The flexible semantic integration of gestures and words

Comparing face-to-face and telephone dialogues

Jennifer Gerwing and Meredith Allison
University of Victoria / Elon University

One measure of the communicative function of gestures is to test how speakers' gestures are influenced by whether an addressee can see them or not, that is, by manipulating visibility between participants. We question traditional dependent variables (i.e., rate measures), suggesting that they may have been insufficient for capturing essential differences in the gestures speakers use in each condition. We propose that investigating the qualitative features of gestures is a more nuanced, and ultimately more informative approach. We examined how speakers distributed information between their gestures and words, testing whether this distribution was affected by the visibility of their addressee. Twenty pairs of undergraduates took part in conversations that were either face to face ($n = 10$) or on the telephone ($n = 10$). Each speaker described a drawing of an elaborate dress to the addressee. We used a semantic feature analysis to analyze descriptions of the dress' skirt and assessed when words or gestures contributed information about five categories pertaining to features of the skirt's unusual shape. Although speakers' rates of gesturing and number of words did not vary significantly between conditions, speakers contributed more information and conveyed more categories in their gestures when the addressee would see them, while words carried the informational burden when addressees would not see the gestures (p's $< .001$). These results suggest that gestures serve a communicative function. The semantic feature analysis is thus an example of how to explore gestures' qualitative features within a quantitative paradigm.

Keywords: hand gestures, semantic information, dialogue

Observations of gesture use, both inside and outside the laboratory, have led researchers to question the functions conversational hand gestures serve. Gestures' depictive qualities (e.g., their particular form and movement) could be intentional:

Gesture 11:3 (2011), **308–329**. DOI 10.1075/gest.11.3.03ger
ISSN 1568–1475 / E-ISSN 1569–9773

the speaker could form the gesture to convey particular information to his or her addressee (e.g., McNeill, 1992). In contrast, the qualities that happen to be depictive might emerge only as unintentional byproducts of the speaker's cognitive processes (Krauss, Dushay, Chen, & Rauscher, 1995). One test of gestures' function, specifically whether speakers use gestures to communicate, is to manipulate visibility between participants. If gestures primarily serve a communicative function, then speakers should gesture only when the addressee can see them. Interpreting the results of visibility studies, however, has been far from straightforward. In several studies, gesture rate did decrease significantly when the addressees were behind partitions, on the other end of an intercom system, or on the telephone (Alibali, Heath, & Myers, 2001; Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995). However, even when addressees would not see their gestures, speakers often still accompanied their speech with gestures. In this paper, we question whether traditional rate variables are nuanced enough to capture communicative function. We propose that a more informative approach would be to analyze the qualitative features of gestures, including how visibility influences gesture features, speakers' words, and the semantic relationship between gestures and words.

The following two examples, taken from our data, illustrate why investigations of this relationship could shed light on gesture function. The speakers in both examples were describing an elaborate, somewhat bizarre, dress (see Figure 1; picture from Blum, 1982, p. 14). The first speaker, who was talking to an addressee who could see her, said, "and it goes like out." The speaker accompanied these somewhat elliptical words with a gesture (Figure 2 shows three screen shots from this gesture): She began with both hands on her waist, then she extended both hands out, tracing a symmetrical, horizontal line that extended approximately two feet on either side of her. She ended the gesture first by curving her fingers slightly downwards and then dropping them. These gestures disambiguated her words, tracing the shape and size of the top of the skirt. In contrast, a speaker who was describing the dress to an addressee who was on the telephone said, "like it just looks like a table, right." This speaker also gestured. She moved the fingers of her right hand slightly to the right in a horizontal motion then down, forming what looked like the corner of a table. (Figure 3 shows the beginning and end of this gesture.) By tracing a shape like the top corner of a table, this movement was congruent with the speaker's words; however, the gesture did not appear to convey information over and above what the words themselves conveyed.

These examples illustrate semantic coordination between words and gestures, which scholars of language-use have characterized as *integrated messages* (Bavelas & Chovil, 2000, 2006) or *composite signals* (Clark, 1996). Indeed, Kendon (2004) defined an utterance as being "fashioned from both spoken language and gesture",

Figure 1. The stimulus picture (from Blum, 1982, p. 14)

and he stated that "gestures used by speakers while they speak are partnered with speech as a part of the speaker's final product and are as much a part of the utterance's design as the speaker's words" (p. 5). However, even though these two examples provide demonstrations of integrated messages, they appear to differ in several respects. Whereas the speaker in face-to-face dialogue traced a life-sized outline

Figure 2. Participant depicting the shape of the skirt with a gesture: "and it goes, like, out"

Figure 3. Participant depicting the shape of the skirt with a gesture: "Like it just looks like a table"

of a skirt, the speaker in the telephone dialogue made a small gesture that was not necessarily suggestive of a part of a dress. In contrast, the words of the speaker in the face-to-face dialogue conveyed little information, but those in the telephone dialogue provided an evocative demonstration of the overall shape of the skirt. These examples suggest that the mere presence of gestures (typically measured in terms of gesture rate or frequency) might not be sufficient to address whether gestures serve communicative functions. In the literature review that follows, we

will explore this issue in more detail: In visibility studies, do researchers focus on gestures and words as integrated messages?

