Given-New Effects on the Duration of Gestures and of Words in Face-to-Face Dialogue

Judith Holler, Janet Bavelas, Jonathan Woods, Mareike Geiger & Lauren Simons

Published online: 05 Sep 2022.

Submit your article to this journal

Article views: 649

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Routledge
Taylor & Francis Group

Given-New Effects on the Duration of Gestures and of Words in Face-to-Face Dialogue

Judith Holler[a,b], Janet Bavelas[c], Jonathan Woods[c*], Mareike Geiger[a,b], and Lauren Simons[d]

[a]Donders Institute for Brain, Cognition & Behaviour, Radboud University; [b]Max Planck Institute for Psycholinguistics; [c]Psychology Department, University of Victoria; [d]Centre for Language Studies, Radboud University

ABSTRACT

The *given-new contract* entails that speakers must distinguish for their addressee whether references are new or already part of their dialogue. Past research had found that, in a monologue to a listener, speakers shortened repeated words. However, the notion of the given-new contract is inherently dialogic, with an addressee and the availability of co-speech gestures. Here, two face-to-face dialogue experiments tested whether gesture duration also follows the given-new contract. In Experiment 1, four experimental sequences confirmed that when speakers repeated their gestures, they shortened the duration significantly. Experiment 2 replicated the effect with spontaneous gestures in a different task. This experiment also extended earlier results with words, confirming that speakers shortened their repeated words significantly in a multimodal dialogue setting, the basic form of language use. Because words and gestures were not necessarily redundant, these results offer another instance in which gestures and words independently serve pragmatic requirements of dialogue.

During a conversation, interlocutors introduce information that is new to their present dialogue. When they later refer to the same information, it becomes *given* information, that is, information "already supplied by the previous linguistic context" (Crystal, 2001, p. 135). Signaling when information is new and when it is given is vital for successful communication. Chafe (1974) proposed that, during the course of a conversation,

> Virtually every sentence a speaker utters is a mixture of what, following Halliday (1967), I will call *given* material, which the speaker assumes is already in the addressee's consciousness, and *new* material, which he assumes is not. As he converts this mixture into sound, the speaker does not treat the given and new material in the same way: typically, he will attenuate the given material in one way or another, e.g., by pronouncing the items that convey such material with lower pitch and weaker stress, or by the attenuated specification or pronominalization of such items. (p. 112)

Clark and Haviland (1977) took a stronger, interactional position. Their *given-new contract* proposes that speakers are obliged to ensure that their addressee can distinguish what is given from what is new. In order to do so, the given version must resemble the new version but must also differ in a recognizable way; for example, by attenuating pronunciation, pitch, or stress.

Example 1 (from data in Bavelas et al., 2014).

A speaker was just beginning to describe a sequence of shapes in a drawing that the addressee could not see. The first shape looked like the letter *S*.

> Where it starts, we have (pause) uh (pause) AN ESS [looks up; addressee nods], ess-shape there. [Addressee: Mhm.]

The first reference ("AN ESS") was new information. The speaker's articulation indicated in two different ways that "ess-shape" was a repetition, rather than a second shape, and therefore given information. First, he had stressed "AN ESS," but "ess- shape" was soft and unstressed. Also, "AN ESS" was 0.61 s long while "ess-shape" was 0.4 s long. The addressee's responses indicated that she understood both.

Some necessary distinctions

Given-new effects can be confounded with common ground. It is therefore essential to distinguish between common ground that exists *before* the dialogue and common ground that is created *during* the dialogue. Clark (1996) described common ground as deriving from three domains:

(1) *Communal common ground* (pp. 100–112) is the knowledge shared in cultural or sub-cultural communities prior to the dialogue.

(2) *Personal common ground* (pp. 112–116) is also prior to the dialogue but more specific; it is created when a particular speaker and addressee share a previous experience or current situation.

(3) *Incremental common ground* (pp. 38–39, 221–251) is created during the interaction, in the interlocutors' dialogue. It is the topic of this article.

Often, "having common ground" seems to refer to prior personal common ground, that is, to the cognitive states of the interlocutors before they interact, when they know that they share some of the same knowledge, beliefs, or experiences (Clark, 1996). In contrast, incremental common ground does not exist before the conversation. It develops within a dialogue, is specific to that particular dialogue, and is the direct result of linguistic and physical co-presence. The accumulation of incremental common ground is a social, interactive process whereby interlocutors build up a joint track record of information as it is being exchanged (Clark, 1996; Clark & Marshall, 1978, 1981) and mutually confirmed through the process of grounding during their dialogue (Clark & Brennan, 1991). It follows that the development of incremental common ground is directly observable in the communicative processes that create it, especially in transitions from new to given forms. So, a spontaneous dialogue is the first requirement for the study of incremental common ground.

The most common environment for dialogue is face-to-face interaction, where manual gestures form an integral part of human communication (e.g., Bavelas, 2022; Enfield, 2009; Goldin-Meadow, 2003; Kendon, 2004; McNeill, 1992; Streeck, 2009). In the gesture literature, Gerwing (2003; published as Gerwing & Bavelas, 2004) distinguished both theoretically and empirically between prior and incremental common ground. Typical experiments on prior common ground manipulate the interlocutors' shared knowledge before their interaction by presenting the same vs. different information to speakers and addressees and then observing the effects on gestures. These effects are immediately apparent in the first reference to the mutually shared information (e.g., Gerwing, 2003; Gerwing & Bavelas, 2004, Analysis 1; Holler & Stevens, 2007; Holler & Wilkin, 2009). The usual common ground study is a between-subjects design in which the same or different information is randomly assigned to the dyads prior to their dialogue. The analysis typically focuses on the effect of common ground on gesture frequency, gesture rate, or gesture form (such as size, space, or precision). These measurements are usually aggregated over all gestures in a dialogue (or per speaker) rather than matching gestures for the same referent (see Holler & Bavelas, 2017, for an overview).

Studies of incremental common ground require a different research design and analysis. They must use a within-subjects design in which the speaker repeats some references to the addressee. The analysis is a precise word-by-word or gesture-by-gesture comparison of the new and given forms of each reference. If speakers are modifying particular gestures as they become given, then the analysis must focus on the contrast that this modification is creating for the addressee. The focus of such an analysis is the contrast in form of speakers' new and given gestures.

Following Gerwing (2003; Gerwing & Bavelas, 2004, Analysis 2), we propose that sequential gestures for the same referent within a dialogue also follow the given-new contract; for example,

Example 2 (from data in Bavelas et al., 1995).

The dyad's task was to watch a "Road Runner" cartoon with a particularly long and complex series of events, then to retell it together. As one speaker picked up where the other had left off, she said,

> Alright. Ok. So the watering can tips o–so I don't know if we can remember all this [while laughing together]–the watering can tips over, and it . . .

At the first underlined words, she gestured water pouring out of the tipped watering can like a waterfall. This gesture took 1.2 s. After she had interrupted herself and laughed about the task with her partner, she resumed and made the same gesture with the next underlined words. This gesture was an abbreviated version that took 0.48 s and matched fewer words, making it clear to the addressee that she was repeating where she had broken off, not describing a second watering can or a second tipping over.

In Examples 1 and 2, the speakers met the given-new contract by producing an attenuated form, thereby ensuring that their addressee could distinguish their second reference as given rather than new information. Indeed, a given-new effect can only be observed by comparing the new and given forms of the same information. Based on previous research with words, we chose to focus on a quantitative comparison, namely, gesture duration.

Previous research

Given-new effects on words

As Chafe (1974) pointed out, there are many ways to mark given status in words, such as lower pitch, weaker stress, or pronominalization. Research on narrative speech has shown that speakers use full lexical forms to (re-)introduce referents, and pronouns or zero anaphora for maintaining them (e.g., Hickmann & Hendriks, 1999; Marslen-Wilson et al., 1982). Speakers thus use fuller lexical forms for new referents, which would be less accessible for the listener, and less full forms for given referents, which would be more accessible (Arnold et al., 2013; Givón, 1983).

