Gesture Categorisation and Understanding Speaker Attention to Gesture

Lauren Gawne, Barbara F. Kelly and Annie Unger
University of Melbourne
b.kelly@unimelb.edu.au

Abstract. The field of gesture classification has been an area of intense scholarship in recent decades. This article provides a brief overview of the area in seeking to understand how this theoretical framework relates to the way speakers attend to gestural information. 48 native English speakers participated in a web-based survey centred on a short narrative. The gestures focused on in the film narrative were based around McNeill's common gesture typology. Half of the participants watched the video with sound and the other half without to help ascertain whether the presence of speech affects how people attend to gestural information. Participants were asked to count the total number of gestures and list what they thought the five "best" examples of a gesture were. While there was no significant difference between the number of gestures counted by each group, the categories of gesture which were attended to varied between the two groups. Those with sound were more likely to include iconic gestures while those without were more likely to attend to beat gestures. This indicates that the presence or absence of sound has no affect on how many gestures participants observe, but it does affect what gestural information they pay more attention to.

Keywords: gesture, non-verbal, interaction, movement classification

Selected Papers from the 2009 Conference of the Australian Linguistic Society,
edited by Yvonne Treis & Rik De Busser. http://www.als.asn.au. 2010.

1. Introduction

The study of gesture as a component of linguistic communication has exploded over the past few decades. Much work has examined the form, context and function of gesture in relation to speech. Researchers have investigated bodily actions encompassing similar movements and functions and from this a basic typology of gestures has evolved within the discipline. While the research community's understanding of gestures has grown to be expansive and complex very little work has examined the intuitions of language users regarding their perception of gestural practice. This paper examines participant perception of gestures in relation to the variable of speech sound, and uses this as a basis for an initial assessment of the current categorisation typology used in gesture studies.

2. Background

In the study of gesture production and perception, research has largely focussed upon movements made by the hands and the arms (see McNeill 1992, 2005; Goldin-Meadow 2003; and Kendon 2004). While the hands and arms are frequently used to gesture, other body parts are also employed communicatively, for example the head may be used to nod or a leg and foot may be used to mark emphasis. Manual gestures appear to be salient for researchers, however it is not known what range of body movements are oriented to communication and are salient for interlocutors in interaction.

A great deal of gesture study has been focused on analysts' attempts to categorise various bodily phenomena and explain their functions (Kendon 2004). McNeill (1992, 2000, 2005) proposed a continuum of bodily movement, which he named "Kendon's Continuum" after Adam Kendon, which allows analysts to categorise body movements on the basis of their communicative functions, as can be seen in Figure 1 below.

– Gesticulation > > Pantomime > > Emblem > > Sign Language –

Figure 1. Kendon's continuum (adapted from McNeill 1992)

As we move from left to right along the continuum two main changes occur. Firstly, the obligatory presence of speech decreases, and secondly, the stability of the meaning of the gesture increases (McNeill 1992). At one end of the continuum are speech-accompanied gesticulations. These require the presence of speech and tend to be highly idiosyncratic. At the other end of the continuum is sign language, e.g. Auslan. At this end we can see that the movement is highly codified since there is an absence of speech. In the middle of the continuum we find pantomimes and emblems. Pantomimes are a crude combination of gestures and simplified speech that have limited communicative abilities and may often be idiosyncratic. Emblems are those culturally established gestures with defined meanings, such as the thumbs-up sign, the peace sign or nodding to signal 'yes'. These gestures can be understood in absence of speech and have standards of well-formedness not found in gesticulations. While gesture categorisation has been useful for understanding when to expect certain types of gesture phenomena, there has been no research that investigates orientation to gesture types, for example, whether communication recipients pay attention to some types of gestures, such as emblems, more readily than to others, such as metaphorics).

