existing literature lies in the fact that the coding of forms and functions of gestures has not been clearly differentiated. This paper first described a recently developed Database of Speech and GEsture based on independent annotation of gesture forms and functions among 119 neurologically unimpaired right-handed native speakers of Cantonese (divided into three age and two education levels), and presented findings of an investigation examining how gesture use was related to age and linguistic performance. Consideration of these two factors, for which normative data are currently very limited or lacking in the literature, is relevant and necessary when one evaluates gesture employment among individuals with and without language impairment. Three speech tasks, including monologue of a personally important event, sequential description, and story-telling, were used for elicitation. The EUDICO Linguistic ANnotator software was used to independently annotate each participant's linguistic information of the transcript, forms of gestures used, and the function for each gesture. About one-third of the subjects did not use any co-verbal gestures. While the majority of gestures were non-content-carrying, which functioned mainly for reinforcing speech intonation or controlling speech flow, the content-carrying ones were used to enhance speech content. Furthermore, individuals who are younger or linguistically more proficient tended to use fewer gestures, suggesting that normal speakers gesture differently as a function of age and linguistic performance.

Keywords Gesture form · Gesture function · Database · Cantonese · Nonverbal communication

A. P.-H. Kong (✉)
Department of Communication Sciences and Disorders, University of Central Florida, Orlando, FL, USA
e-mail: antkong@ucf.edu

A. P.-H. Kong · S.-P. Law · C. C.-Y. Kwan · C. Lai · V. Lam
Division of Speech and Hearing Sciences, The University of Hong Kong, Hong Kong SAR, China

Introduction

Gesture refers to the arm and hand movements that synchronize with speech. They are commonly used together with spoken language in human communication (McNeill 1992). Research on gesture employment has been conducted from different perspectives and approaches, each with different aims. For example, in the field of linguistics, psychology, and social sciences, gesture research mainly focuses on the association between language and human thoughts. Researchers in neuroscience and neurolinguistics are interested in exploring gestures that accompany oral production and their neural correlates. In the discipline of communication sciences, such as studies involving speakers with aphasia or hearing impairment, researchers examine the relationships between gestures and deficits in oral language. Finally, scientists with a background in computer sciences and computer engineering have evaluated and reported various methods and algorithms for real-time gesture and posture recognition.

There is great variability and individual differences in the use of gestures, especially when one considers factors such as a speaker's age (e.g., Cohen and Borsoi 1996; Feyereisen and Havard 1999), level of language proficiency (e.g., Kong et al. 2012), presence of language impairment (e.g., Kong et al. 2010; Lanyon and Rose 2009), or cultural and language background (e.g., Wilson 1998). Co-verbal gestures, which are considered to be complex performance uniquely found in humans, contain content bearing semantic and visual information (Quek et al. 2002). Construction of an extensive database that contains rich information on gestures produced by a large number of normal speakers, with carefully controlled linguistic content using specific communication contexts and standardized language tasks, therefore, should allow one to achieve the following: (1) to properly evaluate the role of gestures in human communication, and (2) to systematically capture how gesture employment varies across individual speakers. As such, databases that are developed by using controlled linguistic targets, such as carefully selected themes or speech content, and standardized elicitation methods based on speakers who are language impaired can further provide important input to the theoretical work on the compensatory and/or facilitating functions of co-verbal gestures (Ahlsén 2011).

Existing Gesture Databases in the Literature

A number of publicly accessible gesture databases have been made available since about two decades ago. The majority of them were constructed with an intention to facilitate computerized automatic detection, recognition, and/or classification of hand gestures. They include databases of static (e.g., Just et al. 1996; Triesch and von der Malsburg 1996) and dynamic hand postures (e.g., Hwang et al. 2006; Marcel et al. 2000), which researchers can refer to as common benchmarks for designing computer algorithms to recognize hand gestures and postures (Dadgostar et al. 2005). Given that manual annotations of gestures can be very time-consuming, these databases should in principle enable automatic annotation tools of gestures based on features such as hand positions, orientation and/or movements. When linguistic information of speech is considered at the same time, the accurate and efficient alignment of gestures one can get from these databases may be crucial to researchers who investigate co-verbal gesture employment.

To allow systematic investigation of the relationship between gestures and oral speech events, a few databases that included video recordings of speech and gestures that spontaneously accompany speech were developed. For example, the 'Gesture Database' (GDB) put forth by the Max Planck Institute (2001) included video clips of narration of spontaneous speech tasks (e.g., storytelling or description of local environment and route direction) with annotated gestures in each recording. With a special interest in examining the cultural differences in co-verbal gestures, subjects in the GDB included native speakers living in the Netherlands, Italy, the USA, Japan, Turkey, Australian Aboriginal communities, Mexico, Belize, and Ghana. Note that the speech content of individual speakers varied depending on their personal experience. Some other databases consist of multimodal tagging of gestures and speech events in natural conversations and mimic of pre-recorded video clips, such as the one reported by Hayamizu et al. (1999) that contains information on the shape, trajectory, and orientation of hand motions, position of face, and speech content. According to Hayamizu et al., the required tasks of free conversation and mimics allowed them to capture linguistic and gestural content encompassing a range of variation across speakers. The 'Bielefeld Speech and Gesture Alignment corpus' (SaGA; Lücking et al. 2010) is another database that details the semantic characteristics, communication goals, and themes of content within verbal speech, in addition to the coding of morphology (e.g., hand shapes and wrist positions) and movement features of accompanied gestures. All participants were engaged in a 'spatial communication task' where they were exposed to the same virtual reality environment and had to describe what they had seen and to provide route direction to examiners. A more comprehensive multimodal database that consists of information regarding gestural (e.g., motion traces and shape of hand motions), speech (e.g., breath pauses and speech dysfluencies), and acoustic signals (e.g., fundamental frequency and speech power amplitude) was also made available by Quek et al. (2002). The speech content produced by Quek et al.'s subjects was less controlled as it involved free description of individual subjects' living quarters to an interlocutor. While these databases contain recordings of normal individuals, there are also some recent ones that are smaller in scale but focus on different populations, such as speakers impaired in language production (Ahlsén 2011) or wheelchair users who are limited in using their hands for co-speech gestures (Anastasiou 2012). The former study involved participants engaging in free face-to-face interaction with experimenters and the latter used participants' daily communication scenarios; variations, therefore, existed in the content of the collected speech samples.