In a classic study, Fowler and Housum (1987, Exp. 1) proposed that *shortening* could mark a word as given. They measured the duration in milliseconds of new and repeated words in a story by a radio personality and in television interviews with public figures. They found a significant decline in duration from new to given words over the course of the story or interview. Although Fowler and Housum (p. 489) described their data as monologues, Fowler (1988) hypothesized that "the reductions *help indicate to listeners* that the reduced words refer back to earlier-presented information in the discourse" (p. 308; italics added). In three later experiments, Fowler (1988) found (i) no significant shortening when reading a list of words, (ii) some shortening when reading meaningful prose, and (iii) significant shortening when talking spontaneously to the experimenter. So mere repetition could not account for the given-new effect.

Albeit without citing Clark and Haviland (1977), both of the Fowler articles recognized shortening as a social effect. For example, Fowler and Housum (1987) "found that listeners can identify words as old or new, and they can use information that a word is old to facilitate integration of related material in a discourse." (p. 501). Fowler (1988) proposed further that

> The talker may produce shortening for the listener's benefit, and so shortening may be more likely to occur when there is a listener present signaling his or her understanding of the talker's utterance . . . (Fowler, 1988, p. 314)

The experiments in this article built on Fowler's work by investigating true face-to-face dialogues. They also address the multimodal nature of face-to-face dialogue by investigating given-new effects for gestures as well as words.

Given-new effects on gestures

Co-speech gestures contribute to reference-tracking in narrative by providing visuo-spatial anchors that help create referential cohesion. For example, many studies showed that (re-) introductions of referents tended to be accompanied with gestures whereas this was less typical for maintained

references (e.g., Gullberg, 2003, 2006; Levy & McNeill, 1992; Levy & Fowler, 2000; McNeill & Levy, 1993; Perniss & Özyürek, 2015; Yoshioka, 2008). However, the picture is more complex than this, because gestures accompanying references in narrative tend to differ in their form depending on the accessibility of the referent and interactions with other variables, such as definiteness of the referential expression, semantic focus of the gesture, and the precise temporal alignment and semantic overlap between gesture and referential expression (see Debreslioska & Gullberg, 2019, 2020a, 2020b; Debreslioska et al., 2013; Foraker, 2011; Wilkin & Holler, 2011). Significantly, Levy and McNeill (1992) hypothesized that "the accumulation of other surface devices, such as increased duration of articulation [for the first presentation] (Fowler & Housum, 1987; Fowler & Levy, 1991) should contribute to this process as well" (p. 298, citations in original).

Gerwing (2003, Gerwing & Bavelas, 2004) was the first to obtain direct evidence of changes in gesture form due to prior common ground (Analysis 1) and to incremental common ground (Analysis 2) in spontaneous face-to-face dialogues. The focus of Analysis 2 (pp. 171–182 & Appendix A) was explicitly the gestural depiction of given versus new information. First, the (adult) interlocutors individually saw and manipulated two toys (e.g., a whirligig and a finger cuff). When they met, knowing that they had had different toys, they were to discuss briefly what they "did with the toys." They were free to talk as they wished. So a speaker's information started as new, then became given over the course of their dialogue. Typically, in their first presentation,

> [Speakers] often exaggerated the important features (e.g., made it larger than life), made it very precise, or drew attention to it with an extra movement. Sometimes they abstracted the salient feature from the whole so that the feature stood out on its own. (p. 175)

In contrast, the later, given gestures for the same action confirmed the hypothesis that

> given information, because it could draw on antecedents in previous gestures, would appear as transformed versions of those gestures. Aspects of the gestures depicting given information would be smaller or less precise versions of previous gestures. (p. 176)

A detailed grid analysis (pp. 181–195) for each of the gestures of the 10 dyads showed how the speakers successively transformed the given gestures.

Here we built on the above qualitative study with a quantitative measure, gesture duration. The experiments were also informed by the Fowler studies on the duration of words in monologues, extended to gestures and words in face-to-face dialogue. Thus, rather than looking at the accessibility of referents in narrative and resulting changes in referential form, we focused on the duration of gestures (and words) when they were repeated in dialogic interaction.

Methodological issues

There have been several gesture experiments in which a speaker presented the same information to the same addressee more than once but that were not about the given-new contract (Jacobs & Garnham, 2007 [conditions 1 & 2]; Holler et al., 2011; Holler & Wilkin, 2011a; De Ruiter et al., 2012; Galati & Brennan, 2014; Alibali, Nathan, Wolfgram, Breckinridge Church, Jacobs, Johnson Martinez & Knuth, 2014; Hoetjes et al., 2015; Vajrabhaya & Pederson, 2018). None was designed or intended to test given-new effects. Some were studying repetition or repeated reference for its own sake. Others used repetition to the same vs. a different addressee as a way of testing whether gestures are communicative or part of speaker-oriented speech production processes (i.e., for the speaker or for the addressee).

Those research goals are interesting and valuable in their own right, but they illustrate methodological choices that could not (and were not intended to) test a given-new effect. For example, several studies precluded the spontaneous dialogue processes that create incremental common ground by using a between-groups design rather than making within-speaker comparisons. Some used a confederate or addressee instructed to be minimally responsive, which may prevent or significantly alter the spontaneous grounding process necessary for creating incremental common ground and can produce different results (Bavelas, 2022; Bavelas & Healing, 2013; Kuhlen & Brennan, 2013). Also,

most of the studies used overall gesture rate, an aggregated measure that cannot assess the relationship between the first and later gestures for the same referent. Studies that did compare individual gestures did not necessarily compare gestures referring to exactly the same semantic feature or referent. Instead, the criterion was that the first and later gestures only had to refer to the same broader stimulus event, each of which tended to elicit multiple gestures. Although these first and second gestures would have some semantic connection, speakers would not have treated the second gesture as depicting given information if it did not depict exactly the same referent within the stimulus event.

These are not criticisms of the above studies ; our point is simply that they do not present evidence for or against given-new effects because they did not intend to do so. Whether the given-new contract specifically affects the duration of repeated gestures, as has been shown for words, remained an open question, addressed in the two experimental studies presented here. Specifically, we investigated given-new effects in face-to-face dialogues with a focus on two core domains of human communication: instruction (that is, demonstrations, as in teaching) (Experiment 1, based on Woods, 2005) and referring (Experiment 2). Experiment 1 focused on gesture only, Experiment 2 analyzed both gestures and words.

Design and predictions

The key criteria shared by the two experiments were (a) two participants in a spontaneous face-to-face dialogue, visible to each other at all times and (b) tasks that were likely to elicit repeated gestures. The only specific instructions were the role assignments within the dyads and the goals of their interaction, both of which required a speaker to describe easily gestured material to their addressee several times. Both experiments used a within-participants design and compared the duration of first and later gestures for the same referent. All participants were fluent English speakers.

There are also several differences that contribute to generalizability (see Table 1 for a complete overview): In Experiment 1, the speakers were teaching their addressee dance steps enacted with their hands on the table between them. That is, the referent was a prescribed sequence of motions that reoccurred over several trials. In Experiment 2, participants were doing the Tangram task (e.g., Clark & Wilkes-Gibbs, 1986), which requires a Director and Matcher to align on referring expressions for static geometric figures over several trials. The dance steps involved several prescribed gestures while the Tangram figures could be depicted by a single, improvised gesture (or a combination of gestures) of the Director's choice. Also, the dance steps had to be enacted with gestures, while the Tangram figures have been successfully described entirely in words (as in the original Clark & Wilkes-Gibbs experiment with a partition between Director and Matcher). Face-to-face versions of the Tangram task (Holler & Wilkin, 2011a; De Ruiter et al., 2012) have shown that participants do use gestures; the question is whether these gestures would serve given-new functions when words alone could do so.

Table 1. Differences between Experiments 1 and 2.