Within the gesticulation categorisation a further typology has been established that is based on both the form and function of the action. Several schemas have been proposed (for thorough reviews on past typologies see Rimé & Schiaratura 1991, McNeill 1992 and Kendon 2004), and all have many similar features. Currently within the field, the most often used and in many ways most basic typology of gestures is that proposed by McNeill (1992) in which there are four main types of gesture.[1] The first are **iconic** gestures, which refer to concrete objects, such as framing the shape of a person in describing them. Secondly, there are **metaphoric** gestures, which generally have the same physical form as iconic gestures, but refer to abstract concepts and tend to be culturally specific in meaning, such as a notion of an idea as a vessel being represented by hands depicting it as though it were a physical object. The third category is that of **deictic** gestures, which refer to locations and ob-

[1] A survey of articles in *Gesture* from Vol 1, 2002 – 7(1), 2007 shows a total of 32 papers, 15 of which did not categorise gestures but use McNeill's terminology, 17 of which explicitly address features of Kendon's continuum.

jects that may or may not be visually present, such as pointing toward an object with a finger. The fourth category of gesticulation is beats. These are repetitive biphasal strokes of the hand that have an emphatic quality, drawing attention to what is being said with the beat falling on the stressed elements, e.g. *Don't you pinch the cat!* Since they don't refer to entities, real or abstract, and have more of a discourse function than the other categories McNeill (1992:15) notes that "of all gestures, beats are the most insignificant looking". However, in some arenas, such as Arandic sand-stories (Green 2009) and orchestral conducting (Parton 2007) it appears that beat gestures are oriented to substantially by interaction recipients.

While the gesture typology has found favour with many researchers it has been criticised by some members of the field. Farnell (1994) argues that it creates false divisions between various forms of movement and is self serving. Kendon (2004) is also wary of the gesture typology, claiming that while a typology may have use for a particular task "gesture can not be pinned down into a typology in any fixed way". (2004:84). Even McNeill (2000, 2005) has relaxed and expanded his original taxonomy, conceding that a gesture may exhibit "dimensions" of several categories simultaneously. In particular, from a technology perspective, in the building of interactive agents Karam & Schraefel (2005) have found only some features of the original gesture typology useful in Human-Computer interaction work, and Poggi (2002) has suggested that a subset of features are important for communication. The creation of a typology of gestures has been a useful tool for analysts in attempts to study gesture phenomena but to date there has been little research investigating whether these categories have any saliency for interactants – both speakers and recipients – who use them in everyday interaction.

In one study that has explored the attitudes of language users toward the various forms of bodily movement in communication, Kendon (1978) explored whether certain types of movement that constituted "significant action", such as gestures, received more attention from the recipients than others, such as fidgeting. The participants in Kendon's study were native English speakers who were Australian residents. They all viewed a silent film of a man speaking at an event in Papua New Guinea. After viewing it multiple times the participants were asked to point out where they observed movements. Kendon reports that all the participants mentioned the "significant" movements first and only after that did they mention the non-com-

municatively significant movements. This study indicates that, on some level, people have the ability to distinguish which movements are important to communication, even when the gestures are not culturally familiar and speech is absent. Kendon does not provide detail about which type of gesture events the participants focused on, nor does he test to see whether the presence of speech can affect the perception of "significant action". Several other studies have used videos with alterations to the audio and visual channels to investigate the perception of gesture in relation to speech (see Green 2009 for an examination of this in Arrernte narratives, Feyereisen et al. 1988, and also Feyereisen & Lannoy 1991 for a review).

Despite the fact that there have been considerable refinements in the categorisation of meaningful speech-accompanied movement there is much still to explore regarding a recipient's conscious understanding of the form and function of gesture. To this end the study reported here investigates recipient perception of various types of gesture and non-gesture bodily movement during speech, with a view to determining whether this reflects the current theoretical understanding of communicative gesture posed by analysts.

In light of the above discussion, the following research questions have been developed:

1. Does the perception of bodily action in co-speech gesture communication change with the presence or absence of sound?
2. Do the gestures that language users pay attention to in any way reflect the gesture schema currently used by analysts?

3. Methodology

3.1. Participants

Data for the study comes from an internet-based survey of 48 anonymous people who participated without remuneration. None of the participants had any experience within the field of gesture research. All of the participants were English first language speakers residing in Australia. The first part of the study, where people were asked to count the total number of gestures, was a survey of 48 participants, of these 17 were male and 31 were female aged between 18 and 58 with a mean age

of 25.54. In the second part, where participants were asked to name the five "best" gestures in the video, two participants were excluded for non-compliance leaving a total of 46 participants in the second section of the study, of which, exactly half (23) viewed the video with sound, and half (23) viewed the video without sound.

3.2. Data Collection

All participants were asked to complete an online survey based on a video filmed by the researchers. The video was 52 seconds long, and showed a young woman discussing a meal at her neighbour's house. The woman was seated, and the shot showed from her head to just below her knees (see Figure 2, below).