Effect of visibility on gesture rate

Comparing how speakers' gesture rate responds to whether the addressee can see those gestures has been the most ubiquitous and straightforward test of gestures' communicative function. In all studies, when addressees could not see them, speakers continued to gesture, either just as much as when the gestures would be seen (Bavelas, Gerwing, Sutton, & Prevost, 2008; Gullberg, 2006; Lickiss & Wellens, 1978; Rimé, 1982) or significantly less (Cohen, 1977; Cohen & Harrison, 1973; Emmorey & Casey, 2001; Krauss et al., 1995). It is difficult to argue that speakers intend for their gestures to communicate when speakers gesture whether the addressee sees the gestures or not.

Some gestures appear more illustrative than others, but even when focusing just on these gestures, results of visibility studies do not necessarily support their communicative function. Whereas some researchers reported that the rate of gestures illustrating the topic of discussion was unaffected by visibility (Bavelas, Chovil, Lawrie, & Wade, 1992; Bavelas et al., 2008), others reported that these gestures were more frequent when they would be seen (Alibali et al., 2001). But again, in both cases, speakers continued to contribute illustrative gestures when speaking to addressees who would not see those gestures. It seems that a focus on gesture rate alone does not provide clear evidence as to whether speakers intend for gestures to serve a communicative function. In the next section, we describe research that explored whether gestures that could be seen had any distinguishing features compared to those that could not be seen.

Effect of visibility on the qualitative features of gesture

In the examples we provided above, the qualities of the two gestures were different. Whereas the gesture for the visible addressee was large and clearly depictive of the shape of the dress, the gesture for the addressee on the telephone was small and more ambiguous. Although it may be more challenging to attempt to operationalize features of gestures that could be associated with a more communicative function (e.g., size, clarity of form), such analyses could provide a more nuanced focus than analyzing gestures' mere presence. However, in contrast to gesture-rate variables in visibility studies, only a few researchers using this particular manipulation have analyzed how qualitative features of gestures are influenced by whether the addressee could see them. In five studies where researchers operationalized and compared qualitative features, we found more predictable and conclusive findings.

Bangerter (2004) presented pairs of participants with an array of photos. One in each pair had to help the other to identify particular sets photos. By manipulating visibility between participants and the distance that the array was from them, Bangerter found that participants pointed more frequently when they could see each other and when the referent array was close enough for the pointing to indicate photos unambiguously. Speakers in a study by Gullberg (2006) studied a cartoon narrative and then relayed the story either to addressees who could see them or who could not. When addressees could see the gestures, speakers were significantly more likely to use anaphoric gestures that were well-defined and differentiated spatially and that adhered to a more stable location during that narrative. Bavelas et al. (2008) instructed speakers to describe a picture of a dress to an addressee. In the face-to-face condition, speakers produced gestures that were life-sized (often placed around the speaker's body). In contrast, speakers on the telephone created significantly smaller gestures that matched the size of the picture. Kimbara (2008) asked participants to retell a cartoon together and found that when they could see each other, their gestures became very similar in form. Using a referential communication task, Holler and Wilkin (2011) found that gestural mimicry occurred significantly more frequently in face-to-face dialogue than in dialogues where participants were separated by a screen, and it played a vital role in how the participants established mutual understanding. In each of these studies, the form of gestures (not just their rate) was influenced by visibility between participants. If the speakers' gestures were serving purely cognitive functions, there would be little reason to expect that their form would change in any respect according to whether the gestures would be visible, let alone change in a way that would make the gestures a more effective communicative resource for the addressee. These qualitative variables have provided clearer evidence for gestures' communicative function than rate measures have.

Effect of visibility on the relationship between gestures and words

Another test of gestures' function is to analyze the relationship between gestures and words. If gestures are serving a unique communicative function, in face-to-face dialogue, they should convey information that is otherwise missing from the words. Furthermore, this semantic relationship should be predictably influenced by whether the gestures will be seen by the addressee: When participants cannot see each other, the distribution should shift so that the speaker's words should convey more information. This opposite pattern would indicate that speakers in face-to-face dialogue may indeed be using gestures intentionally to communicate; that is, the information conveyed by those gestures would not be an unintentional byproduct of speakers' cognitive processes.

We have found three visibility studies that included variables investigating this semantic relationship. These studies present results about either deictic expressions referring to gestures (Bangerter, 2004; Bavelas et al., 2008; Emmorey & Casey, 2001) or gesture-word redundancy (Bavelas et al., 2008). When addressees were able to see the speakers' gestures, speakers used significantly more deictic expressions referring to those gestures (such as "like *this*" while gesturing the shape of a part of a dress, "over *here*" while pointing in a particular direction) than when addressees could not see the gestures (Bavelas et al., 2008; Emmorey & Casey, 2001). Similarly, when pointing would be an effective way of distinguishing a referent, participants used pointing with deictic expressions significantly more often (Bangerter, 2004). Bavelas et al. (2008) analyzed gesture-word redundancy in speakers' depictions of the neckline of a dress and the shape of the skirt. In both kinds of descriptions, speakers in the face-to-face condition were significantly more likely to convey non-redundant information in gestures than they were in the telephone condition. That is, compared to telephone dialogues, gestures in the face-to-face dialogues were significantly more likely to convey information that was otherwise missing from the speakers' words.