Experiment 1 (Dance steps)	Experiment 2 (Tangrams)
Conducted in Canada.	Conducted in the U.K.
Participants were Psychology students.	Participants had diverse backgrounds.
Task: Teaching six dance steps in prescribed sequences.	Task: Referential communication; describing 12 figures.
Four assigned within-subject variations in these sequences.	Using gesture for any given figure at any given time was spontaneous.
Gestures were required; words were not well suited.	Both words and gestures would suit the task, but it could be completed with words alone.
Participants were sitting, using hand gestures.	Participants were standing, using hand and bodily gestures.
There was a practice phase.	There was no practice phase.
Instruction to use hands as feet to gesture.	No instruction to gesture (or how).
Preparatory phases were not meaningful, hence not included in duration measure.	Preparatory phases were often informative and included in the duration measure.
Gesture units were defined by the point at which understanding had been reached. Therefore, the gesture units measured for duration always included multiple sequential gesture strokes.	Gesture units consisted of individual gesture strokes or groups of strokes (i.e., several strokes with temporal overlap).

Table 2. Occurrences of the dance steps within and across trials.

Dance Step Trial number	Trial Number					
	1	2	3	4	5	6
	Position within the Trial					
Criss-Cross	1st	4th	4th	1st	7th	3rd
Heel-Toe	-	-	1st	3rd	3rd	7th
Forward-Back-Side	3rd	1st	6th	6th	5th	1st
Kick-Up-Kick-Back	6th	-	-	-	1st	5th
Clap, Arms Out	2nd	2nd	5th	2nd	6th	6th
Clap	5th	5th	2nd	5th	2nd	2nd
Turn Right 90°	4th	-	3rd	4th	-	-
Turn Left 90°	-	3rd	-	-	4th	4th

The first four dance steps in the table are the target steps; the others were fillers. Dashes indicate that the step was not included in that dance trial.

Predictions

(1) Experiments 1 and 2 would replicate the given-new effects on gestures in Gerwing (2003; Gerwing & Bavelas, 2004, Analysis 2), this time using the duration measure that Fowler and Housum (1987) had used for words. The results would provide generalizable, quantitative evidence that speakers in face-to-face dialogs use the duration of their gestures to mark given versus new information, independently of whether they can also do so in words.

(2) Experiment 2 would also replicate Fowler's (1988) given-new effects on words, this time in a spontaneous face-to-face dialogue.

Experiment 1. The dance steps

Woods (2005) collected data from dyads engaged in a dance step task that formed the basis of his study. Each dyad comprised a Teacher and a Learner. The Teachers' task was to teach their Learner how to use hand gestures to perform six different dance sequences over six experimental trials, while sitting across a table from each other. The dance sequences were made up of combinations from a set of eight dance steps, with each trial portraying a different sequence of the dance steps; see Table 2. To test the range of the given-new hypothesis, there were four within-subject experimental conditions in which four of the dance steps repeated in different sequences across trials. That is, two of the four target steps occurred in all six trials, which tested the primary hypothesis that steps would be shorter after their first appearance. We varied the inclusion of the other two target steps to investigate the robustness of the given-new effect in different sequences: One dance step occurred only in trials 3 through 6 (i.e., it was introduced relatively later in the interaction). The other occurred in trial 1, then not again until trials 5 and 6, which tested the effect of a delay between the new and given information. The dependent measure was the duration of a Teacher's gestures for each the four target dance steps.

The given-new contract predicts that gesture duration for the initial (i.e., new) presentation of a dance step would be significantly longer than for subsequent (i.e., given) presentations of the same step in later trials, regardless of where the new and given steps occurred within the six trials. The other four steps were fillers, easily described verbally and not relevant to the experiment.

step was not included in that dance trial.

Method

Participants

Thirty-four students in a first-year psychology course at the University of Victoria (Canada) participated in return for two bonus credits (1% toward their course grade). The 29 women and five men formed 17 dyads (13 female-female, one male-male, and three mixed-gender dyads) with random

assignment to the roles of Teacher and Learner. All participants were unacquainted, in their late teens or early twenties, proficient in English, and had normal vision and hearing. Their participation in this experiment was approved by the University of Victoria Human Research Ethics Board.

Equipment

The Human Interaction Laboratory in the Psychology Department had four remotely controlled, tightly synchronized Panasonic WV-CP474 color cameras and two special effects generators (a Panasonic WJ-5500B overlaid on a customized Panasonic/MicroImage Video Systems). Two opposing cameras recorded a front view of each participant, and a third camera recorded a side view of both participants; see Figure 1. The data were digitized and analyzed on a 14-inch color monitor using Broadway (https://www.omegamultimedia.com/Data_Translation_BroadwaynPro_DVD.html), a commercial viewing and editing software program that was our preference at the time. Figure 2 shows the Broadway display, including the measurement precision to hundredths of seconds.

Materials

Participants received an Initial Consent form, which included permission to be video-recorded, as well as a post-experimental Permission to View form, which asked them to choose or reject various levels of permission to view the data (e.g., to be viewed only by the analysts, played at a professional conference, or reproduced as a still photo in a journal article). There were also two kinds of schematic drawings for the Teachers. Figure 3 shows the Dance Steps drawing, which depicted each of the steps except for two "filler" steps that could easily be described verbally and were therefore not of interest in the analysis. During the experiment, the Teachers also received an Instructional Dance Sheet for each trial. These used the same drawings to show the prescribed order of the dance steps in that trial. The sheets stood in an upright stand facing the Teacher and were not visible to the Learner.

Procedure

Practice phase
After a brief getting-acquainted conversation, random assignment determined the roles of Teacher (hereafter referred to as female) and Learner (hereafter referred to as male). The Teacher had an initial practice

Figure 1. Video of three-camera split screen. The Teacher (in white shirt) is demonstrating the crisscross step.
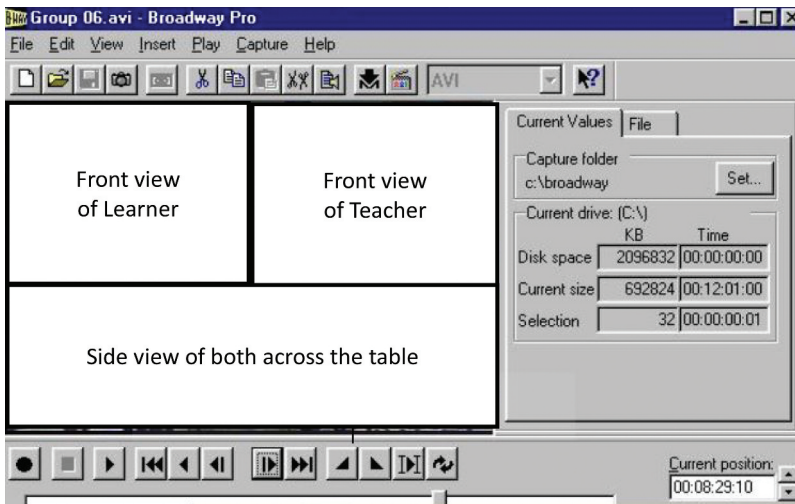
Figure 2. Broadway analysis system.

phase in the absence of the Learner, during which the experimenter stood beside her and showed her how to perform the six steps by actually doing each of the steps. Then she learned how to match each step with its corresponding schematic image on the Instructional Dance Sheets (e.g., Figure 3) and how to combine a sequence of steps into a dance. The training dances used trials 1, 3, and 6, which together covered all six steps. The experimenter often went over specific steps within each trial more than once when the Teacher was having difficulty mastering one of them. Then the Teacher and experimenter sat across from one another at a table as the Teacher demonstrated Trials 1 and 6 (counterbalanced across dyads). While seated, the Teacher could now only demonstrate a dance step by combining her words with hand gestures representing feet. This practice phase usually lasted less than 10 minutes and was complete when the Teacher could execute each step correctly and perform each of the training dance trials using the Instructional Dance Sheets without assistance from the experimenter.