The above-mentioned databases containing abundant verbal and nonverbal information have greatly facilitated conducting research on employment of gestures during spontaneous production of language. However, mixed conclusions focusing on the part of the speakers and listeners have been made regarding the communicative functions of gestures. For example, speakers who employ gestures would benefit from the facilitative effects on lexical retrieval (e.g., Goldin-Meadow 1999), and the gestures listeners see during verbal interactions have been reported to assist speech content decoding or conceptualization of the visuo-spatial aspects of a message (e.g., Wu and Coulson 2007). While both perspectives can be true and these facilitative effects can sometime co-occur for both the speakers and listeners during verbal interaction, most of these existing databases have not specified whether the coding framework and analyses focused on those who convey or receive the message. In this paper, we introduce a recently developed Database of Speech and GEsture (DoSaGE) based on listeners' perception of and independent annotation of the forms and functions of gestures employed by 119 neurologically unimpaired right-handed native speakers of Cantonese.

Previous Gesture Coding Systems

Objective quantification of gestures is not as straightforward as that of language production. A clear and comprehensive classification framework for gestures can therefore help us identify their referential values in communication (McNeill 1992) and potentially compare performance across individuals. Several gesture classification systems have been proposed over the years, with earlier work mainly classifying gestures based on their relationship to speech content. For example, Ekman and Friesen (1969) identified five types of gestures, including (1) pictorial gestures—movements that can reveal speech content by 'drawing' a picture of the target, such as using the hands to mimic the shape or size of the object, (2) spatial gestures—using hands to show spatial distances, such as both palms facing each other to indicate the distance of the platform gap in the train station, (3) rhythmic gestures—movements that stress particular phrases or movement along the flow of speech, (4) kinetic gestures—movements a speaker performs within speech content, such as arm moving to pretend the action of running, and (5) pointing gestures, or the use of finger(s) to point out specific objects. While most gestures in this system were identified as one of the five types, there were cases when multiple coding was given to one particular gesture, indicating the overlapping nature and potential ambiguity of the gesture types in the framework.

McNeill (1992) modified the coding system of Ekman and Friesen's (1969) so that one gesture could only be coded once. More specifically, all hand movements representing actions or features of objects (kinetic and pictorial gestures, respectively) were referred to as 'iconic gestures'. All movements associated with the flow or rate of speech (rhythmic gestures) were classified as 'beats'. In addition to movements providing spatial information, those that conceptualized abstract ideas were coded as 'metaphoric gestures'. As for pointing gestures, they were grouped together with other movements that indicated location and identified as 'deictic gestures'. The author also included two additional codes in the system, 'emblems' and 'pantomime'. The former was defined as any conventionalized signs, such as a thumbs-up gesture, that are culturally specific with standard forms and significances but vary from place to place, while 'pantomime' was a single gesture or a sequence of them that allowed one to convey a narrative line, with a story to tell, produced without speech. McNeill referred to the above gestures as 'gesticulations', which are within a continuum of motions that embodies a meaning relatable to the accompanying speech. More importantly, he believed that this modified coding system not only widened the definition for each type of gesture, but also provided a clearer framework to analyze gesture production during spontaneous speech.

Crowder (1996) also developed a classification framework that focused on the degree of information overlap between gestures and speech content. In particular, co-verbal gestures were dichotomized into gesture-speech pairs and non-meaningful gestures. All gesture-speech pairs were further categorized into three types: (1) redundant gestures, of which no new information is added to speech content, (2) enhancing gestures, which can expand speech content, and (3) content-carrying gestures, which present additional meaning absent in speech. Non-meaningful gestures, on the other hand, were those identified during pauses or filling of speech gaps. Mather (2005) considered these non-meaningful gestures as 'regulators', which were usually used during speech for filling gaps of speech initiation, continuation, shifting, or termination.

One major limitation of the foregoing gesture investigations lies in the fact that coding of forms and functions of gestures was not clearly differentiated, such as Ekman and Friesen (1969) and McNeill (1992). In particular, the rhythmic and pointing gestures in

Ekman and Friesen (1969) and the beats in McNeill (1992) were coded based on their function served during the co-occurring speech segment, but the remaining ones were related to the gesture forms or patterns of gestural movement. Other previous studies have illustrated that one particular form of gesture can serve one or more communicative functions. For instance, iconic gestures in normal speakers have been found to facilitate word finding process by providing an alternate route to the lexicon (Butterworth and Hadar 1989) as well as to help listeners decode speech content more accurately (Beattie and Shovelton 1999). Reported functions of deictic gestures in the literature also include lowering demand on speakers' cognitive resource and facilitating speech formulation (Goldin-Meadow 2003) and helping listeners decode messages by directly pointing to and/or illustrating objects in the real world (Ekman and Friesen 1969). These examples clearly indicate that under different conditions, a particular form of gesture may carry different functions. The mixed coding of gesture forms and functions within one quantification system is, therefore, conceptually problematic and may create confusion when it comes to interpreting gesture employment. One should also note that the terminologies used in the above-mentioned systems to refer to a particular type of gesture were also inconsistent, which could lead to difficulties when one attempts to compare and contrast them.

Xu et al. (2009) suggested that language and gesture are originated from the same system, but could function independently to compensate each other. In particular, it seems reasonable to suggest that, at least in normal speakers, language and gesture are activated in an antagonistic manner where they can compensate each other. Although the above-mentioned coding frameworks allowed one to code and quantify the employment of gestures in a more objectively manner, most of these reports did not accompany thorough investigations of the relationship between the coded gestures and speakers' characteristics, such as age and language competency. Following Xu et al.'s claim, speakers' age and linguistic competency for language tasks may affect their nonverbal performance during speech production.

Gesture Use as a Function of Age

Montepare and Tucker (1999) have argued that investigating age-related issues concerning the use of language and its associated co-verbal gestures may enable one to better understand the role of nonverbal behaviors in forming impressions and interpersonal communication. Apart from decoding meaning of a message, dynamic body movements can also provide additional cues to normal listeners to identify emotions within a verbal context (Montepare et al. 1999). However, the impact of aging on the use of gestures among normal individuals has not been extensively examined, with only a few studies reported in the literature. For example, Feyereisen and Havard (1999) compared the quantity of conversational hand gestures used by 23 young and 19 older adults (average age of 21 and 70 years, respectively) and found a lack of differences on how frequently gestures were produced between the two age groups. However, the use of representational gestures, such as iconic, metaphoric or deictic gestures as compared to beats, was significantly less frequent in older adults. These findings were consistent with Cohen and Borsoi (1996) who examined the differences of employing gestures in an object description task by 24 young and 24 elderly adults (average age of 20 and 69 years, respectively). Specifically, it was reported that there were fewer descriptive gestures (i.e., those that were closely linked to the content of the speech that may or may not carry portion of the verbal message in a nonverbal form) in the elderly group, but the two age levels were comparable in their use of non-descriptive gestures. Ska and Nespoulous (1987) specifically studied how different age groups of normal adults speakers demonstrated different patterns of using pantomimic gestures. Subjects were divided into three age groups, including 30 young, 30 middle-age, and 36 elderly speakers with an average age of 25, 45, and 71 years, respectively. In response to a verbal command of executing pantomimes, only very limited differences were found quantitatively in the performance among the three age groups. Qualitatively, on the other hand, there was a tendency for the aged subjects to use a part of their body to serve as the relevant imaginary object(s), compared with the younger subjects' whose performance was mainly constituents of self-oriented one-hand gestures. The results of these studies generally suggested a decrease of spontaneous hand gestures related to the speech content with aging.