The redundancy finding, in particular, raises some interesting questions. As we outlined in an earlier paper (Gerwing & Allison, 2009), a redundancy analysis gives primacy to gesture meaning and is therefore vulnerable to overestimating gestures' contribution: Analysts can imbue any representative aspect of a gesture with meaning that may be missing from the words. Furthermore, a redundancy analysis does not directly address the informative value of the words themselves, nor does it specify what information the gesture did contribute. Although the redundancy analysis in Bavelas et al. (2008) was promising, we feel it suggests that a more comprehensive and informative measure of the communicative function of gesture use is justified. Such a measure would ideally compare the semantic contribution of gestures and words in differing conditions. If speakers combine these two modalities into integrated messages or composite signals, they should do so with sensitivity as to how addressees can perceive and interpret them.

In the study that follows, we applied a *semantic feature* analysis to the data. For this analysis, researchers use characteristics of the referent (i.e., what the speaker was describing) to stipulate categories of information and then assess whether words or gestures (or both) contribute information about each category. Results of this type of analysis have demonstrated specific differences in how gestures and words represent meaning (Beattie & Shovelton, 1999, 2002; Gerwing & Allison, 2009; Holler & Beattie, 2002, 2003, 2004; Holler & Stevens, 2007; Holler & Wilkin, 2009). Although it would appear to have potential as a useful test of whether speakers adjust their relative contributions in response to visibility, to our knowledge, it has not been used to explore this question.

We developed a semantic feature analysis to reveal how speakers distributed five kinds of information between gestures and words when they were speaking face to face, and we tested whether that distribution changed when they spoke to each other on the telephone. Each speaker described a drawing of an elaborate dress to an addressee. Half of the speakers and addressees had their conversations in face-to-face dialogue, and half had them on the telephone. Our hypothesis was the following: If speakers use gestures to communicate, then speakers will change the relative semantic distribution between gestures and words according to whether the addressee will see the gestures or not. Accordingly, we had four specific predictions. The first two focused on each conversational context on its own, exploring the distribution of information between gestures and words:

1. In face-to-face dialogues, speakers would convey information about the skirt more often in their gestures than in their words. To test this prediction, we conducted a within comparison of information conveyed in gestures and words in the face-to-face condition.
2. In telephone dialogues, speakers would compensate for the lack of visibility by conveying more information in their words than their gestures. To test this prediction, we conducted a within comparison of information conveyed in gestures and words in the telephone condition.

The next two predictions focused on each modality separately, comparing gestures across the two conditions and then comparing words across the two conditions.

3. Gestures produced in face-to-face dialogues would be qualitatively different from those in telephone dialogues; specifically, gestures that addressees would see would be more informative than gestures they would not see. To test this prediction, we conducted a between comparison of the information conveyed by gestures produced in the face-to-face condition vs. the telephone condition.
4. Words produced in face-to-face dialogues would be less informative than words produced in telephone dialogues. To test this prediction, we conducted a between comparison of the information conveyed by words produced in the face-to-face condition vs. the telephone condition.

Method

The data for this analysis are from the face-to-face and telephone conditions in a previous study, which is reported in detail in Bavelas et al., 2008. That study included a third condition (participants described the dress alone in a room into

a tape recorder), which we did not include in our analysis because it was not relevant to the questions we were posing. Data collection and preparation are described here only briefly.

Data collection

Forty undergraduate psychology students participated in pairs, and they were randomly assigned to the role of speaker or addressee. All participants consented in writing to being videotaped while doing a number of conversational tasks. Ten pairs of participants did these tasks in a face-to-face dialogue and ten on the telephone (the speaker remained in the lab to be filmed, and the addressee participated by telephone from another room). The task of interest for this paper was for the speaker to describe a black and white drawing of an intricate 18th century dress (see Figure 1) to the addressee, who could not see the picture and who would later have to choose the described dress from an array of similar dresses. The speaker and addressee could interact freely while completing the task.

Three tightly synchronized cameras videotaped these conversations: The top two-thirds of the screen showed a face-on view of both the speaker and the addressee, the bottom third showed a side view of the two participants (as in Figure 2). In the ten telephone dialogues, the screen configuration stayed the same, even though only the speaker could be seen on videotape (as in Figure 3). After the tasks were completed, the experimenters debriefed all participants, who saw the video of their participation and had the opportunity to give or withhold permission to have their video used for a variety of purposes (e.g., for analysis, for example at conferences or classes).

Analysis

Data preparation

Analysts in the previous study (Bavelas et al., 2008), had transcribed the words. They reliably located all topic gestures, which were meaningful hand movements synchronized with speech that depicted some aspect of the topic of conversation (ibid, see pp. 507–508), which was, in this case, the dress. We used excerpts from each speaker's description during which the speaker described the unusual shape of the skirt from the woman's waist outwards to and including the rounded corners. Our analysis included all words and topic gestures from these excerpts.

Definitions of semantic categories

We analyzed the excerpts using five semantic categories that were based on the distinguishing features of the skirt's shape. If the speaker wanted the addressee to be able to distinguish this dress from other possible dresses, these are five features that he or she could express in words or in gesture. Full operational definitions are available from the first author.

Identity. In each excerpt, the speaker described the skirt (as opposed to the neckline, bodice, or sleeves). Speakers could identify the skirt as being the topic of the description with their words (e.g., "the *skirt* goes like this") or their gestures (e.g., placing the gestures around the lower half of their body). Analysts did not resolve pronouns when making decisions; that is, the word "it" did not count as identity, even if it was referring to the skirt.

Waist. The skirt began at the woman's natural waist (as opposed to above it, as in empire style, or below it, as in a mermaid style). Speakers could use words to depict this location (e.g., "it starts at the waist") or they could use gestures (e.g., by touching their own waist).