Dialogue phase
The formal experiment began in the dialogue phase, when the Learner entered the room and sat across the table while the Teacher taught him Trials 1 through 6, each made up of a different combination of the dance steps (see Table 2). Both participants remained seated across from one another, so the Teacher could not stand or use her legs to demonstrate. There were no restrictions on the words or upper-body movements of either participant nor limits on how long they could spend practicing each dance step, just that they were not to move onto the next step until they mutually decided that the Learner had understood the current one. Teachers often needed to make several enactments of parts or the whole dance step within a trial before the Learner clearly understood. Therefore, the unit of analysis was the full grounding cycle for each time a dance step was presented in a trial, rather than any individual gesture within it. However, only the cumulative periods of gesturing of the Teacher within that unit was summed to establish the duration of Teacher's overall gesture duration for a dance step on a given trial; pauses or speech without gesture were not included. At the end of each trial, the Learner had to demonstrate his understanding of the entire trial while still remaining seated. When both parties felt that the Learner had sufficiently demonstrated his understanding of that trial, they continued on to the next one. This reflects the principle, cited earlier, that incremental common ground is mutually confirmed through the process of grounding during a dialogue (Clark & Brennan, 1991). We predicted that the new information would take longer to establish in the first trial and would then be shorter in later trials because it was given.
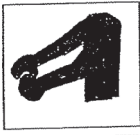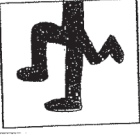
Dance Sheet | Name used by researchers
(Not used with the Teacher or Learner)

Criss-cross

Clap, arms out

Forward, back, slide

TURN RIGHT 90°   Turn right 90°

Clap

Kick up, kick back

Figure 3. Example of teacher's instructional dance sheet (Trial 1).

After the six trials, the experimenter explained the purpose of the experiment to the participants, showed them the video of their participation, answered their questions, and asked them to complete the Permission to View form individually.

Gesture analysis

In McNeill's (1992) system, the dance steps were iconic gestures. Recall that the unit of analysis was not each individual gesture but each trial, within which a dance step would be grounded to the pair's satisfaction. The focus was on the duration of a Teacher's gestures depicting each of the four target dance steps (Criss-Cross, Forward-Back-Side, Heel-Toe, and Kick-Up-Kick-Back) within each of the trials in which it appeared. This *gesture unit* (Kendon, 1980) often included several gesture strokes, until the Teacher and Learner had established that the Learner understood the dance step.

JW analyzed all of the data (323 presentations), and a second analyst (JG) viewed a random 10.5% of the data (34 presentations, two by each Teacher). The analysts made their decisions in three stages, assessed inter-analyst agreement for each stage, and resolved any disagreements for that stage before moving on to the next. The successive decisions were as follows:

(1) Whether the Teacher gestured during her presentation of the dance step; 100% agreement. Teachers gestured during 33 of the 34 presentations.

(2) Whether the gesture was analyzable (e.g., clearly visible and depicting the current dance step); 93.9% agreement. Thirty of the 33 gestures were analyzable.

(3) The duration of analyzable gestures during a presentation. This was guided by the software, which displayed the chosen beginnings and endings to hundredths of a second; agreement was $r = .995$. The duration did not include any initial preparatory phase (e.g., when the Teacher placed her hands on the table to start). Timing began when the Teacher indicated which foot to start with or at the first movement of the step. Timing ended when her hands stopped moving, but included any post-stroke holds up to two seconds as part of the gesture (this arbitrary cutoff point was applied to both new and given gestures):

> A post-stroke hold, that is, when a speaker sustains the articulator of a stroke in position after performing the stroke action . . . seems to be a way by which the expression conveyed by the stroke may be prolonged . . . . We shall refer to the phase of action that includes the *stroke* and any *post-stroke hold* as the *nucleus* of the gesture phrase. (Kendon, 2004, p. 112, italics in original)

As anticipated above, a Teacher might make more than one gesture for a dance step during a trial (separated by a period without gestures or by a gesture unrelated to the dance step). The analysts timed each of these by the same rules and combined them later. After assessing their agreement, the analysts discussed and resolved any discrepancies greater than one second in order to arrive at the final value.

Statistical analysis

The data were divided into two datasets: one subset with the two dance steps that occurred on all six trials (Criss-Cross and Forward-Back-Side) and one subset with the two dance steps that did not occur on all trials (Kick-Up-Kick-Back and Heel-Toe). Both datasets were analyzed using R (Version 4.0.4, R Core Team, 2020) and the lme4 package (Version 1.1.26; Bates et al., 2015). Both datasets were arranged by trial as well as by repetition. For the dance steps occurring on all six trials, these two data arrangements were very similar. Any small differences between the datasets were due to some participants not producing any (or any analyzable) gestures in a given grounding cycle (i.e. on an individual trial). Those 'no gesture' or 'na' cases were omitted when the data were arraned by repetition. Because the data arranged by trial resulted in a singular fit error, we analyzed the data by repetition rather than by trial. Moreover, analysis by repetition was the only way in which the data for the dance steps not occurring on all trials could be analysed, and it was most comparable to the analysis applied in Experiment 2. To test whether repetition had an effect on gesture duration, a generalized linear mixed model was fitted by maximum likelihood (Laplace approximation) with repetition as the fixed effect, and with eacher-Learner pair and dance step as random effects. The response family was inverse Gaussian with a logit link. In addition, a generalized linear mixed model was fitted with a matching structure to the first one, but with repetition treated as a categorical variable to explore the effect of the individual repetitions.

Teachers sometimes did not gesture for a dance step on a given trial. Such cases were entered as "no gesture." Other times, teachers gestured but the gestures were not analyzable. In such cases, the occurrence of gesture was entered but the duration value set to "NA."

Results

Duration of dance steps occurring on trials 1 through 6

Tables 3a and 3b and Figure 4 summarize the data by trial and by repetition. As laid out above, the statistical analyses were based on the data arranged by repetition. The general linear mixed model showed a significant effect of repetition on gesture duration ($\beta = -0.26$, SE = 0.02, t = −11.59, $p < .001$). Table 3a shows that gesture duration decreased from R0 to R5 (and the same pattern can be seen when looking at the data by trial, Figure 4 and Table 3b). The model comparing gesture duration for R0 with each following repetition showed that gesture duration for R0 was significantly longer than the duration of gestures for all subsequent repetitions (R0–R1: $\beta = -1.36$, SE = 0.2, t = −6.95, $p < .001$;
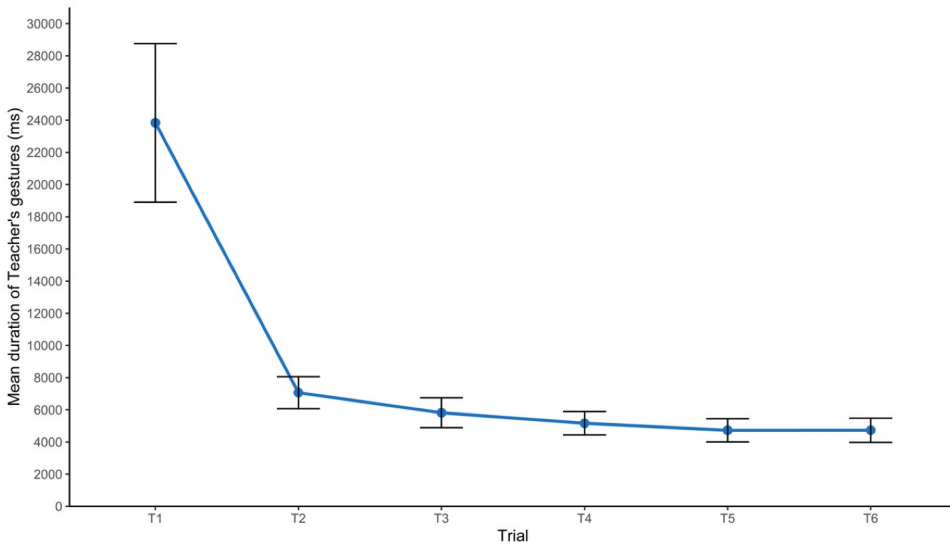
Figure 4. Duration of dance steps occurring on trials 1 through 6. Error bars represent 95% confidence intervals.

Table 3a. Mean duration (and SD) of dance steps repetitions occurring on all trials in milliseconds.