Gesture Use as a Function of Linguistic Command

Conflicting findings about the relationship between gesture and linguistic command have been reported in the literature. Jacobs and Garnham (2007) found that gestures were seldom used for retrieving words on the part of the speakers; instead they were produced for listeners to assist understanding of speech content. On the other hand, Goldin-Meadow (1999, 2003) suggested that there are beneficial effects of gesture use on reducing a speaker's linguistic demand and/or promoting formulation of thoughts and ideas in conversing complicated ideas, such as explaining a mathematic problem, among normal speakers. In other words, when one gestures more cognitive capacity for the speech task can be spared, thus freeing up capacity for retrieving words from memory (Rauscher et al. 1996). Alibali and DiRusso (1999) reported consistent findings that typical speakers could remember more words in a word reciting task when gesturing was involved, suggesting gestures might help lower the demand of working memory in a speech task. On the other hand, Hostetter and Alibali (2007) reported that individuals with low phonemic fluency skills, as reflected by a lower ability to generate single words with a specific initial consonant, but high spatial visualization skill would gesture most often. The rate of producing representational gestures, however, did not necessarily vary as a function of the competency of phonemic fluency alone, suggesting that individual differences in using gestures were associated with individual differences in cognitive skills.

Similarly, results in Colletta et al. (2010) suggested a positive correlation between the frequency of gestures and linguistic skills in normal developing children. In particular, they revealed that the use of non-representational gestures, such as head nodding for agreement or hand gestures that indicate a speaker is searching for a word or expression, which allowed speakers to mark and cue their narrative cohesion, increased in frequency with developmental maturation (and therefore linguistic skills). In contrast, some other studies focusing on normal adult speakers found the opposite. For example, Crowder (1996) reported that while adult speakers with lower lexical diversity produced more target-enhancing or content-carrying gestures to supplement their speech, those with richer lexical diversity tended to gesture less or produce gestures that were redundant to speech content. Fluency of speech was found to positively correlate with frequency of gesture employed (Mayberry and Jaques 2000). In particular, there is a tendency of a reduced amount of gesture production in speakers who stutter more. These conflicting findings on how gesture production is related to linguistic integrity may be due to different age groups of participants (e.g., adults vs. children), and different measures used for quantifying subjects' linguistic performance.

Aims of the Present Study

The first aim of this paper is to introduce a coding system that independently codes and quantifies gesture forms and functions for the development of a database of gesture employed during spontaneous speech tasks, namely the Database of Speech and GEsture (DoSaGE), based on 119 neurologically unimpaired right-handed native speakers of Cantonese. The DoSaGE incorporates and integrates many categories of gestures used in previous systems but critically differs from other systems in that gesture forms and functions are distinguished. Speech content for each participant was carefully controlled by using standardized speech elicitation tasks and procedures following the AphasiaBank protocol (MacWhinney et al. 2011). One major advantage of this protocol lies in the fact that it has been widely adopted for studies of (non-)linguistic investigation of speakers with and without language impairment worldwide, including those with aphasia, dementia, stuttering, and traumatic brain injury, which may facilitate any subsequent cross-linguistic and/or cultural comparison studies.

The second aim was to examine how speakers' age and linguistic performance were related to the frequency of gesture employment using a large sample of normal speakers. Specifically, it was hypothesized that the number of gestures would increase with a speaker's age but decrease with language skills. Unlike most previous studies, systematic and detailed quantification of each participant's discourse output was done in the current study, with the use of sensitive linguistic measures that reflected performance from the lexical to sentence levels.

Method

Participants

One hundred and nineteen right-handed native speakers of Cantonese, with no history of neurological lesions that would affect everyday communication, were recruited in Hong Kong. They were distributed into three age groups (Young 18–39 years, Middle-aged 40–59 years, and Senior: 60 or above), and two education levels (using secondary school for the two younger and primary school for the oldest group, respectively, as cut-off of "low/high"). Their demographic information is given in Table 1.

Data Collection and Analysis

Following the AphasiaBank protocol (MacWhinney et al. 2011) but with adaptation to the local Chinese culture, each participant was asked to narrate an important event in their life, to tell two highly familiar stories (Turtle and Hare and Crywolf) after presentation of picture cards, and to describe the procedures of making a ham and egg sandwich in front of photos of the ingredients. All production was videotaped in a sound proof room and subsequently transcribed orthographically as files in the Child Language ANalyses computer program (CLAN; MacWhinney 2003). Each language sample and its corresponding digitized video were linked using the EUDICO Linguistic ANnotator (ELAN; Max Planck Institute for Psycholinguistics 2002). Three independent tiers were generated in each ELAN document to annotate the (1) linguistic information of the transcript, (2) forms of gestures appeared, and (3) function for each gesture used.

Table 1 Number of participants in various age and education groups (N = 119)

Age (in year)	Male		Female	
	Low education	High education	Low education	High education
18–39	6	12	9	12
40–59	6	9	12	12
60 or above	8	5	22	6
Total	20	26	43	30

Language Performance

For each participant's language samples, the total numbers of simple and complete sentences (and therefore the total number of sentences) were tallied. A simple sentence referred to that consisting of one independent clause with a subject and predicate (verb plus complement or modifier). A complete sentence was one with a structure that consisted of one or more clause or a phrase with specific intonation. It can be (i) a sentence with a subject plus predicate, (ii) a sentence of a predicate only and could exist alone, (iii) a sentence of a predicate only but was grammatically accurate only when it was accompanied with the previous sentence, (iv) a single-word sentence, or (v) a compound and/or complex sentence. The samples were then linguistically quantified using five indices: (1) type-token ratio—dividing number of different words by total number of words; (2) percentage of simple sentences; (3) percentage of complete sentences, (4) percentage of regulators (utterances used for initiation, continuation, shift, and termination in conversations; Mather 2005)—dividing its number by total number of sentences; and (5) percentage of dysfluency (pauses, interjections, repetitions, prolongations, or self-corrections; Mayberry and Jaques 2000)—dividing frequency of dysfluency by total number of sentences.