Figure 2. Screen shot of short narrative used in the survey

The speaker worked from a loose script with the gestures included at specific spots. 12 movement events, predetermined by the researchers, featured in the video. These are detailed in Tables 1, 2 and 3 further below. The narrative was not overly scripted to allow the event to be as natural as possible which meant minor non-gestural movements such as blinks and in-breaths occurred with some regularity. For a transcription of the narrative see the appendix.

Movements in the video were carefully constructed and controlled. To the best that could be enacted, none of the gestures used were blends of categories, but prototypical examples of distinct categories. Metaphoric gestures were not included as they represent abstract entities as solid objects and therefore generally take the same form as iconic gestures. Due to such similarity in form it was felt that their

presence would result in a "doubling up", especially for viewing the video without sound, and were therefore not included.

Two of each type were used to give a greater range of data and to see if the participants focused upon certain categories more than others. For the deictic gestures one referred to an object within sight (*me*) and the other to an object out of view (*neighbour's house*). With the iconic gestures one referring to *mock dogs* had the hands represent the entity, while a gesture referring to 'brandy' used the hands acting out their role pouring the bottle. There was also a short sharp beat gesture and a longer less rigid one. For emblems, there were two performed with the hand (a thumbs-up sign and an ok sign) as well as two performed with the head (a nod and a shake). This was to allow the researchers to see whether gestures performed with the head were given any more or less attention than gestures performed with the hand. One of each gesture type was included in the first half of the video, and all the types were then repeated in the second half. This maintained balance and allowed the researcher to assess whether more focus was being given to the start or end of the film.

The body movements used by the narrator in the film are broken down into the categories of "gesticulations", "emblems" and "non-gestural movement" and are presented in the following tables. Table 1 below presents gesticulations via a description of the physical movement and the type coding according to the gesture typology discussed above. These are time-aligned with accompanying speech.

Gesticulations			
Time	**Description**	**Type**	**Accompanying speech**
3.3s-4.9s	Left hand points across body to the right side of screen, index finger extended.	deictic	My neighbour's house, she lives across...
6s-8.7s	Both hands held in front of body creating an oval	iconic	...ade these ahh vegetarian hot-dogs I think s..
12.2s-14.8s	Flat hand, palm down, two straight strokes	beat	and they were were so dis-gusting...
23s-29s	Partially closed hand, 5 rolling strokes	beat	... had custard and cream and ice-cream and jelly and cake uh...
30.2s-33.7s	Hand grasping imaginary bottle, pouring	iconic	um. it also had a lot of brandy as well. I was feeling a littl...
48.8s-50.5s	Flat hand pressed to chest	deictic	I think I'll cook myself din-ner befor...

Table 1. Gesticulations included in video

Table 2 below lays out the emblems present in the video. It presents a description of the gesture movement and a culturally relevant term for the gesture together with the time-aligned accompanying speech.

Emblems			
Time	**Description**	**Term**	**Accompanying speech**
10s-12.1s	Index and middle fingers extended and retracted several times	quote marks	...s like mock dogs
15s-16.3s	Head moves backward and forwards	head shake	couldn't bring myself to each th…
41.3s-43.2s	Thumb and index finger come together to create circle, other three fingers extended, facing up, palm outwards	ok sign	...ages, it was so good.
44.8s-48.3s	Head moves up and down	nodding	get myself invited back there but maybe just for dessert

Table 2. Emblem gestures included in video

Table 3 below shows the location in the video of the two most prominent non-gestural body movements.

Non-gestural movement		
Time	**Description**	**Accompanying speech**
16s-18.1s	Right hand brushes hair behind right ear	t'eat them um fort...
35s-37.6s	Speaker uncrosses legs and adjusts seating position	um. so uhh

Table 3. Non-gestural body movements included in the video

After developing a definition of gesture, participants were presented with the researchers' very specific and intentionally narrow definition as follows:

> A gesture is a movement you make while speaking that may mean something, or add meaning or emphasis to your speech.

This was to ensure that participant and researcher expectations were broadly in line with each other. Participants who agreed to work with the researchers' definition then progressed to the next phase of the survey. At this stage participants were randomly divided into two groups: 1) a group which received sound when viewing the video, and 2) a group which viewed the same video without sound. The participants were asked to view the video once, and then asked to view again, this time providing the total number of gestures they saw. After this, they were once again presented with the video and asked to describe what they saw to be the "best" gesture. The next screen then prompted them to describe four other gestures they also saw as being good examples of a gesture. By asking participants to describe these five gestures, the researchers were able to determine which gestures the participants prioritised as the most salient examples of gestural phenomena.