Flatness. As can be seen in the picture of the dress, the fabric did not drape down gracefully from the woman's waist, instead the top of the skirt protruded to the woman's left and right sides in a horizontal, flat fashion. Speakers could use words to depict this flat shape (e.g., "it shoots straight out"), or they could use gestures (e.g., tracing a horizontal straight line).

Size. The skirt was also unusually big, which speakers could convey verbally (e.g., "big" or "massive") or gesturally (by making a big gesture).

Corners. The fabric draped down, forming almost right-angle corners; the draping fabric was perpendicular to the horizontal, flat top. Speakers could depict these corners verbally (e.g., "it kind of curves downwards") or gesturally (using a rounded handshape).

Note that we designed our operational definitions and coding procedures to address the possible confound presented by the fact that speakers in one condition were holding a telephone in one hand. In all definitions, gesturing the same feature with one or two hands would not affect the ultimate dependent measure. For example, while a speaker in the face-to-face condition could use both hands to trace symmetrical, horizontal lines extending from his body, a speaker in the telephone could use only one hand to do so. In each case, the speaker would receive only one point for the semantic category of flatness.

Analysis of words

Analysts used detailed operational definitions to decide whether each speaker's words conveyed information pertaining to each of the five categories described above. They listened to the audio from the excerpts (without viewing), using a transcription of just the words (without gesture location) to record their decisions. Each contribution of information in a category was scored as a point in that category. For example, the words "it's like massive" would score "1" in the size category and "0" in the remaining categories. "The skirt's massive" would score "1" in the size and identity categories and "0" in the remaining three categories. Each speaker's raw scores for the words were the number of times words in the description invoked each category.

One analyst scored all the verbal descriptions, and a second analyst, who was blind to condition, scored a randomly selected 6 of the 20 speakers. For calculating reliability, whenever the analysts scored a word or phrase, we checked whether the analysts agreed on the score for each of the semantic categories. For example, if analysts agreed that "it's like massive" counted as "1" for the size category, they received credit for that agreement as well as for agreeing that it was "0" in the other four categories; that is, their agreement was 5/5 for that phrase. On the other hand, if one analyst decided "it's like massive" conveyed both identity and size, but the other analyst decided it conveyed only size, their agreement would be 4/5 for that phrase. Agreement for the two analysts' decisions was 231 for the 250 decisions (92.4%).

Analysis of gestures

For the gestures, analysts used detailed operational definitions for how a gesture could contribute information about any of the semantic categories. They watched the video excerpts (with audio) and recorded their decisions on a transcription of the words that included the location of the topic gestures.

One analyst scored all the gestures, and a second analyst scored a randomly selected subset of approximately 20% of the total number of gestures (32 gestures, or 160 decisions as to whether the gesture was depicting information pertaining to each of the semantic features). Agreement was calculated as described above. Analysts agreed on 137 of the 160 decisions (85.6%).

Aggregating scores

For each speaker, we calculated proportions for the contributions of words and of gestures in each semantic category using the following method: The total number

of contributions for each semantic category provided the denominator, and the contributions from words and gestures provided the numerators for each of their proportions. Our results are based on the mean proportions for each category across the 20 speakers.

Results

First we compared how many words the speakers used, their rate of gesturing, and the duration of their gestures. Participants speaking face to face used an average of 94.80 words ($SD = 49.17$), and participants on the telephone used an average of 71.40 words ($SD = 50.20$). A between-subjects t-test indicated that the difference between these two means was not significant ($t_{(18)} = 1.05$; $p = .31$). In the face-to-face condition, the average gesture rate was 11.16 gestures per 100 words ($SD = 3.56$) and on the telephone, the mean gesture rate was 8.86 gestures per 100 words ($SD = 6.35$). Again, these rates were not significantly different ($t_{(18)} = 1.00$; $p = .33$). Further, gestures in the face-to-face condition were the same duration ($M = 6.12$ words; $SD = 1.99$) as gestures on the telephone ($M = 5.96$ words; $SD = 1.73$). Whether the speakers described the skirt face to face or on the telephone, they used the same number of words, gestured at the same rate, and made gestures that were the same duration in relation to their words.

Distribution of information conveyed

For the semantic feature variables, all of our predictions were directional. Therefore we used one-tailed tests, and our significance levels are reported accordingly.

First, we compared the mean proportions of the semantic contributions of words vs. gestures in each of the conditions (regardless of semantic category), testing differences using paired-sample t-tests; these results are presented in Table 1. In the face-to-face condition, participants put more information into their gestures than into their words ($p < .001$). Conversely, in the telephone condition, participants put more information into their words than into their gestures ($p < .01$). We also compared how the conversational context (i.e., speaking face to face or on the telephone) affected how much information speakers conveyed in each modality. We tested this difference using independent-samples t-tests. Whereas speakers' gestures in the face-to-face dialogues were significantly more informative than speakers' gestures on the telephone ($p < .001$), speakers' words in the face-to-face dialogues were significantly less informative than speakers' words on the telephone ($p < .001$).

Table 1. Means and Standard Deviations of Proportions of Contributed Information in Gestures and Words in each Condition

Conversational context	Modality	M (SD)	$t_{(df)}$
Face-to-face	Gestures	.74 (.10)	$t_{(9)} = 7.42^{*}$
	Words	.26 (.10)	
Telephone	Gestures	.27 (.18)	$t_{(9)} = -4.02^{**}$
	Words	.73 (.18)	

Note. Tested with paired-sample t-test
* $p < .001$; ** $p < .01$.