Form and Function of Gesture

In the current study, 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 its resting position, gestures would be divided by either a pause in the movement or an obvious change in shape or trajectory (Jacobs and Garnham 2007). During the flow of speech, self-adapting motions, such as touching the face or changing hand position from the lap to the desk, were excluded from the analysis as they lack semantic attachment (Jacobs and Garnham 2007). Six forms of gestures, modified based on the classification by Ekman and Friesen (1969), Mather (2005), and McNeill's (1992), were proposed:

1. Iconic gesture: refers to gestures that model the shape of an object or the motion of an action. For example, a speaker arcs his/her fist to form a cup and drink from it when saying 'I drink from a cup'.
2. Metaphoric gesture: refers to gestures that show pictorial content to communicate an abstract idea. For example, when a speaker says 'I want to show you something', he/she raises a hand up to form a circle-like sign and offers the listener that sign, representing the concept of 'something'.

3. Deictic gesture: refers to familiar pointing, indicating objects in conversational space. For example, when a speaker says 'I walked up the stairs', he/she points upward.
4. Emblem: gestures with standard properties and language-like features. For example, the OK sign has a culturally agreed upon meaning; and it is necessary to place the thumb and index finger together in order to form the sign.
5. Beat: It refers to rhythmic beating of a finger, hand or arm. It can be a simple hand or arm flick or a moving motion of finger(s), hand(s), or arm(s) in an up-and-down or a back-and-forth fashion.
6. Non-identifiable: refers to an uncodable finger, hand, and/or arm movement due to its ambiguous connection or lack of a direct meaning to the speech content. A gesture that was partly visually obstructed but cannot be categorized as any of the above five forms is also coded as non-identifiable.

In the dimension of functions, gestures were classified by their primary role in communication. Acknowledging that sometimes a specific gesture could serve one or more communicative purposes, only the primary function in relation to the speech content was annotated. Eight functions of gestures adopted from several previous studies were proposed:

1. Providing additional information to message conveyed: a gesture gives information in addition to the speech content (Goldin-Meadow 2003), i.e., the content of the gesture gave additional information related to the speech. For example, when a speaker says the word 'open', he/she gives a twisting motion in the air, representing the action to open a jar. This gesture gives additional information to the speech content (about the opening of a container, but not other objects).
2. Enhancing the speech content: a gesture gives the same meaning to the speech content (Beattie and Shovelton 2000). This might help listener to decode speech content. For instance, a speaker uses his finger to show the action of switching on the light when saying 'I switched on the light when I entered the room'.
3. Providing alternative means of communication: a gesture carries meaning that is not included in speech content (Le May et al. 1988). Speech could be superseded by the gesture employed. For example, a speaker produces the OK sign solely to answer the question of 'Are you ready to start?'
4. Guiding and controlling the flow of speech: a gesture that reinforces the rhythm of the speech (Jacobs and Garnham 2007). The rate of gesture movement is synchronized with the pace of speech.
5. Reinforcing the intonation or prosody of speech: usually coded when a speaker uses a gesture to emphasize his/her meaning. For example, a speaker gives a strong flick at the target word.
6. Assisting lexical retrieval: a gesture that is intended to facilitate lexical access (Krauss and Hadar 1999). This is coded when a speaker attempts to produce a target word after gesturing at times of word-finding difficulty, such as an observable time delay when trying to retrieve the target word(s), stating 'I know that word, but I can't think of it', interjections, circumlocution, word or phrase repetitions, and word substitutions within a speech (Mayberry and Jaques 2000). For example, this type of gestures may occur when a speaker demonstrates a noticeable pause or prolongation of speech sound in the middle of a sentence where a target word is finally retrieved or when a speaker reformulate or substitute a phrase or sentence during speech (German 2001).
7. Assisting sentence re-construction: a speaker shows modification of syntactic structure after using a gesture (Alibali et al. 2000). Alternatively, this gesture is produced when

there is difficulty in sentence construction and refinement of sentence structure is noticed after the use of this gesture.

8. No specific function deduced: a gesture that does not show a specific function in relation to the speech content or an unclassifiable function other than the seven mentioned above.

Note that the most important criteria in differentiating the first three functions was to determine if the presence of a gesture, irrespective to its form, could provide nonverbal information that was additional to (functions 1 and 3) or same as (function 2) the associated speech content. Functions 2 and 4 differed from function 5 in that the former ones did not involve a speaker's intensifying or accentuating a target element in the speech. If needed, coders could also use prosodic patterns that were specific and consistent to the speech content to make the distinction. As for functions 6 and 7, a particular annotated gesture should accompany a delay or struggle of word finding and sentence formulation, respectively, as described above. To ensure consistency across the coding, only one rater (author CK) was involved in generating the codes. Frequency count of each gesture form and function was obtained for each participant using the 'Annotation Statistics' option in ELAN.

Reliability of the Gesture Coding System

Data from 12 subjects (10 %) were selected randomly and re-analyzed in terms of the forms and functions of gestures by author CK and another trained independent rater to obtain intra-rater and inter-rater reliabilities, respectively. Kendall's tau correlation was employed to calculate if there were significant differences between the ratings. In addition, the same data set was used to obtain the point-to-point percentage agreement, both within (author AK) and across raters (author AK and CK), on (1) the total number of gestures coded, (2) coding each gestures form, and (3) coding the function of each identified gesture. The results of Kendall's tau coefficients for inter-rater and intra-rater reliabilities showed that all coefficients were significant at $p < .05$ or better, with the coefficients for intra-rater agreement higher than those of inter-rater agreement. Concerning the across-rater point-to-point percentage agreement, only two out of a total of 186 gestures showed disagreement (1.08 %) where one more and one less gesture was identified in two different subjects. An agreement of 97.85 % (182/186) and 96.24 % (179/186) was obtained for the forms and function of gestures coded, respectively. The within-rater point-to-point percentage agreement was found to be better, with a 100 % agreement on coding the total number of gestures. The agreement for coding the gesture forms and functions was both 98.92 % (184/186). Major within- and across-rater disagreements occurred in annotating the form of 'non-identifiable gestures' (where they were coded as deictic or metaphoric gestures) and the function of 'non-specific' (where they were coded as enhancing the speech content).