The use of the internet as a medium of distribution for this survey was chosen as it proved highly convenient for both the researchers and the participants. For the participants, the website allowed them to access the survey from any computer with internet connection and sound, at any time. For the researchers, the medium allowed not only convenient distribution of the survey, but also convenient collation and storage of data. The use of a website allowed information to be easily given to, or withheld from, the participants until the time that the researchers saw fit; the researchers' definition, for example, was not given to the subjects until after they had given their own. The participants were also very easily divided between the sound and non-sound groups in an arbitrary and inconspicuous way via this process.

3.3. Data Analysis

Results for the study were then collated, coded and analysed. Since the participants gave their five best gesture examples in qualitative form it was necessary to convert the data to numerical form. The data were coded on the basis of the category that each gesture belonged to. Two researchers then coded the data independently and the third moderated any discrepancies.

An independent *t*-test was run on the total number of gestures found by participants in each group to ascertain any significant differences between the total number found based on the variable of sound. A chi-square was then run to establish whether the presence or absence of sound affected which gestures a participant was

likely to focus on. The frequencies of gesture categories selected by the group with access to sound were also tallied and compared. A final analysis was performed tallying the number of gestures described by the group with sound which were found in the first and second half of the video, ensuring the variable of video length did not interfere with results.

4. Results and Discussion

Overall, results indicate that the variable of sound does not affect the frequency of gestures made salient to the recipients but it does have a bearing on the saliency of certain gesture types. While recipients focused on a broader scope of gesture than those found in the analysts' schema, overall recipients oriented to the gestures found in the schema.

We now turn to the first research question:

> Does the perception of bodily action in co-speech gestural communication change with the presence or absence of sound?

An independent *t*-test showed no significant difference between the number of gestures counted by the group with sound versus the group without sound. However the types of gestures oriented towards did vary between the groups, as can be seen in Figure 3 below. Figure 3 represents the results of a chi square and as such we not only see the total number of each category for those with and without sound, but also the expected total for each category. The "other" group are movements from the video that do not fall into any of the categories, namely the intended non-gestural movements incorporated into the video or the unintended non-gestural movements.

For the group who viewed the video with sound, a mean of 15.57 gestures was counted (sd = 5.358), for the group who viewed the video without sound, a mean of 16.87 gestures was counted (sd = 6.676). The difference in the total number of gestures the participants in each group recorded was found to be non-significant ($t = 0.731$, $p = 0.469$), indicating that while the group without sound did count slightly more gestures on average than the group with sound, the total number of gestures did not vary significantly between these two groups.