Distribution of semantic information in each conversational context

Next we separated the data into the five semantic categories. We predicted that in each semantic category, in the face-to-face condition, gestures would convey more information than the words, and in the telephone condition, words would convey more information than the gestures.

Face-to-face condition. For four of the semantic categories, when the addressee could see their gestures, speakers conveyed significantly more information in their gestures than they conveyed in their words (see Table 2). Speakers used those gestures to convey information about the skirt's unexpectedly large size, the way it began at the waist, its horizontal, flat shape, and the way the fabric of the skirt draped down at the corners.

Table 2. Face-to-face condition: Means and Standard Deviations of Proportions of Contributed Information in Words vs. Gestures

Semantic Category	Modality	M (SD)	$t_{(df)}$
Identity of skirt	Words	.35 (.29)	*n.s.*
	Gestures	.65 (.29)	
Waist	Words	.26 (.18)	$t_{(9)} = -4.18^{**}$
	Gestures	.74 (.18)	
Flat shape	Words	.25 (.16)	$t_{(9)} = -4.89^{**}$
	Gestures	.75 (.16)	
Size	Words	.27 (.14)	$t_{(9)} = -5.27^{**}$
	Gestures	.73 (.14)	
Shape of corners	Words	.17 (.22)	$t_{(9)} = -4.67^{**}$
	Gestures	.83 (.22)	

Note. Tested with paired-sample t-test
** $p < .001$.

Table 3. Telephone condition: Means and Standard Deviations of Proportions of Contributed Information in Words vs. Gestures

Semantic Category	Modality	M (SD)	$t_{(df)}$
Identity of skirt	Words	.89 (.20)	$t_{(9)} = 6.11^{**}$
	Gestures	.11 (.20)	
Waist	Words	.60 (.46)	$t_{(9)} = 1.81^{*}$
	Gestures	.20 (.34)	
Flat shape	Words	.49 (.35)	*n.s.*
	Gestures	.50 (.35)	
Size	Words	.93 (.12)	$t_{(9)} = 11.32^{**}$
	Gestures	.07 (.12)	
Shape of corners	Words	.17 (.27)	$t_{(9)} = -2.12$
	Gestures	.53 (.44)	

Note. Tested with paired-sample t-test
$^{*}p = .052$; $^{**}p < .001$.

Telephone condition. When the participants were speaking on the telephone, the words were significantly more informative than the gestures were for the categories pertaining to the identity of the skirt, and its size (see Table 3). (Note that for the waist category, the difference approached significance; $p = .052$). For example, when conveying size, the speakers' gestures were often small, but their words reflected the size of the skirt (e.g., "it goes way beyond that of her shoulders", or "if you were to hold the sides of your dresses out and curtsy, but hold it out really far"). However, participants on the telephone conveyed the shape of the corners of the skirt in their gestures more frequently than they did in their words.

Distribution of semantic information in each modality

Next we tested how conversational context (face-to-face dialogue vs. telephone dialogue) affected each modality. We predicted that gestures in the face-to-face condition would be more informative than gestures in the telephone condition, and words in the telephone condition would be more informative than those in the face-to-face condition. We used independent *t*-tests to compare means.

Effect of conversational context on the gestures. When comparing gestures in the two conversational contexts, we found that in all five categories, gestures in the face-to-face conversations were significantly more informative than those in the telephone conversations (see Table 4). As we were conducting analysis, we had the impression that many of the gestures in the face-to-face descriptions conveyed several categories at once, just as the gesture in the example from our introduction

Table 4. Gestures: Means and Standard Deviations of Proportions of Contributed Information in the Telephone vs. Face-to-Face Conditions

Semantic Category	Conversation Context	M (SD)	$t_{(df)}$
Identity of skirt	Telephone	.11 (.20)	$t_{(18)} = 4.90^{**}$
	Face-to-face	.66 (.30)	
Waist	Telephone	.20 (.34)	$t_{(18)} = 4.35^{**}$
	Face-to-face	.74 (.18)	
Flat shape	Telephone	.50 (.35)	$t_{(12.64)} = 1.97^{*}$
	Face-to-face	.75 (.16)	
Size	Telephone	.07 (.12)	$t_{(18)} = 11.42^{**}$
	Face-to-face	.73 (.14)	
Shape of corners	Telephone	.53 (.44)	$t_{(13.31)} = 1.89^{*}$
	Face-to-face	.83 (.22)	

Note. Tested with independent-samples *t*-tests
$^{*}p < .05$; $^{**}p < .001$.

did. In contrast, gestures on the telephone often appeared to convey only one category of information at a time. For example, speakers might move one hand only a few centimeters from left to right, which would convey the flat shape of the skirt but nothing else. We tested our impression statistically. Indeed, gestures in the face-to-face dialogues conveyed more categories at once ($M = 3.13$ categories; $SD = 0.54$) than gestures in the telephone dialogues did ($M = 1.04$ categories; $SD = 0.70$). This difference was statistically significant; $t_{(16)} = 7.14$; $p < .001$. In other words, although participants gestured at the same rate when speaking on the telephone, these gestures were not as informative as the gestures the participants used when speaking face to face. Furthermore, it seemed that the gestures in the telephone condition might be less informative because they depicted fewer semantic features at once.