Statistical Analysis

Prior to running any statistical analysis, normality and homogeneity of variance of the normative data were tested using Kolmogorov–Smirnov test and Levene test, respectively (Field 2009). For data that were not normally-distributed due to zero in gesture counts, Box-Cox transformations (Box and Cox 1964) were employed to normalize negative skewness (Osborne 2010). Natural log was done when positive skewness of the data was observed (Field 2009). Non-parametric statistical analyses were implemented if the assumption of normal distribution was still violated after data transformation.

To investigate the relationship between linguistic performance and frequency of gesture used, the 119 subjects were ranked with reference to the total number of gestures they used. The top 40 subjects (i.e., 33.3 %) who used the highest number of gestures were regarded as the "high gesture group"; while the bottom 40 subjects were regarded as the "low gesture group". Comparisons between the high and low gesture groups were made among the five language indices mentioned above using independent *t* tests. Bonferroni adjustment of alpha value of .01 (.05/5) was applied to control for the occurrence of Type I errors.

Non-parametric Kruskal–Wallis test was used to determine the effect of age. Post-hoc tests of Mann–Whitney test were followed to determine any significant differences between the three age groups. Bonferroni adjustment of alpha value was applied to control for Type I errors. In addition, Kendall's tau coefficients were used to examine how each gesture form was related to the linguistic measures and to the age of the participants.

Results

Normative Data in DoSaGE

A total of 3,061 gestures were annotated from 82 participants throughout all tasks (see Table 2). The remaining 37 speakers (31.1 %) did not produce any gestures. Concerning the forms of these gestures, the majority was non-content-carrying (i.e., non-identifiable: 84.1 % and beats: 3.4 %). Note that only very few non-identifiable gestures (about 0.4 %) were partly visually obstructed. The distribution of content-carrying gestures was as follow: deictic (5.8 %), iconic (3.3 %), metaphoric (2.3 %), and emblems (1.1 %). As for their functions, it was found that 83.7 % of the gestures did not serve a specific function. This was followed by enhancing the speech content and reinforcing intonation or speech prosody.

Regarding the relationship between the forms and functions of these coded gesture, the results suggested that content-carrying gestures mainly functioned as enhancing the speech content (82.9–92.7 %). Beats, on the other hand, served to reinforce the intonation or prosody of speech during incidents of emphasizing speech content (78.6 %) or to guide and control the speech flow (21.4 %). For non-identifiable gestures, more than 99 % of them had no specific functions.

The average length of the total transcript was 792.3 words or 89.05 sentences per participant. To understand the effect of linguistic performance in relation to gestural production, five linguistic parameters were compared between speakers in the groups of high and low gesture use. The descriptive statistics on linguistic performance of speakers in high and low gesture group is shown in Table 3. Results of independent *t* tests indicated that speakers who gestured more had a lower type-token ratio ($p = .01$) and lower percentage of regulators ($p = .001$) than those in the low gesture group. No significant differences between the two gesture groups were found for the other three parameters (percentage of complete sentences, $p = .63$, simple sentences, $p = .24$, and dysfluency, $p = .61$).

Given the high proportion of non-specific gestures in our normative data set, the examination of how gesture use differed as a function of age was conducted by separating the gestures into (a) all but non-specific functions and (b) solely non-specific gestures. The

Table 2 Distribution of various forms and functions of gestures employed (N = 119)

	Content-carrying gestures				Non-content-carrying gestures		% of functions
	Iconic	Metaphoric	Deictic	Emblem	Beat	Non-identifiable	
% of forms	3.3 (102)[a]	2.3 (70)	5.8 (178)	1.1 (34)	3.4 (103)	84.1 (2,474)	
Functions (%)							
Provide[b]	10.8	1.4	4.5	5.9	0.0	0.0	0.7 (22)[a]
Enhance[c]	83.3	82.9	92.7	88.2	0.0	0.0	11.0 (338)
Alternate[d]	1.0	1.4	0.0	2.9	0.0	0.0	0.1 (3)
Guide/ Control[e]	0.0	0.0	0.0	0.0	21.4	0.0	0.7 (22)
Reinforce[f]	0.0	5.7	0.0	0.0	78.6	0.0	2.8 (85)
Lexical[g]	4.9	4.3	2.2	2.9	0.0	0.5	0.9 (26)
Sentence[h]	0.0	0.0	0.0	0.0	0.0	0.2	0.1 (4)
No specific[i]	0.0	4.3	0.6	0.0	0.0	99.3	83.7 (2,561)

Total number of gestures = 3,061

[a] Frequency (total number of gesture) is given in parenthesis

[b] Providing substantive information

[c] Enhancing speech content

[d] Providing alternative means of communication

[e] Guiding and controlling flow of speech

[f] Reinforcing intonation or prosody of speech

[g] Assisting lexical retrieval

[h] Assisting sentence re-construction

[i] No specific function

Table 3 Linguistic performance of the high and low gesture groups

Linguistic parameters	High gesture group ($n = 40$)		Low gesture group ($n = 40$)		t test
	Mean	SD	Mean	SD	
Type-token ratio (TTR)	0.32	0.06	0.35	0.08	*
Percentage of simple sentence	0.93	0.05	0.91	0.04	
Percentage of complete sentence	0.90	0.05	0.90	0.08	
Percentage of regulators	0.02	0.02	0.04	0.05	**
Percentage of dysfluency	0.51	0.29	0.46	0.25	

* $p = .01$; ** $p = .001$

results of the Kruskal–Wallis test, with a Bonferroni's adjustment of alpha level to .025, revealed a main effect of age on the production of all but non-specific gesture forms [$H(2) = 14.976, p = .001$)]. Mann–Whitney tests, with a Bonferroni's adjustment of alpha level to .0167, were then employed to examine the effect in each group. For all but non-specific gesture forms, significant differences were only found between the senior and young group ($U = 466.5$, $p = .000$) as well as senior and middle-aged group ($U = 545.5$, $p = .008$). No significant differences were found for the solely non-specific gestures.

Table 4 displays the correlations between individual gesture forms and participants' age as well as performance on the speech tasks. Three of the four content carrying gestures correlated significantly with age. The non-content carrying gestures did not seem to relate to speakers' age. Out of the five linguistic parameters, the results indicated that while type-token ratio (TTR) had significant correlations with both content and non-content carrying gestures, percentage of simple sentence correlated better with the content carrying gestures. The other three measures did not have significant relationship with the gesture forms annotated.