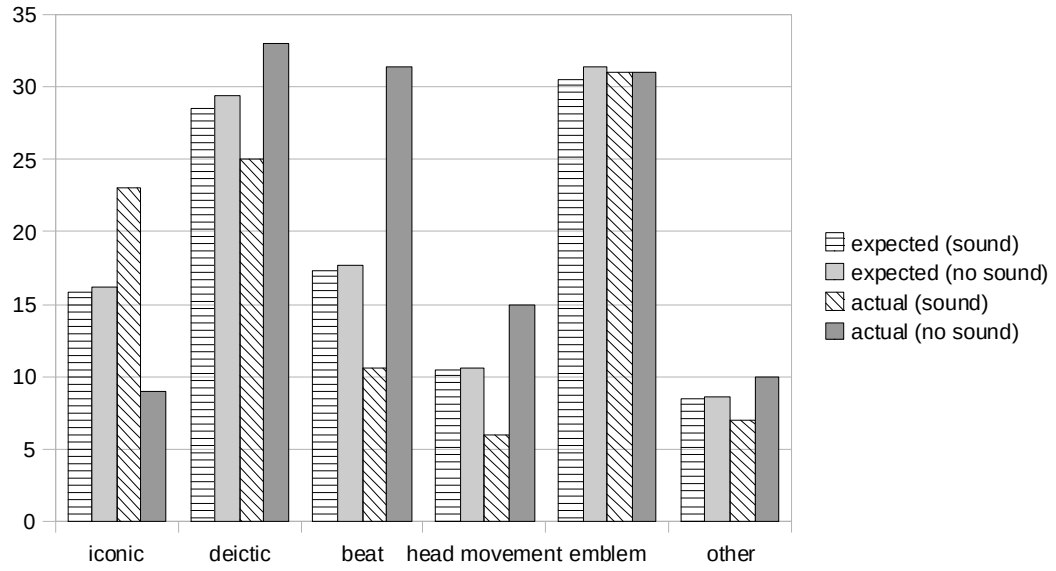

Figure 3. Chi square expected and actual totals for both groups

The average number of gestures counted in each group was well above those transcribed by the researchers; both groups counted an average of 50% more than the researchers' tally. It must be noted that while each group counted a similar number of gestural events, the study design does not allow confirmation that the participants in each group counted the same movements. In an augmented study of 23 new participants it was found that participants consistently oriented towards the same movements in their gesture counts. Further research in this area aims to determine exactly what events participants orient towards (Gawne & Kelly, in prep.). The similarity in the number of gestures counted by each group implies there is no observable difference in frequency of perception of bodily action when sound is removed. This finding supports Kendon's (1978) original research using the silent film to examine participants' impressions of gesture events, since the presence or absence of speech would have made no difference to the frequency of gesture identification in his data.

While the number of gestures counted remain the same in relation to the sound variable, we can see that participants' orientation towards certain gesture types is affected by the presence or absence of sound. A chi-square was run on the total number of gestures counted by participants in each category, and is summarised in Fig-

ure 3. The chi-square recorded a significant difference between expected and actual gesture category counts for both the sound and non-sound groups ($\chi^2 = 11.834$, $p = 0.036$). As can be seen from the representation of the chi-square (Figure 3 above) the strongest difference in frequencies can be seen in iconic gestures, the group with sound recording a higher frequency than those without. Head gestures and deictic gestures also prompted discrepancies in frequencies with a higher frequency for those without sound. Frequency differences between beat gestures, emblems, and other movements were minimal.

The number of emblems prioritised did not vary significantly in frequency between the two groups. This was largely due to the overwhelming focus on these two gesture events; all but five participants prioritised one, if not both of the emblem gestures shown, regardless of the sound variable. This leads to very limited room for variation between the two groups and implies that there is a great deal of saliency attached to these gestures, partially due to their ability to function symbolically with or without speech.

Iconic gestures showed the highest variation of any gesture category, with more participants in the with-sound group prioritising them over participants without access to sound. Those with sound included iconic gestures 23 times, while those without sound only included iconic gestures 9 times. The finding that these gestures were favoured so highly by those with access to sound, but less prioritised by those without sound, provides evidence that the speech context is important for the interpretation of these gestures. Those participants without sound had to resort to describing the hand movements for the gesture, while those with sound had instant access to the spoken object that the speaker was trying to describe, appearing to make the gesture much more meaningful and memorable. For example, a participant with access to sound described the gesture as "**the hot dog** shows the shape" (emphasis added); the gesture itself was described easily, and the motivation for the gesture could be accurately described. A participant without access to sound described the gesture as "**hands in circle** – showing the size of something?" (emphasis added); the participant appears less sure of the motivation for this gesture.

There was also a small variance in frequencies of deictic gestures and head movements, which follows a converse pattern. In this case participants without sound

were more likely to prioritise these movements, possibly motivated by the factor discussed above; participants without sound were more likely to focus on the deictic gestures and head movements as they are less dependent on the co-speech context for their meaning than iconic gestures.

Beat gestures only minimally featured in participants' orientations towards gestures, and showed very little variability in frequency between the two groups. Those without sound counted slightly fewer beat gestures (16) than those with sound (19), but this is not significant and the saliency of meaning for beat gestures is not affected by the variable of sound. This in turn implies a generally low level of salient meaning found in these gestures, which supports McNeill's (1992) observation that beat gestures are often viewed as the least significant gestures, regardless of the sound variable. While the variable of sound does not appear to affect the way participants viewed the video when only the total number of gestures counted is analysed, we can see differences in the two groups when we look at what types of gestures participants are orienting towards.

We turn now to the second research question:

> Do the gestures that language users pay attention to in any way reflect the gesture schema currently used by analysts?

As can be seen in the data above, participants counted a much higher number of gestures than analysts, indicating that participants oriented towards a wider range of phenomena.

In comparing the way language users and analysts perceive gesture we chose to examine only the data of the group with access to sound since these worked as a control group for the experiment. The choice of gestures focused on by the participants was examined and the frequency of gestures tallied by category.