Effect of conversational context on the words. Although speakers did not use more words in the telephone condition, for four of the semantic categories (the skirt's identity, the waist, the skirt's size, and its flat shape), the speakers' words conveyed significantly more information than speakers' words had in the face-to-face condition (see Table 5). Speakers were often quite efficient in their verbal contributions on the telephone, especially when they used similes or metaphors to describe the skirt. For example, when the speaker in our example compared the dress to a table, she conveyed information about all five categories at once, or when another said, "it's kinda like in a big rectangle shape", she conveyed three of the categories. Note that Bavelas et al. (2008) conducted an analysis of figurative language on the larger data set from which we selected the excerpts used here.

Table 5. Words: Means and Standard Deviations of Proportions of Contributed Information in the Telephone vs. Face-to-Face Conditions

Semantic Category	Conversation Context	M (SD)	$t_{(df)}$
Identity of skirt	Telephone	.89 (.20)	$t_{(18)} = -4.90^{**}$
	Face-to-face	.34 (.30)	
Waist	Telephone	.60 (.46)	$t_{(11.79)} = -2.19^{*}$
	Face-to-face	.26 (.18)	
Flat shape	Telephone	.49 (.35)	$t_{(12.64)} = -1.97^{*}$
	Face-to-face	.25 (.16)	
Size	Telephone	.93 (.12)	$t_{(18)} = -11.43^{**}$
	Face-to-face	.27 (.14)	
Shape of corners	Telephone	.17 (.27)	*n.s.*
	Face-to-face	.17 (.22)	

Note. Tested with independent-samples *t*-tests
* $p < .05$; ** $p < .001$.

These authors found that speakers on the telephone used figurative expressions significantly more often than speakers in face-to-face dialogue.

Discussion

One central test of whether gestures serve a communicative function is to vary visibility between participants and observe changes in gesturing. Many such visibility studies have compared the rate of gestures between conditions. We have proposed, however, that rate measures do not capture the communicative nature of gestures because they neither take into account qualitative features of gestures nor reflect the integration between gestures and words. We set out to test the communicative hypothesis by varying visibility between participants and by using a semantic feature approach to analyze excerpts from speakers' descriptions of a dress. We focused our analysis on excerpts during which the speakers described the shape of the skirt to an addressee.

In order to contrast the sensitivity of the semantic feature approach with traditional rate measures, we first compared gesture rate, number of words, and gesture duration in excerpts from the two conditions. We found no significant differences between conditions on any of these three variables. If our analysis had focused only on these measures, it would have appeared as though there were no differences between the telephone and face-to-face conditions. These results, on their own, would not have supported the hypothesis that gestures serve a communicative function.

The results of the semantic feature analysis provided a contrast to the rate dependent variables, which were apparently misleading. The semantic feature results strongly support the communicative hypothesis. In the face-to-face conversations, speakers, in general, contributed significantly more information in their gestures than they did in their words. These results held even when divided into the five separate semantic categories. In telephone conversations, speakers, in general, contributed significantly more information in their words than in their gestures. When divided into semantic categories, the analysis showed that this difference was driven primarily by significant differences with the identity of the skirt and its size. We also compared gestures in the two conditions and we found that those produced in face-to-face dialogues were significantly more informative than those produced on the telephone, even when the results were divided into the separate semantic categories. Furthermore, individual gestures in face-to-face dialogues were more "dense" with information, conveying significantly more categories of information at once than the gestures in the telephone dialogues did. Words were also predictably influenced by condition: Words in the telephone dialogues were significantly more informative than those in the face-to-face dialogues, and this difference held in four of the five semantic categories.

These findings matched our predictions and support the hypothesis that gestures serve a communicative function. Gestures and words form integrated messages (Bavelas & Chovil, 2000), and speakers are flexible in how they distribute information, being sensitive to the context of the conversation. If the addressee can see the gestures, speakers form gestures so that they carry information; if the addressee cannot see the gestures, even though speakers often gesture, the words convey the information.

Why did participants distribute the information they conveyed in these systematic ways? Bavelas and Chovil (2000), in their discussion as to the symbolic nature of hand gestures, drew a distinction between *analogic* and *digital* encoding. Gestures are analogically or naturally encoded; that is, they resemble their referent. This quality of hand gestures makes them the ideal modality for expressing things that hands do (such as hand actions) as well as some aspects of shape, size, and the relationship between things. Participants in our study could use their hands to trace the outline of the skirt as though they were wearing the dress. The gestures they traced around their bodies bore an analogic relationship to the actual shape of the skirt. Words, on the other hand, are digitally encoded: They are not like what they represent, and they bear an arbitrary relationship to their referents. Words can express many things that gestures cannot, and with a shared lexicon and syntax, human expressiveness is limitless. In the face-to-face condition, participants could encode all aspects of the skirt easily into gestures, but when they were speaking on the telephone, they had to (and did) rely on their words. We pro-

pose that participants' choice of what to use in any given situation is driven simply by what is quickest and easiest at that moment. In other words, we agree with Clark (1996), who said that speakers are *opportunistic* in how they use language. When gestures are available to convey those aspects of the referent that could be encoded analogically, the gestures, combined with words, are the most efficient means to achieve mutual understanding. If gestures are not available, as with the telephone condition, then speakers no longer have them as an available resource, and they shift the burden of information to words. The systematicity in our data was due the demands of the conversational setting and the ease of encoding supplied by words or gestures.