Discussion

We have presented a gesture coding system with independent annotation of gestures with reference to their forms and functions during spontaneous speech production. A Database of Speech and GEsture (DoSaGE) that included speech samples and video recordings elicited from 119 typical Cantonese speakers in Hong Kong using the AphasiaBank protocol was also developed. Unlike other coding systems reported in previous literature (e.g., Ekman and Friesen 1969; McNeill 1992), this framework can be considered more refined for gestural analyses given its independent coding of gesture forms and functions. Our results also showed that all forms of gestures exhibited at least two functions. For example, while beats were found to be associated with stressing a word (McNeill 1992), we observed that they were also used by speakers to emphasize speech content (78.6 %) or to regulate flow of conversation (21.4 %).

We believe that it is a necessary step to examine gestural production among normal speakers as a reference when one investigates gestural employment in speakers with communicative disorders. Specifically, information regarding the pattern and mechanism of gesture use among typical unimpaired speakers can serve as important baseline when one attempts to measure how an individual with language impairment may deviate from his or her normal counterpart. However, researchers have seldom investigated, collected, and reported data for a large number of normal speakers. Based on the newly proposed framework, normative data of gesture use by 119 individuals in the Cantonese-speaking community in Hong Kong was obtained from standardized data collection procedures and speech elicitation tasks. Two interesting findings that have yet been reported regarding normal speakers' gestural employment are revealed. First, about one-third of subjects in our pool of speakers did not gesture throughout the three narrative tasks. While this observation may well reflect individual differences in gestural employment, it also has important clinical implications when one compares nonverbal behaviors among speakers of normal and clinical groups. Second, out of the 3,061 gestures produced by our speakers, the proportion of non-content-carrying gestures was drastically higher than its content-carrying counterparts (87.5 vs. 12.5 %). One possible reason was the lack of a need to use gestures, particularly content-carrying ones, in language tasks for our participants. According to Goldin-Meadow (1999), two major functions of gestures on the part of speakers were compensating speech content for listeners and assisting lexical retrieval for speech output. However, with intact language abilities among our typical speakers, a need to employ gestures for these purposes was not obvious (Jacobs and Garnham 2007). Krauss et al. (2000) proposed that lexical and gesture production systems interacted with each other during speech, i.e., activation of gestural information associated with a target is part

Table 4 Correlations between gesture forms and age as well as linguistic parameters

	Kendall's tau						
	Content-carrying gestures				Non-content-carrying gestures		Total
	Iconic	Metaphoric	Deictic	Emblem	Beat	Non-identifiable	
Age	0.23**	ns	0.29***	0.21*	ns	ns	0.19**
Type-token ratio (TTR)	−0.16*	ns	−0.20**	−0.16*	−0.15*	−0.18**	−0.16*
Percentage of simple sentence	0.18*	0.15*	0.22**	ns	ns	ns	ns
Percentage of complete sentence	ns	ns	−0.17*	ns	ns	ns	ns
Percentage of regulators	ns	ns	ns	ns	ns	ns	ns
Percentage of dysfluency	ns	ns	ns	ns	ns	ns	ns

* $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$; *ns* non-significant

of the word retrieval process in normal speakers. Gesture production would therefore be superfluous when the feature carried by the gesture was realized in the speech content.

Besides establishing normative data on gesture use, factors affecting gestural production in normal Cantonese speakers were investigated. First, linguistic and gestural performances were found to be related in the current study. Speakers with a higher linguistic proficiency (as reflected by a higher type-token ratio in discourse) tended to use fewer gestures, and vice versa. This observation of trade-off between the use of gestures and language was consistent with the conclusion in Xu et al. (2009). Moreover, it is worth mentioning that the use of regulators in our participants was negatively related to the frequency of gesture use. Based on the comparison of speakers in high and low gesture groups, a lower percentage of regulator usage was found in speakers who gestured more. This is consistent with the findings of Smith (1987) in which gestures could serve as silent regulators for conversation initiation, continuation, shift, and termination and could substitute the use of verbal regulators. This observation could also be accounted for by the relationship between gesture use and linguistic loading. As gestures could help reduce cognitive burden and free up effort for thought formation (Alibali and DiRusso 1999), gesture might serve the function of regulators in speech when speakers required more effort to formulate an idea.

The age of our speakers was found to influence the frequency of gesture use, in which older speakers demonstrated an increased frequency of gestures. Moreover, as compared to non-content carrying gestures, gestures that are content-carrying seemed to be related more to speakers' age. This age effect on gesture use has extended the findings in Colletta et al. (2010), who suggested a constant increase in frequency of gestural use with age in adults and a tendency of using more co-speech gestures due to semiotic resources that developed with aging. Furthermore, it has been reported that aging could lead to lengthening of reaction time (Gooch et al. 2009) and lower efficiency in allocating resources in memory tasks (Old and Naveh-Benjamin 2008). We propose that the cognitive demand for the same language task might be higher for speakers in the senior group than the other two younger groups. As a result, older speakers showed a tendency to gesture more spontaneously so as to alleviate their linguistic demand and free up more cognitive capacity and effort for speech tasks. Also worth highlighting is the contrast between our findings and those of

Feyereisen and Havard (1999) and Cohen and Borsoi (1996); these studies reported similar number of gestures, irrespective of their types, for young and older speakers. We propose that the exclusion of gestures with a non-specific function from our current analysis might have in part contributed to the discrepancy.

Our current findings suggested that most content-carrying gestures functioned to enhance speech content. This seems to be consistent with Goldin-Meadow (2003), who proposed the facilitative functions of content-carrying gestures for listeners' comprehension of speech content, through the extraction of meanings from gestures by normal communicative partners. As a post hoc analysis, we also observed that our speakers used emblems, iconic, and metaphoric gestures more frequently when they were referring to action words (64.1 %), as compared to objects (21.3 %) or adjectives (14.6 %). Deictic gestures, on the other hand, tend to be associated more with speakers' production of nouns (85.4 %), as compared to verbs (2.2 %) or other parts-of-speech (12.4 %). This unexpected observation may provide preliminary information and new insights into how the use of various gesture forms is associated to production of different word classes, but more systematic investigations should be carried out in the future.