The results of the with-sound participants' choice for the "best" example of a gesture, that is, the one that they wrote down first, appear to indicate a strong preference for emblematic gestures, over all others. Figure 4 shows the raw score numbers for each category of gesture and the order in which they were written by participants. This indicates that these gestures with culturally encoded names and

meanings, such as the ok sign, have a salient meaning for speakers. More idiosyncratic gestures, such as iconic or beat gestures, were less prevalent in this stage.

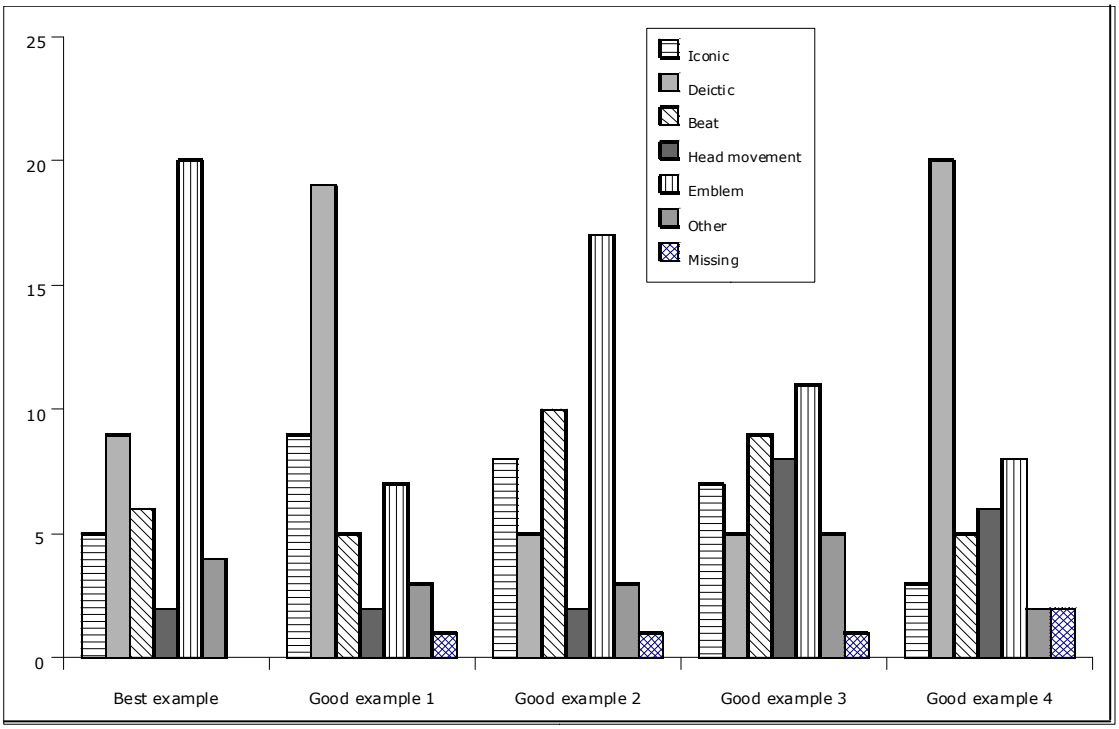

Figure 4. With-sound group: Gesture numbers across category for five best examples

However, when the results of all five "good" gestures are combined, we see that the prevalence of the different categories, while not equal, were still broadly distributed, as shown in Figure 5 below which indicates the best examples of gestures across gesture categories.