R0	R1	R2	R3	R4	R5
23835.76	7063.87	5647.58	5073.33	4710.94	4981.07
(13890.08)	(2712.68)	(2578.56)	(2125.51)	(2036.34)	(1838.01)

Table 3b. Mean duration (and SD) of dance steps occurring on trials 1 through 6 in milliseconds.

| T1	T2	T3	T4	T5	T6
23835.76	7063.87	5817.42	5163.94	4728.71	4731.29
(13890.08)	(2712.68)	(2526.02)	(2046.81)	(1965.44)	(2051.83)

R0–R2: $\beta = -1.65$, SE = 0.19, t = −8,57 $p < .001$; R0–R3: $\beta = -1.76$, SE = 0.19, t = −9.25, $p < .001$; R0–R4: $\beta = -1.82$, SE = 0.2, t = −9.33, $p < .001$; R0–R5: $\beta = 1.8$, SE = 0.2, t = −9.1, $p < .001$).

Duration of dance steps not occurring on all trials

Figure 5 and Tables 4a and 4b show that gesture duration decreased from the first time the gestures described a dance step to the last time the gestures described the same dance step, even when a dance step occurred for the first time on a later trial (Figure 5, panel A), and even when trials without the respective dance step happened in between (Figure 5, panel B). This underlines the robustness of the phenomenon in question.

To test for this pattern statistically, we arranged the data such that the first occurrences of each of the dance steps were compared to the second occurrences and the third occurrences, irrespective of the trial number (i.e. we aligned them by the number of repetition of a dance step's grounding cycle, rather than by trial, see Table 4a and the section on Statistical Analysis), to allow for sufficient data points at each measurement. The general linear mixed model confirmed the reduction in gesture duration as significant ($\beta = -0.33$, SE = 0.04, t = −8.55,

Panel A
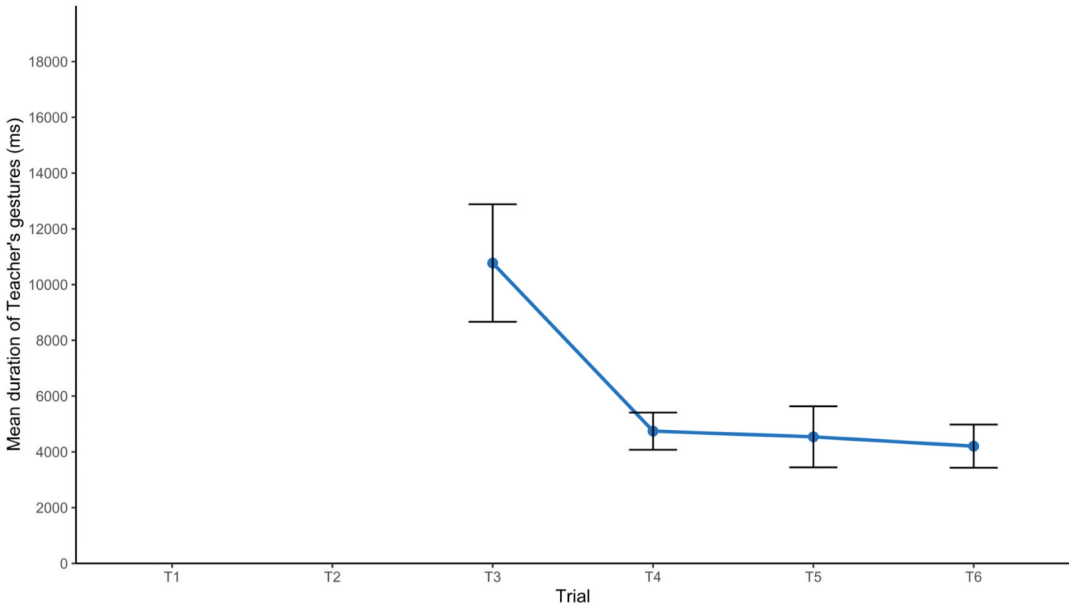
Panel B
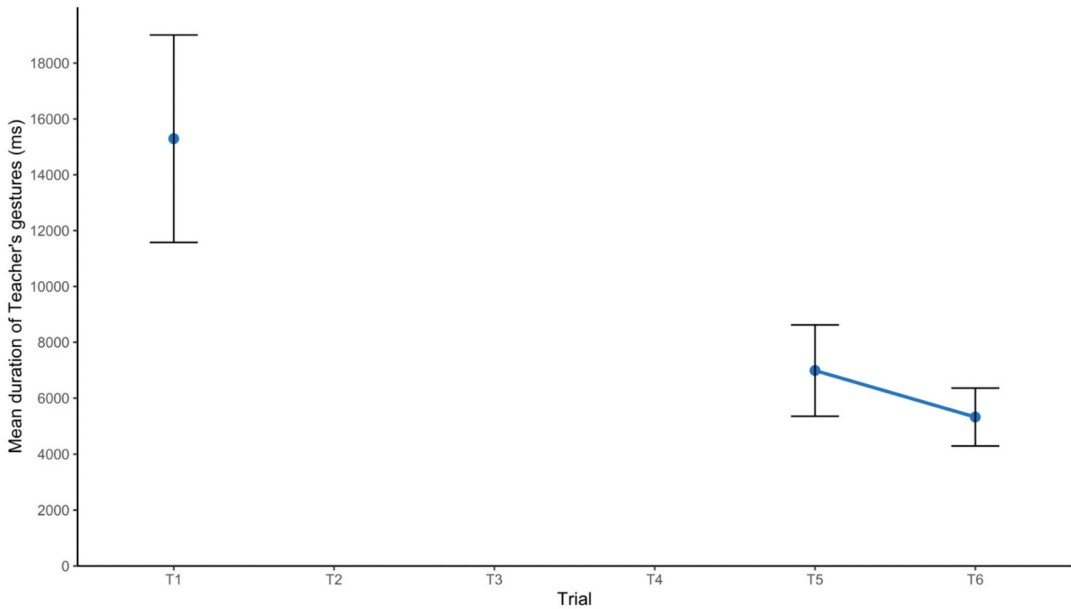
Figure 5. Duration of dance steps not occurring on all trials (Panel A: Heel-Toe, Panel B: Kick-Up-Kick-Back). Error bars represent 95% confidence intervals.

$p < .001$). Gesture duration was significantly longer on the first measurement than the duration on all repetitions (R0-R1: $\beta = -0.82$, SE = 0.11, t = -7.4, $p < .001$; R0-R2: $\beta = -0.99$, SE = 0.11, t = -8.94 $p < .001$; R0-R3: $\beta = -1.03$, SE = 0.13, t = -8.19, $p < .001$).

Table 4a. Mean duration (and SD) of repetitions of dance steps not occurring on all trials in milliseconds.

R0	R1	R2	R3
13032.35	5832.12	4944.19	4204.38
(6224.84)	(2557.52)	(1967.65)	(1450.47)

Table 4a. Mean duration (and SD) of dance steps not occurring on all trials in milliseconds.

T1	T2	T3	T4	T5	T6
Heel-Toe					
–	–	10772.35	4744.71	4539.33	4204.38
–	–	(4100.86)	(1298.13)	(1978.19)	(1450.47)
Kick-Up-Kick-Back					
15292.35	–	–	–	6987.5	5323.75
(7228.3)	–	–	–	(3064.71)	(1942.51)

Discussion

The results confirmed that the given gestures for the four different target dance steps were significantly shorter than their initial presentation. On average, the duration of dance steps that occurred on all trials, went from 24 seconds when new to 5 seconds when given. There was a similarly reduction for the dance steps that did not occur in all sequences: from 15 and 11 seconds when new to 5 and 4 seconds when given). These findings replicate Gerwing (2003, 2004) with a new task and a quantitative measure. These results further extend the given-new effect from words to co-speech gestures, as Levy and McNeill (1992) had proposed. The results also contribute theoretically. Clark and Haviland (1977) had explicitly proposed that the effect was part of the obligation of speakers to mark given information for their addressee. Fowler and Housum (1987) and Fowler (1988) proposed the same, yet their data were ambiguously dialogic. The present results are the first evidence with clearly dialogic data.