Findings in relation to past research

As mentioned earlier, the videotaped material we used for this study is a subset from a study reported elsewhere (Bavelas et al., 2008). Not surprisingly, our subset had the same results on rate measures; in both analyses, there was no significant difference in gesture rate or number of words in the face-to-face and telephone conditions. (Note that Bavelas et al. did not conduct an analysis of gesture duration.) In addition, these researchers conducted an analysis of gesture-word redundancy on the same subset of excerpts analyzed in the present study, and they found that speakers' gestures in the telephone condition were redundant with the words significantly more often than in the face-to-face condition. However, the average proportion of non-redundant gestures in the telephone condition was .42. In other words, almost half the gestures in this condition contributed information that was not conveyed in the words.

We propose that this surprisingly high proportion was due to the analysis procedures. First, when assessing redundancy, Bavelas et al. compared gestures only to the directly accompanying words; analysts did not consider the semantic contribution of words that were not accompanied by gestures. In our analysis, we analyzed all words, whether they were accompanied by gestures or not. For example, one participant on the telephone said, "I guess, the first thing that I notice is, right at, the BOTTOM part of the dress, <u>SHOOTS out, quite a bit on both sides</u>. Like WAY beyond, that of her shoulders." During the underlined words, she did a gesture; all of the other words were unaccompanied by gestures. From this example, one can see that the words without gestures conveyed information about the identity of the skirt ("the bottom part of the dress") and the size ("like way beyond that of her shoulders"), neither of which would have been considered in the redundancy analysis. We propose that a redundancy analysis underestimates the contribution of words in the telephone condition (Gerwing & Allison, 2009). Furthermore, the definition for a *non-redundant gesture* in Bavelas et al. was that

the gesture contribute any kind of information that was not in the accompanying words. Given how overly-inclusive this definition is, it is not surprising that analysts could rate a gesture as conveying unique information, likely resulting in the contribution of gestures being overestimated (Gerwing & Allison, 2009).

When comparing the gestures in the two conditions, we focused on gesture features that might make them more communicative. Specifically, we looked for handshapes that depicted the five categories of information, either singly or in combination. Whether, and indeed how, the gestures depicted these categories differed systematically in the two conditions, indicating that speakers were aware of the contextual variable provided by the addressee (who could either see those gestures or not). While our literature review focused on visibility studies, it is interesting to note that other manipulations of social context have yielded similar results. For example, gesture form was influenced systematically by the seating arrangement between speaker and addressee (Özyürek, 2000, 2002) or the common ground that they shared (Gerwing & Bavelas, 2004; Holler & Stevens, 2007).

We chose a narrow focus for our analysis, but this is a common approach, for example, in studies using referential communication tasks (e.g., Clark & Wilkes-Gibbs, 1986; Holler & Wilkin, 2011; Schober & Clark, 1989). Tasks in which one participant has all the information and must convey it to an unknowing addressee are analogous to everyday conversations, such as describing how to do something or get someplace. Furthermore, a semantic feature analysis requires a systematic, close examination of words and gestures, comparing each with their referent. It therefore necessarily sacrifices breadth for depth (Gerwing & Allison, 2009).

Our data were from dialogues where participants were able to interact freely, and other than fulfilling the demands of the task (describing the dress), the participants in these conversations were not restricted in any way. Future research in this area would benefit a great deal from explorations of how participants use gestures to collaborate in dialogue (e.g., Clark & Krych, 2004; Furuyama, 2000), including an analysis of how they use gestures and incorporate them into the immediate organization of their discourse to establish mutual understanding. While Holler and Wilkin (2011) have contributed findings in regard to the role of gestural mimicry, there appears to be a paucity of research on how participants integrate gestures and words to establish mutual understanding. This facet of communication requires reciprocity and spontaneity and has likely been greatly impeded by the use of confederates as addressees (see discussions in Bavelas & Gerwing, 2007; Holler & Wilkin, 2011).

In conclusion, on one level, all speakers had the same task: Describe the dress they were observing on the placard. All participants were free to gesture and speak spontaneously, improvising how to describe what was arguably difficult to put into words. If the gestures we analyzed functioned merely to help the speakers find

words or expressions that were not coming easily, there is no reason why those gestures should have differed in their semantic content in the face-to-face vs. telephone conversations. Nor are there any reasons why the semantic content of the words should have differed between the two conditions. While we do not interpret our results as meaning that the gestures served an exclusively communicative function, we do believe that the results provide strong support for speakers' systematic and intentional attention not only to the information they are conveying in words and in gestures, but also to what the addressee will be able to perceive and interpret.

References

Alibali, Martha W., Dana C. Heath, & Heather J. Myers (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, 44, 169–188.

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

Bavelas, Janet B. & Nicole Chovil (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, 19 (2), 163–194.

Bavelas, Janet B. & Nicole Chovil (2006). Hand gestures and facial displays as part of language use in face-to-face dialogue. In V. Manusov & M. L. Patterson (Eds.), *Handbook of nonverbal communication* (pp. 97–115). Thousand Oaks, CA: Sage.

Bavelas, Janet B., Nicole Chovil, Douglas A. Lawrie, & Allan Wade (1992). Interactive gestures. *Discourse Processes*, 15, 469–489.