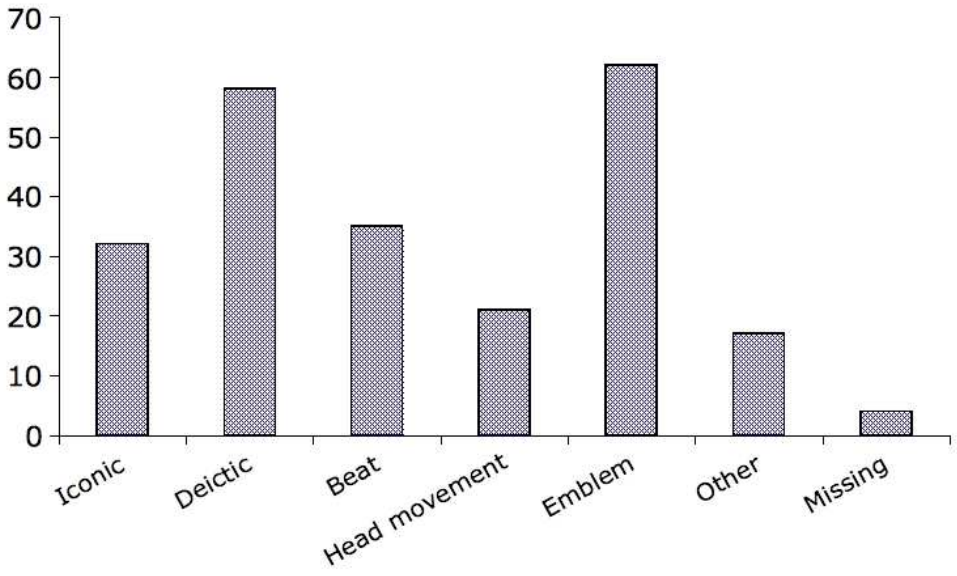

Figure 5. With-sound group: Total number of best-gesture examples, by category

If the participants' views of a gesture were entirely divorced from the linguist's typology, we would have expected a different distribution to that seen in Figure 5, and the highly idiosyncratic gestures would have featured, only minimally. Hence, we can see that the participants' choices were distributed among a range of categories.

Of interest to researchers concerned about the current gesture typology being narrow and self-serving (Farnell 1994, Kendon 2004, Poggi 2002) is that while some participants included movement events not considered to be gestures within the standard gesture typology, these were less prioritised by the subjects when compared to those found within the typology. This study also provides post hoc support for Kendon's (1978) study that people will consistently focus on specific "significant actions" during discourse, while also providing more information, as to the hierarchy of this focus.

Participants described several communicative movement events that are not generally considered in current gesture schemas, such as posture, eye movement and clothing adjustment.[2] This was evidenced by the fact that the number of gestures

[2] A single participant intriguingly drew communicative inferences from the name of the band on the t-shirt worn by the woman in the video.

counted by each participant was no lower, and in many cases much higher than those counted by the researchers in the study and coded by the gesture analysts. While it is possible that beat gestures may have been separated into their component parts by the participants without sound, this increase alone would not explain the variance between the 10 gestures counted by the researchers, and the average of 15.57 gestures counted by those who watched the video with sound, and average of 16.87 for those who watched with no sound. Nor can it explain the large numbers counted by those participants with access to sound.[3] It does seem clear that movements which fall outside the researchers' typology are apparent in the participants' counts, however, the number of these choices which were prioritised when participants listed their five "best" gestures was quite low, and as such does not bring into question the validity of the current typology. The frequency of head gestures described was comparatively low. In the current classification head movements are considered emblems, however, the two head movements in the video received far fewer mentions than the two manual emblems, implying an inherent hierarchy of gestures held by the participants which preferences hand movements over non-hand movements. This is interesting since only 9 people in this group defined gesture with explicit reference to the hands, and in all of those definitions they included other body parts that could also be used to make gestures. This implies that the presence of a subconscious gesture hierarchy would benefit from further exploration.

As with any study that draws participants' attention toward the topic of focus there is an inherent risk they will become hyperaware of the phenomena being discussed. This may account for the fact that nobody recorded a gesture count of less than ten, and many people counted a considerable number more. These participants may have been seeking out gestures that they may have not normally considered as such. It is difficult to find a workable way to eliminate this, and as such all results should always be viewed with this in mind. There is also the risk of the "subject expectancy" effect (Brown 1988) where participants attempt to second-guess the intentions of the research. However, this phenomenon was minimised though the use of

[3] Interestingly, of the 19 qualitative descriptions of beats, only one described individual components.

a relatively de-personalised web tool which gave the participants greater anonymity thereby potentially reducing the need to "impress" the researchers.

While there are potential flaws in the video medium, such as its length encouraging participants to focus only on earlier or later gestures, or the occurrence of unavoidable bodily movements which could be seen as distracting, we do not believe that these could in any way be seen to invalidate the experiment. An analysis was performed tallying the number of gestures described by with-sound group found in the first and second half of the video, to test whether video length interfered with results. It was found that exactly 50% of gestures described occurred in each half of the video. 100% of participants also described gestures from both the first and second half of the video. Gestures focussed on were found to be evenly distributed in the time frame of the video, and the natural movements were not so great nor so frequent as to negatively impact the experiment. As a medium for studying gesture perception this short clip proved highly successful.