An alternative explanation of these effects could be that speakers would shorten their repeated gestures regardless of their addressee, perhaps because of effort. If so, the shortening would presumably happen gradually over all trials rather than on the first repetition. Also, Fowler (1988) found no

Figure 6. Still shot from the split-screen recordings overlaid with the Tangram stimuli used, and showing task Director (left panel) and Matcher (right panel).

significant shortening when reading words from a list and "more shortening of content words produced in a communicative context than in the same discourse, transcribed and read into a microphone" (p. 307). Thus, the reduction of gestures' duration observed here is likely to be the direct result of the given-new contract between speaker and addressee, as proposed by Clark and Haviland (1977).

Experiment 2. Tangram figures

Experiment 2 was an independent experiment performed in a different laboratory using a referential communication task (Glucksberg et al., 1966) with Tangram figures (geometrical shapes often perceived as resembling animate entities (shown in Figure 6). The same figures occurred repeatedly (in different orders) over several experimental trials, providing a test of the given-new contract, but with an important difference from Experiment 1. As noted earlier, it would be difficult to describe the dance steps solely in words, whereas the Tangram task has traditionally involved dyads who could and did use only words (e.g., Clark & Wilkes-Gibbs, 1986; Schober & Clark, 1989). Although Clark and Wilkes-Gibbs's analysis of the data did not refer specifically to a given-new process, it provided suggestive evidence for its effects: The number of words speakers used to describe the same figure to their addressee decreased significantly over the trials.

This task difference creates an opportunity for investigating how the availability and suitability of different communicative modalities – speech and gesture – may affect interlocutors' strategies for signaling given-new status. For example, it is possible that in the dancestep task, interlocutors resorted to shortening their gestures to signal given-new status because the gestures were doing most of the communicative work. In the Tangram task, where words are able to carry all of the communicative burden (see Clark & Wilkes-Gibbs, 1986) and gestures would be spontaneous, we might observe a different pattern. Shortening might happen primarily in the verbal modality while gesture duration remains comparatively unaffected. Alternatively, it is possible that speakers tend to signal given-new status not only for repeated words but also for gestures when they are repeated (not necessarily at the same time as words). In such cases, the given-new status would be conveyed in the relevant modality so that the two modalities would each show given-new effects.

In the Tangram task, each dyad has a Director and a Matcher. The Director's task is to instruct the Matcher on the order in which to place 12 Tangram figures. They did so repeatedly over six consecutive trials with the figures in different orders. Thus, the trials required that Directors refer to the individual figures (or their features) a minimum of six times. The dependent measure here was the duration of each new and repeated (gestural or verbal) reference in seconds. The given-new contract predicts that the duration of references for given information would be significantly shorter than when the information was new. Repeated references in this task did occur and were measured in speech and gesture independently (i.e., a verbal reference was not necessarily accompanied by a gestural reference, and vice versa). We predicted shorter durations after the first reference (for both modalities), regardless of where the new and given references occurred within the six trials.

Method

Design

Suitable Tangram videos were available from an experiment by Holler and Wilkin (2011a) and Holler et al. (2011), which was originally a mixed design with one within-participants factor (trials 1–6) and one between-participants factor (face-to-face vs. screen). Our purpose here was to show whether both words and gestures in face-to-face dialogue would show given-new effects, not to test the effects of visibility (which are complex in themselves; Bavelas & Healing, 2013). Therefore, the present analysis used only the data from the face-to-face condition, and the data in all following sections are from this condition only.

Participants

Sixteen participants from the University of Manchester (U.K.) participated in the face-to-face condition (12 females, 4 males; *M* = 26.50 years, *SD* = 6.37), from a variety of academic (students) and non-academic backgrounds (UofM employees). They formed eight Director-Matcher dyads (five female-female, one male-male, and two mixed-gender dyads.) All participants were native English speakers, with no known motor or speech impairments. Before the experiment started, all participants consented to being video-recorded as well as to their videos being retained by the experimenter for analysis. Participants were compensated for their participation financially or with course credits. Their participation in this experiment was approved by the University of Manchester School of Psychological Sciences Ethics Committee.

Materials

The Director and Matcher had identical sets of cards showing 12 different Tangram figures (Clark & Wilkes-Gibbs, 1986, p. 11, originally from; Elffers, 1976). The 10 cm x 13 cm figures were laminated onto individual paper squares (Figure 6). Each participant's set of cards was placed on their 75 cm x 150 cm table. The two tables were placed opposite one another, about 3 m apart. Each table had 12 empty 10 × 13 cm blank paper squares glued onto the table in two horizontal rows of six. These represented the ordinal positions of the individual cards for that trial. Each table also had a 20 cm high cardboard barrier on the side facing the other person, so the participants, who were both standing, could see each other but not each other's cards.

Equipment

Two wall-mounted Panasonic cameras provided color recordings of the participants in split-screen format (see Figure 6). We used ELAN version 5.3 (https://tla.mpi.nl/tools/tla-tools/elan/; Wittenburg et al., 2006) to annotate and measure the gesture data and used Praat version 6.1 (http://www.fon.hum.uva.nl/praat/; Boersma & Weenink, 2019) to measure the speech data. Both ELAN and Praat show time to 1000ths of a second.

Procedure

Before each trial, the Director's cards were on the paper squares in the correct order for that trial. The Matcher's cards had been shuffled and placed in a random order in one long line along the bottom edge of the table (i.e., beneath the blank paper squares where the Matcher would place them during the trial).

Assignment to the roles of Director and Matcher was random. The task implicitly required them to collaborate in order to converge on an unequivocal reference for each card so that the Director could instruct the Matcher how to place his or her cards in the same order as the Director's cards. They could talk to each other as much as they wished, and there was no explicit mention of gestures. The participants were to inform the experimenter when they thought they had achieved a full match for that trial (i.e., after the Matcher had placed all 12 cards). The participants then stepped away from their tables while the experimenter recorded the Matcher's order, rearranged the Director's cards into the pre-determined order for the next trial, and reshuffled the Matcher's cards. The experimenter did not provide any feedback about the performance after individual trials.

After the sixth trial, participants received feedback about their performance, a full debriefing, the financial compensation or course credit, and our thanks for their participation. The experiment took approximately 45 minutes in total.

Analysis

The analysts measured the duration of references to a Tangram figure (or to a specific feature of a figure) in gesture and then, independently, in speech. It is important to point out that the verbal and gestural references considered in these analyses often did not occur together. For example, the director might say "It looks like someone flying, they've got the arms out to both sides." The words "someone flying" might enter into the analysis of word duration, while a gesture depicting both arms being held out to the sides that accompanied the words "arms out to both sides" might enter the analysis of gesture duration. Moreover, gestures could be about individual figure features, and these may not have had any equivalent mention in the speech (i.e., they depicted complementary information). Also, individual referential descriptions were often accompanied by several different gestures occurring in succession (each of which may or may not have been repeated on a subsequent occasion). In yet other cases, figures may have been referred to exclusively in speech. This means that the data points considered in the analyses of gesture duration and word duration are not corresponding sets. For this reason, we analyzed words and gestures separately.

Gestures

The analysis focused solely on representational gestures that depicted the shape or size of a Tangram figure or its components (e.g., with the hand; see Figure 7) or that acted out the postures or actions of a figure (e.g., with the whole body; see Figure 8). In McNeill's (1992) system, these are iconic gestures. We excluded most deictic gestures (McNeill, 1992) because they did not describe a figure (i.e., Directors used them primarily to refer to the stimulus cards on the table). However, we did include iconic gestures with deictic elements, such as tracing a shape in the air with the index finger or depicting the motion of a character as running from left to right with the extended index finger, where the hand depicts the motion path as well as pointing to the motion end point. Also included were

Figure 7. Still frame of an individual movement counting as one gesture.

Figure 8. Still frames illustrating two different gestural movements (movement 1: hand on lower back; movement 2: leg up) produced in temporal overlap (right panel), thus counting as one combined construction (and as one gesture).

deictic gestures that were part of the depiction of the figures' actions or shapes, such as when saying "and he's pointing to the right," accompanied by the speaker pointing to the right. Finally, we excluded interactive gestures (Bavelas et al., 1992, 1995), such as using an open hand with the palm up to invite a response. By definition, these gestures refer to the Director-Matcher interaction rather than representing any figure. None of the gestures was double-coded. The final dataset for the present analysis comprised 96 different gestures (referred to below as R0) with at least one repetition (referred to below as R1, R2, R3, etc.). The total number of gestures in the present analysis, including all repetitions, was $N = 306$ gestures.