For non-content-carrying gestures, our conclusions about beats, which mainly emphasized speech content and regulated the flow of speech, were also compatible with previous findings reported by Goldin-Meadow (2003). One should note that almost all non-identifiable gestures coded had no specific functions (99.3 %). Albeit the lack of significant relationship between percentage of dysfluency and frequency of gesture use, employment of non-specific gestures were found to always accompany incidents of dysfluency, such as pauses, interjections, repetitions, prolongations, or self-correction. This corroborates with Mather (2005) who reported 'regulators' as non-meaningful gestures that did not serve communication purposes and occurred when there were gaps within a speech. In the current study, an additional post hoc test revealed that our speakers' production was significantly related to incidents of dysfluency, $\tau = .213$, $p < .01$. This condition further supported the hypothesis that speech and gesture compensated each other (Butterworth and Hadar 1989). In our case, a high proportion of the non-specific gestures appeared when speech production was disrupted. With reference to the current results, one may question the dominance of gestures that did not serve a specific communicative function during a speech task. We argue that the current coding system did not fail to capture the majority of gestures produced by our speakers. Instead, the results lay an interesting and important foundation to further extend our project on other populations. Specifically, because of the potentially higher demand of non-oral means for communication in other speaker groups, such as those with difficulty in spontaneous oral production or children during the period of language acquisition, one may also wonder if the same high proportion of non-specific gesture will be found. Therefore, how the presence or absence of this type of gestures is related to human communication warrants further investigation. Note also that gesture-speech mismatch, a communicative act in which information conveyed in a gesture is different from that conveyed in the accompanying speech (Goldin-Meadow 1999), was not found in the current dataset. It would be interesting to further examine if this impaired communicative act is prevalent among individuals with language difficulties.

Traditional analysis of gesture production was mainly done by paper-based manual coding that was time-consuming (Ekman and Friesen 1969; McNeill 1992). With the advancement of digital media technology and linguistic engineering, ELAN (Max Planck Institute for Psycholinguistics 2002) has become available for public use to support annotations between linguistic and nonverbal behaviors in a speech context. In the past decade, ELAN has mainly been applied to the coding of conversational gesture employment among normal speakers (e.g., Bressem 2008; Brugman and Russel 2004; Lausberg and Sloetjes 2008, 2009). Similar to Lausberg and Sloetjes (2008, 2009), we also found that the use of ELAN in the current study allowed efficient coding as well as clear display of co-verbal gestures. In particular, annotations of a sentence or word in a language transcript can be time-aligned with the description of any nonverbal features observed in the corresponding video media. This multimodal annotation of one's gestural and linguistic performance within a carefully controlled language tasks enabled us, as well as future investigators who adopt this analytic approach, to effectively visualize all speech and non-speech events during specific moments of communication and/or breakdown. Therefore, systematic examination of the relationships between oral and nonverbal production can be done. Based on the experience of author CK, the average time needed to process data of each subject (four sets of language samples) in ELAN was approximately 45 min. As one becomes more experienced with the software, annotating gestures and performing qualitative and quantitative analyses of also became more efficient.

According to Wu and Coulson (2007), judgment of the function of a gesture often involved some degree of subjectivity. While coding of the six gesture forms was straightforward given the distinctive feature of individual forms, we believe that the level of subjectivity in annotating gesture functions was greatly reduced in the current study with the much more clearly defined coding criteria. The results of the inter-rater and intra-rater reliability measures further suggested that our gesture coding system was reliable. Note that among the different forms of gestures, metaphoric gestures seemed to have the lowest inter-rater agreement ($\tau = .71$). Two non-identifiable gestures and two iconic gestures were coded as metaphoric, possibly due to the different interpretation about the idea of 'abstract' for metaphoric gestures. Concerning gesture functions, emphasizing speech content had the lowest inter-rater correlation coefficient. About 6 % (8 out of 131) of the gestures originally rated as 'reinforcing speech intonation or prosody' was coded as 'no specific function' by another rater. It is believed that coding accuracy will improve with training and more detailed examples for each gesture form and function.

There are several directions to extend the current work. First, for speakers with aphasia who have long lasting disturbance of language production and/or comprehension that can manifest at the single word, sentence, and conversation levels, it has been found that non-verbal means of communicative behaviors are commonly used to act as an alternative to or a facilitator of verbal communication (Herrmann et al. 1988). The proposed coding framework should be applied to a wide range of speakers with aphasia encompassing different severity of language deficits. This will not only allow a more systematic investigation of the relationship between language competency and gesture employment, but also an examination of how nonverbal behaviors can potentially assist speakers with verbal deficits to communicate. Second, one should note that most previous studies analyzed gestural performance in a single speech task (Jacobs and Garnham 2007; McNeill 1992). The method of soliciting speech samples could potentially influence gestural production (McNeill 1992), i.e., the amount and types of gestures found in everyday conversation might be different from a picture description or story-telling task (Ekman and Friesen 1969). A more interactive language tasks can, therefore, be used to study the communicative functions of gestures. Given that the presence of conversation partners may affect the production of gestures (Jacobs and Garnham 2007), future studies may also consider this aspect of communication by investigating the presence of a conversational partner and how different conversation partners, such as strangers versus family members or friends, may change one's gestural behaviors. Finally, with this multi-dimensional approach of independent consideration of gesture forms and functions, one can systematically study the development of nonverbal communication in typically developing children. This new framework can also be tested on its suitability to study gesture use among children with developmental language disorders.

Acknowledgments This study is supported by the grant "Towards a multi-modal and multi-level analysis of Chinese aphasic discourse" (1-R01-DC010398) funded by the National Institutes of Health to Kong, A.P.H. (PI), Law, S.P. (Co-I), and Lee, A. (Co-I). The authors would like to thank all subjects for their help and co-operation during data collection.

References

Ahlsén, E. (2011). Towards an integrated view of gestures related to speech. In *Proceedings of the 3rd Nordic symposium on multimodal communication*, Vol. 15, pp. 72–77.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development, 14*(1), 37–56.

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

Anastasiou, D. (2012). A speech and gesture spatial corpus in assisted living. In *Proceedings of the eight international conference on language resources and evaluation (LREC'12)*, Vol. 8, pp. 2351–2354.

Beattie, G., & Shovelton, H. (1999). Mapping the range of information contained in the iconic hand gestures that accompany spontaneous speech. *Journal of Language and Social Psychology, 18*(4), 438–462.

Beattie, G., & Shovelton, H. (2000). Iconic hand gestures and the predictability of words in context in spontaneous speech. *British Journal of Psychology, 91*(4), 473.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformation. *Journal of Royal Statistical Society (Series B), 26*, 211–246.

Bressem, J. (2008). *Notating gestures*: *Proposal for a form based notation system of coverbal gestures.* Unpublished manuscript.