Bavelas, Janet B. & Jennifer Gerwing (2007). Conversational hand gestures and facial displays in face-to-face dialogue. In K. Fiedler (Ed.), *Social communication* (pp. 283–308). New York: Psychology Press.

Bavelas, Janet B., Jennifer Gerwing, Chantelle Sutton, & Danielle Prevost (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58, 495–520.

Beattie, Geoffrey, & Heather Shovelton (1999). Do iconic hand gestures really contribute anything to the semantic information conveyed by speech? An experimental investigation. *Semiotica*, 123, 1–30.

Beattie, Geoffrey & Heather Shovelton (2002). An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology*, 93, 179–192.

Blum, Stella (1982). *Eighteenth-century French fashion plates in full color: 64 engravings from the "Galeries des Modes", 1778–1787*. New York: Dover.

Clark, Herbert H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, Herbert H. & Meredyth A. Krych (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50, 62–81.

Clark, Herbert H. & Deanna Wilkes-Gibbs (1986). Referring as a collaborative process. *Cognition*, 22, 1–39.

Cohen, Akiba A. (1977). The communicative functions of hand illustrators. *Journal of Communication*, 27, 54–63.

Cohen, Akiba A. & Randall P. Harrison (1973). Intentionality in the use of hand illustrators in face-to-face communication situations. *Journal of Personality and Social Psychology*, 28, 276–279.

Emmorey, Karen & Shannon Casey (2001). Gesture, thought and spatial language. *Gesture*, 1 (1), 35–50.

Furuyama, Nobuhiro (2000). Gestural interaction between the instructor and the learner in origami instruction. In David McNeill (Ed.), *Language and gesture* (pp. 99–117). Cambridge: Cambridge University Press.

Gerwing, Jennifer & Meredith Allison (2009). The relationship between verbal and gestural contributions in conversation: A comparison of three methods. *Gesture*, 9 (3), 312–336.

Gerwing, Jennifer, & Janet B. Bavelas (2004). Linguistic influences on gesture's form. *Gesture*, 4 (2), 157–195.

Gullberg, Marianne (2006). Handling discourse: Gestures, reference tracking, and communication strategies in early L2. *Language Learning*, 56, 155–196.

Holler, Judith & Geoffrey Beattie (2002). A micro-analytic investigation of how iconic gestures and speech represent core semantic features in talk. *Semiotica*, 142, 31–69.

Holler, Judith & Geoffrey Beattie (2003). How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process? *Semiotica*, 146, 81–116.

Holler, Judith & Geoffrey Beattie (2004). The interaction of iconic gesture and speech in talk. In Antonio Camurri & Gualtiero Volpe (Eds.), *Gesture-based communication in human-computer interaction* (pp. 63–69). Berlin: Springer.

Holler, Judith & Rachel Stevens (2007). The effect of common ground on how speakers use gesture and speech to represent size information. *Journal of Language and Social Psychology*, 26 (1), 4–27.

Holler, Judith & Katie Wilkin (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24, 267–289.

Holler, Judith & Katie Wilkin (2011). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, 35, 133–153.

Kendon, Adam (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.

Kimbara, Irene (2008). Gesture form convergence in joint description. *Journal of Nonverbal Behavior*, 32, 123–131.

Krauss, Robert M., Robert A. Dushay, Yinsiu Chen, & Frances Rauscher (1995). The communicative value of conversational hand gestures. *Journal of Experimental Social Psychology*, 31 (6), 533–552.

Lickiss, Karen P. & A. Rodney Wellens (1978). Effects of visual accessibility and hand restraint on fluency of gesticulator and effectiveness of message. *Perceptual and Motor Skills*, 46, 925–926.

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

Özyürek, Aslı (2000). The influence of addressee location on spatial language and representational gestures of direction. In David McNeill (Ed.), *Language and gesture* (pp. 64–83). Cambridge: Cambridge University Press.

Özyürek, Aslı (2002). Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory and Language*, 46, 688–704.

Rimé, Bernard (1982). The elimination of visible behavior from social interactions: Effects on verbal, nonverbal and interpersonal variables. *European Journal of Social Psychology*, 12, 113–129.

Schober, Michael F. & Herbert H. Clark (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, 21, 211–232.

Authors' addresses

Jennifer Gerwing
Department of Psychology
University of Victoria
P.O. Box 3050
Victoria, B.C., V8W 3P5
Canada

jjg@uvic.ca

Meredith Allison
Department of Psychology
Campus Box 2337
Elon University
Elon, N.C., 27244-2010
USA

mallison5@elon.edu

About the authors

Jennifer Gerwing, Ph.D., is an Adjunct Assistant Professor of Psychology at the University of Victoria. Her previous gesture research included the effect of conversational context (common ground or visibility between participants) on gesture's rate and form. Her doctoral work was on social reciprocity in parent-infant interaction, investigating differences in social responsiveness as displayed in home videos of autistic vs. non-autistic infants. She is currently working as a researcher at the Norwegian Centre for Minority Health Research focusing on health care communication in dialogues between health care professionals and non-native speakers (i.e., immigrant patients).

Meredith Allison, Ph.D., is an Assistant Professor of Psychology at Elon University. She received her doctorate in psychology at the University of Victoria. Her previous gesture research was on the co-construction of virtual spaces and the use of interactive vs. topic gestures. Much of her research is in the psychology and law domain where she has studied the use of narrative features in the testimonies of older vs. young adult eyewitnesses and in people's alibis.