Repeated gestures. The next step was to locate repeated gestures, that is, any that resembled a preceding gesture for the same figure and that depicted either the same feature of that figure or that figure in its entirety. So, for example, a gesture depicting a diamond shape only counted as a repetition of an earlier depiction of a diamond shape when both referred to the same aspect of the same figure (e.g., the diamond-shaped head of the figure). Crucially, the gesture was not a repetition if the first occurrence referred to the diamond-shaped head and the second to a diamond shape sticking out at the back of the figure. However, the mode of the gestural representation – that is, the manner in which information was depicted, such as tracing versus molding a shape (Müller, 1998) did not matter for the identification of repeated gestures: On one occasion, the Director may have depicted a diamond-shaped head with index fingers and thumbs of both hands touching each other to outline the shape. A later gesture using the index finger of one hand to trace the outline of the diamond shape counted as a repetition. The rationale for not considering differences in the mode of representation as a criterion was that such changes were likely to be the direct consequence of Directors simplifying their gestures as a result of incremental common ground. This follows Gerwing and Bavelas (2004, Analysis 2 & Appendix A), which showed that speakers physically transformed an earlier gesture in a variety of ways, such as leaving out some features, in order to reflect its status as given information. It is also consistent with Fowler and Housum's (1987) finding that repeated words were less intelligible, and with Holler and Wilkin's (2011a) analysis of mimicked gestures in the process of grounding; their study suggested that interlocutors considered gestures as repeated versions of earlier ones even when these involved slight changes in the mode of representation. Last but not least, changes between different modes of gestural representation appear much more minimal than the changes in form occasioned by a synonym replacement, where one word (e.g., couch) is replaced by an entirely

different one (e.g., sofa); see criteria for speech below. We thus consider changes in the mode of representation in the present task to be more like changes in pronunciation, as in Fowler and Housum (1987).

The Directors' gestures sometimes depicted salient features of a figure with one individual movement (n = 164; Figure 7). However, in many cases, their gestures were combined constructions (n = 142), that is, movements we considered to form a single gestural representation but which consisted of several movements produced in close succession and each depicting a different figure feature. Crucially, each of these feature depictions was held until all of them had been depicted. This created temporal overlap between the individual features, which was the criterion for considering it as one gesture in the analysis (Figure 8). These were only considered repeated if they were later produced as the same combined constructions.

Finally, gestures that were "corrected" by the participants themselves were not included in the analysis of repeated gestures. For example, the diamond-shaped head had been described and depicted as hanging off the left side in trials 1 and 2, but after noticing the error, the Director correctly depicted it as hanging off the right side from trial 3 onwards. In such cases, the first, erroneous gestures were not included in the analysis, but the launch of the new, correct gesture and its following repetitions were.

Gesture duration.   Duration was measured from the beginning to the end of each gestural movement, as identified on a frame-by-frame basis (24 fps, later converted to milliseconds, using ELAN, see Equipment section). In this study, the *preparation phase* of each gesture (i.e. the hand moving from its resting position into the position where the main gestural depiction was performed; Kita et al., 1998) was included in the duration measure because it was considered part of the overall depiction. In many of the preparation phases, the hands had already begun to shape what would ultimately become the full configuration present in the most meaning-bearing part, termed the *stroke phase* (Kita et al., 1998; see Figure 9 for an example). Gesture preparation and gesture stroke together form the *gesture phrase* (Kendon, 1980).

The need to consider the gesture preparation phase is different from Experiment 1 in which the Teachers started each gesture from a prescribed starting position with both hands (as "feet") flat on the table. Only when the Teachers began to move from this position were they beginning to enact the dance step. Therefore, neither moving the hands from wherever they had been to the prescribed initial position on the table nor moving the hands from this static position on the table before beginning the dance step were meaningful parts of the gestural depiction. In both experiments, the end of a gestural movement was the onset of the gesture retraction or the onset of a new stroke (if they did not temporally overlap). In the case of combined gestures, consecutive strokes were considered to form

Figure 9. Example of a gesture preparation phase, with the stills illustrating the gradual unfolding of semantic information (hands depicting a triangle shape) during this phase.

one gesture, as described above. In short, the unit of measurement for gesture duration was Kendon's (1980) *gesture unit* (bar the retraction), sometimes containing single gesture phrases and sometimes a group of gesture phrases. A second analyst, blind to the experimental hypotheses, timed the duration of 16.3% of the repeated gestures; the correlation between the two analysts was $r(64) = .74$, $p < .0001$.

Speech

As in Fowler and Housum (1987), we measured the duration (in msec, using Praat, see Equipment section) of individual lexical items or a small group of words forming a short phrase, which the Director used repeatedly to refer to a figure (or part thereof). Examples from the present dataset include "monk," "rabbit," "baby," "goalkeeper," "bird," "ice skater," "sliding tackle," "kungfu guy," "someone kneeling down," "guy shooting a gun with a flag," "man who looks like a zombie," "woman kneeling," "man going like that," "flying" (see Table 5 for two more detailed examples). The analysis focused on exact repetitions only, which excluded synonyms (e.g., "guy" after "man") and identical lexical items when they differed in grammatical function. For example, if the term "ostrich" had initially appeared as a noun ("the ostrich"), it was not a repetition when later functioning as an adjective ("the ostrich-shaped one"). The rationale was that we know interlocutors are sensitive to synonym replacements, taking them as referring to a different entity when they occur (e.g., Bögels et al., 2015; Brennan & Clark, 1996).

As a general rule, we considered whole noun phrases, but if that was not possible because no complete noun phrase was repeated, then a different lexical item was chosen (e.g., "scare someone," "flying"). If there were several options to choose from (e.g., two different noun phrases were repeated for the same figure), we choose the one that occurred first. We included all repetitions that fulfilled the above criteria, no matter which trial they occurred in or the amount of other speech material that occurred between repetitions. The final dataset for the present analysis comprised 85 different lexical items/phrases (referred to as R0 below) with at least one repetition (below referred to as R1, R2, R3 etc.). The total number of lexical items/phrases in the present analysis, including all repetitions, was $N = 309$.

Statistical analysis

The Tangram task is much less structured than the Dance Step task in Experiment 1 in that interactants could go back and forth between the different figures freely. Thus, trial 3 might include the second repetition for one figure, but the fifth, sixth or seventh for another figure. In order to provide precisely comparable data, we analyzed the gestures or words by the number of the repetition

Table 5. Two examples of a repeated lexical noun phrase (example 1) and a repeated verb (example 2) (only instances in bold print were considered to meet the criteria, the remainder is included for context).

Trial	Repetition	Description
Example 1		
1		It's got you know the head shape where it's like a square on the top at an angle
2	R0	A **woman kneeling** with a triangle sort of feet behind or something
3	R1	A **woman kneeling** with like heels or something
4	R2	A **woman kneeling** with a book or something
5		Someone kneeling with a triangle to the bottom left sticking up two squares and two prongs to the right
6		It's the other similar shape one to that but with the triangle at the bottom left with the almost like someone kneeling forward
Example 2		
1	R0	A diamond shaped head it looks like a figure kind of **flying** across the sky really with two hands sticking out and it's um I think it's called a parallelogram actually as a shaped body
2	R1	Looks like something **flying** through the sky both hands sticking out
3	R2	The thing that looks like it's **flying** with two arms out stretched
4	R3	A figure which is **flying** with its arms sticking out
5	R4	The thing that's **flying** with its arms out
6	R5	The **flying** thing

rather than by the trial in which each occurred: R0 for the first reference, then R1, R2, R3, and so forth for the repetitions. As a result, the total number of repetitions for each figure varied, and dyads also varied in the number of total repetitions (collapsed across figures).