Brugman, H., & Russel, A. (2004). *Annotating multi-media/multi-modal resources with ELAN*. LREC.

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

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

Colletta, J. M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. *Speech Communication, 52*(6), 565–576.

Crowder, E. M. (1996). Gestures at work in sense-making science talk. *The Journal of the Learning Sciences, 5*(3), 173–208.

Dadgostar, F., Barczak, A. L. C., & Sarrafzadeh, A. (2005). A color hand gesture database for evaluating and improving algorithms on hand gesture and posture recognition. *Research Letters in the Information and Mathematical Sciences, 7*, 127–134.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica, 1*(1), 49–98.

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

Field, A. P. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)* (3rd ed.). London: SAGE.

German, D. J. (2001). *It's on the tip of my tongue: Word-finding strategies to remember names and words you often forget*. Chicago, IL: Word Finding Materials Inc.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in Cognitive Sciences, 3*(11), 419–429.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Belknap Press of Harvard University Press.

Gooch, C. M., Stern, Y., & Rakitin, B. C. (2009). Evidence for age-related changes to temporal attention and memory from the choice time production task. *Aging, Neuropsychology & Cognition, 16*(3), 285–310.

Hayamizu, S., Nagaya, S., Watanuki, K., Nakazawa, M., Nobe, S., & Yoshimura, T. (1999). A multimodal database of gestures and speech. In *Proceeding of the sixth European conference on speech communication and technology*, Vol. 6, pp, 2247–2250.

Herrmann, M., Reichle, T., Lucius-Hoene, G., Wallesch, C. W., & Johannsen-Horbach, H. (1988). Nonverbal communication as a compensation strategy for severely nonfluent aphasic? A quantitative approach. *Brain and Language, 13*, 41–54.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture, 7*, 73–95.

Hwang, B. W., Kim, S. M., & Lee, S. W. (2006). A full-body gesture database for automatic gesture recognition. In *Proceeding of the 7th international conference on automatic face and gesture recognition*, Vol. 7, pp. 243–248.

Jacobs, N., & Garnham, A. (2007). The role of conversational hand gestures in a narrative task. *Journal of Memory and Language, 56*(2), 291–303.

Just, A., Rodriguez, Y., & Marcel, S. (1996). Hand posture classification and recognition using the modified census transform. In *Proceeding of the IEEE international conference on automatic face and gesture recognition*, Vol. 2, pp. 351–356.

Kong, A. P. H., Law, S. P., Kwan, C., Lai, C., Lam, V., & Lee, A. (2012, November). *A comprehensive framework to analyze co-verbal gestures during discourse production*. Poster presented at the 2012 American speech-language-hearing association (ASHA) convention, Atlanta, GA, USA.

Kong, A. P. H., Law, S. P., & Lee, A. S. Y. (2010). An investigation of use of non-verbal behaviors among individuals with aphasia in Hong Kong: Preliminary data. *Procedia Social and Behavioral Sciences, 6*, 57–58.

Krauss, R. M., Chen, Y., & Gottesman, R. F. (2000). Lexical gestures and lexical access: A process model. In D. McNeill (Ed.), *Language and gesture* (pp. 261–283). Cambridge: Cambridge University Press.

Krauss, R. M., & Hadar, U. (1999). 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: Oxford University Press.

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

Lausberg, H., & Sloetjes, H. (2008). Gesture coding with the NGCS–ELAN system. In. A. J. Spink, M. R. Ballintijn, N. D. Bogers, F. Grieco, L. W. S. Loijens, L. P. J. J. Noldus, G. Smit, & P. H. Zimmerman (Eds.), *Proceedings of measuring behavior 2008* (pp. 176–177). Netherlands: Noldus Information Technology.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, 41*(3), 841–849.

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

Lücking, A., Bergmann, K., Hahn, F., Kopp, S., & Rieser, H. (2010). The Bielefeld speech and gesture alignment corpus (SaGA). In M. Kipp, J. P. Martin, P. Paggio, & D. Heylen, (Eds). *LREC 2010 workshop: Multimodal corpora–advances in capturing, coding and analyzing multimodality* (pp. 92–98). Republic of Malta.

MacWhinney, B. (2003). *Child language analyses (CLAN) (version 23 September 2003) [Computer software]*. Pittsburgh, PA: Author.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*, 1286–1307.

Marcel, S., Bernier, O., Viallet, J-E., & Collobert, D. (2000). Hand gesture recognition using input/ouput hidden markov models. In *Proceedings of the 4th international conference on automatic face and gesture recognition*, Vol. 4, pp. 398–402.

Mather, S. M. (2005). Ethnographic research on the use of visually based regulators for teachers and interpreters. In M. Metzger & E. Fleetwood (Eds.), *Attitudes, innuendo, and regulators* (pp. 136–161). Washington, DC: Gallaudet University Press.

Max Planck Institute. (2001, February 27). *Gesture database (GDB)*. Retrieved from http://www.mpi.nl/ISLE/overview/Overview_GDB.html/.

Max Planck Institute for Psycholinguistics. (2002). http://www.lat-mpi.eu/tools/elan/.

Mayberry, R. I., & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In D. McNeill (Ed.), *Language and gesture* (pp. 199–214). New York: Cambridge University Press.

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

Montepare, J., Koff, E., Zaitchik, D., & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior, 23*, 133–152.

Montepare, J. M., & Tucker, J. S. (1999). Aging and nonverbal behavior: Current perspectives and future directions. *Journal of Nonverbal Behavior, 23*, 105–109.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential effects of age on item and associative measures of memory: A meta-analysis. *Psychology and Aging, 23*(1), 104–118.

Osborne, J. W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation, 15*(12), 1–9.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X. F., Kirbas, C., et al. (2002). Multimodal human discourse: Gesture and speech. *ACM Transactions on Computer-Human Interaction, 9*(3), 171–193.

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.

Ska, B., & Nespoulous, J. (1987). Pantomimes and aging. *Journal of Clinical and Experimental Neuropsychology, 9*, 754–766.

Smith, L. (1987). Nonverbal competency in aphasic stroke patients' conversation. *Aphasiology, 1*(2), 127–139.

Triesch, J., & von der Malsburg, C. (1996). Robust classification of hand postures against complex backgrounds. In *Proceedings of the second international conference on automatic face and gesture recognition*, Vol. 2, pp. 170–175.

Wilson, F. R. (1998). *The hand: How its use shapes the brain, language, and human culture*. New York: Pantheon Books.

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

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