5. Conclusion

Conversation participants care about body movements. More specifically, they orient to many of the bodily actions gesture researchers have categorised in gesture typologies. The current study has established that accompanying speech sound does not influence the frequency of gestures that recipients consider to be salient. However, differing categories of gesture did receive different levels of focus of meaning depending on the variable of sound. Iconic gestures, which are highly idiosyncratic, and which lack a clearly definable meaning when stripped of their co-speech context, are less likely to be prioritised when viewed in this context. Gestures which possess more salient meaning in a context free of co-speech – such as deictic gestures or head movements – tend to be more highly prioritised in these context, as they appear to have a higher saliency of meaning to those viewing them. Interestingly, while participants did show a clear initial preference for manual emblematic gestures, in their broad discussions of gesture, participants focused on the range of gestures found in current gesture schemas.

This study is an initial step in an on-going research program designed to examine the perception of gesture by English speakers across various genre of communica-

tion including narratives and conversation, and in a lab-based setting rather than the depersonalised internet environment (Gawne & Kelly in prep.). As the study of gesture continues to gain popularity and the amount of research based on the established theoretical framework grows, it is occasionally helpful to examine the field from the vantage point of everyday gesture users. To this end the study reported here has paved the way for a line of enquiry to aid gesture researchers in understanding the discipline from a new angle.

Acknowledgements

The researchers would like to thank Kent Humphries and Angus MacAuley for their invaluable assistance in producing the web interface for this experiment, and Lawrence Cavedon for highlighting the need for such a study. Thank you to the two anonymous reviewers for their comments.

Bibliography

Brown, James Dean. 1988. *Understanding research in second language learning.* Cambridge: Cambridge University Press.

Farnell, Brenda. 1994. Ethno-graphics and the moving body. *Man* 29(4). 929-974.

Feyereisen, Pierre & Jacques-Dominique Lannoy. 1991. *Gestures and speech: Psychological investigations*. Cambridge: Cambridge University Press.

Feyereisen, Pierre, Michèle Van de Wiele & Fabienne Dubois. 1988. The meaning of gestures: What can be understood without speech? *European Bulletin of Cognitive Psychology* 8. 3-25.

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

Green, Jenny 2009. Between the earth and the air: Multi-modality in Arandic sand stories. Melbourne: University of Melbourne PhD dissertation.

Karam, Maria & m. c. Schraefel. 2005. *A Taxonomy of Gestures in Human Computer Interactions*. Technical Report ECSTR-IAM05-009, Electronics and Computer Science, University of Southampton.

Kendon, Adam. 1978. Differential perception and attentional frame in face-to-face interaction: Two problems for investigation. *Semiotica* 24(3/4). 305-315.

Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.

McNeill, David. 2000. Introduction. In David McNeill (ed.), *Language and gesture*, 1-10. Cambridge: Cambridge University Press.

McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press.

Parton, Katharine 2007. Conducting gesture and McNeill's continua. Paper presented at Inaugural International Conference on Music Communication Science (ICOMCS) University of New South Wales Sydney, Australia

Poggi, Isabella 2002. *From a typology of gestures to a procedure for gesture production*. In Ipke Wachsmuth & Timo Sowa (eds.), Gesture and sign language in human computer interaction: Revised papers. International Gesture Workshop, GW 2001, London UK, April 18-20, 2001 (Lecture Notes in Computer Science 2298), 158-168. Berlin & Heidelberg: Springer.

Rimé, Bernard & Loris Schiaratura. 1991. Gesture and speech. In Robert S. Feldman & Bernard Rimé (eds.), *Fundamentals of non-verbal behavior*, 239-284. Cambridge: Cambridge University Press.

Appendix

Transcribed text of the video narrative

A: So what did you do last night for dinner?

B: Last night for dinner I went to my neighbour's house, she lives across the road, and she made these uh vegetarian hot dogs. I think she called them something hilarious like "mock dogs". And they were so disgusting – I couldn't bring myself to eat them. um. Fortunately though she makes the best desserts. um. She made a trifle and it had custard and cream and ice-cream and jelly and cake uhh it was so good. um. it also had a lot of brandy as well. I was feeling a little bit ill by the end of it. um. so uhh it was seriously like the best trifle I think I've had for ages. It was so

good. I'll um I'll have to try and get myself invited back there but maybe just for dessert. I ah I think I'll cook myself dinner before I go.

A: (Laughs)