Furthermore, participants in the Tangram task could freely choose whether to gesture when describing a figure and which components of it they would depict. This led to gestures for the same figure differing greatly between participants. For example, one Director referred to a figure with a simple zig-zag gesture while another referred to the same figure by gesturing the square body shape with zig-zag movements on one of the sides (in a combined gesture). The same applied to lexical items/phrases. One Director repeated the lexical item "ice-skater" while another referred to the same figure with the lexical item "ostrich." We took into account the fact that the gestured figure features and the lexical items/phrases were not exactly the same between the Director-Matcher pairs by modeling them in a nested random structure (see below).

As for Experiment 1, we analyzed the data using R (Version 4.0.4, R Core Team, 2020) and the lme4 package (Version 1.1.26; Bates et al., 2015). We took the same steps for the analyses of both words and gestures: First, to test whether repetition has an effect on duration (gesture or words), we fitted a generalized linear mixed model by maximum likelihood (Laplace approximation) with repetition as the fixed effect and director-matcher pair and gestured figure feature /lexical item in a nested structure as random effects. The response family was inverse Gaussian with a logit link. In addition, for gesture duration, we fitted a generalized linear mixed model with a matching structure to the first one, but with repetition treated as a categorical variable to explore the effect of the individual repetitions. For word duration, we proceeded in exactly the same way to test for the general effect of repetition on duration. However, to test our prediction about the replication of Fowler and Housum's (1987) and Fowler's (1988) effect on word duration, our planned comparisons only referred to R0–R1, since these earlier studies, too, only compared the first and second mention for each item. To do so, we used sum-to-zero contrasts (with R0 coded as +0.5 and R1 coded as −0.5).

Results

The experimental dataset as a whole consisted of 19,143 words and 2230 gestures, of which 1148 were iconic gestures. The large majority of words and gestures occurred on trial 1 (8357 words and 965 gestures, of which 664 were iconic). The following duration analyses refer to the subset of 306 repeated gestures and 309 repeated lexical items/phrases described in the Method section.
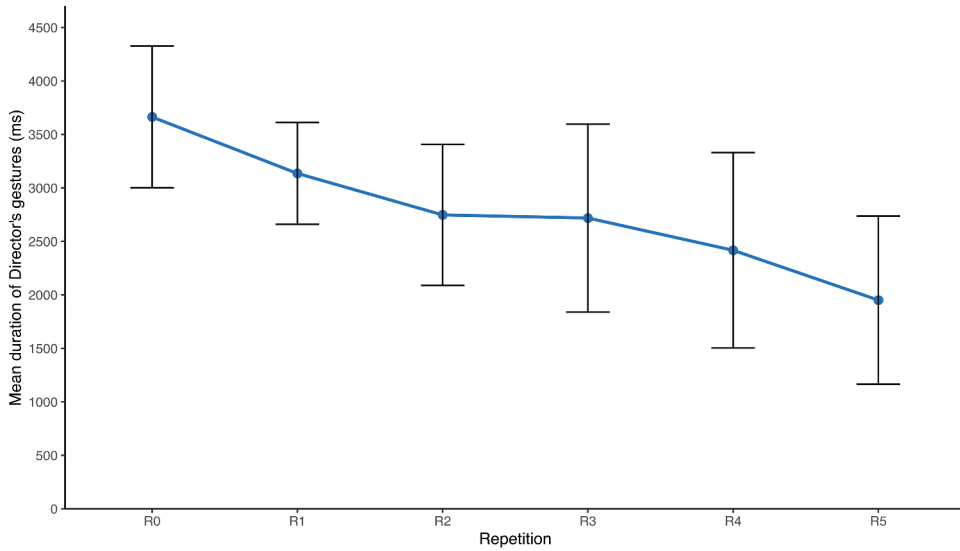
Gesture

Table 6 and Figure 10 show that mean gesture duration decreased from R0 to R5. The general linear mixed model confirmed that the effect of repetition on gesture duration was significant ($\beta = -0.11$, SE = 0.02, t = −4.99, $p < .001$). The results of the model comparing gesture duration for the individual repetitions show that gesture duration for the first gesture occurrence was significantly longer than the duration of all of the following repetitions (R0–R1: $\beta = -0.19$, SE = 0.07, t = −2.56, $p < .05$; R0–R2: $\beta = -0.31$, SE = 0.09, t = −3.57, $p < .001$; R0–R3: $\beta = -0.27$, SE = 0.1, t = −2.62, $p < .01$; R0–R4: $\beta = -0.48$, SE = 0.12, t = −3.91 $p < .001$; R0–R5: $\beta = -0.67$, SE = 0.16, t = −4.09, $p < .001$).

Speech

Table 7 and Figure 11 show that lexical item/phrase duration became shorter across the repetitions R0 to R5. The general linear mixed model showed that the effect of repetition on the duration of lexical items/phrases was not significant ($\beta = -0.008$, SE = 0.01, t = −1.24, $p = .215$). However, the planned post-hoc comparison of the first and second mention of a lexical item/phrase does show a significant reduction (R0–R1: $\beta = 0.059$, SE = 0.022, t = 2.67, $p = .008$).

Table 6. Mean duration (and SD) of gestures across repetitions (in milliseconds).

R0	R1	R2	R3	R4	R5
3663.33	3135.72	2747.55	2718.24	2416.76	1950.89
(3270.42)	(2346.29)	(2344.43)	(2479.16)	(2005.84)	(1022.71)

Figure 10. Mean gesture duration across repetitions. Error bars represent 95% confidence intervals around the means.

Table 7. Mean duration (and SD) of words (lexical items/phrases) across repetitions (in milliseconds).

R0	R1	R2	R3	R4	R5
895.93	820.34	709.93	626.95	589.26	592.46
(512.95)	(451.68)	(392.31)	(219.14)	(182.63)	(249.37)
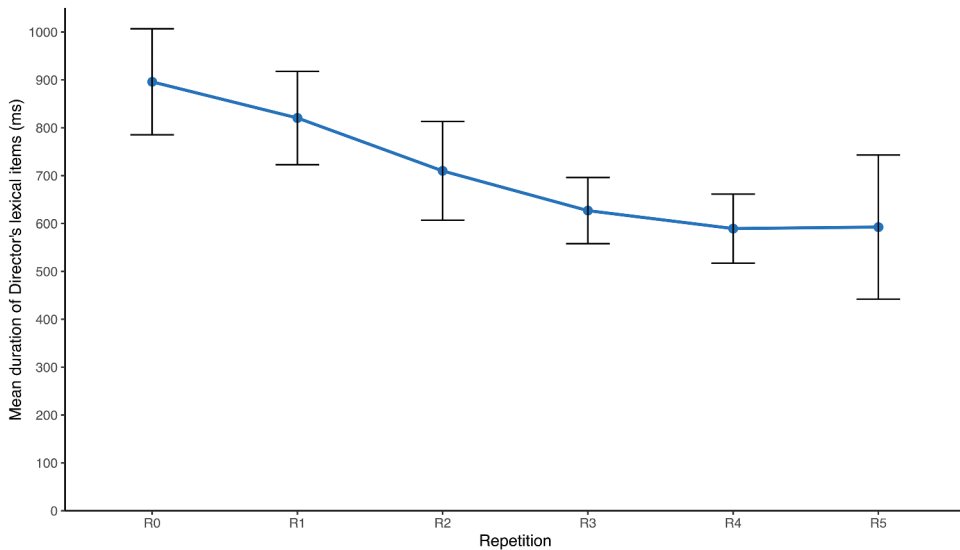
Figure 11. Mean duration of lexical items/phrases across repetitions. Error bars represent 95% confidence intervals.

Discussion

The gesture results in Experiment 2 provide an independent replication of Experiment 1, showing that repeated gestures are shorter in duration than their initial presentation, adding a second quantitative measure to the findings of Gerwing (2003, 2004). This finding shows that the given-new contract applies not only to gestures that carried the bulk of the communicative burden for conveying the relevant information (as in Experiment 1), but also to gestures in a task where speech, in principle, would have been sufficient (as demonstrated by earlier Tangram studies based on experiments without visibility, e.g., Clark & Wilkes-Gibbs, 1986). That is, signaling givenness appears to be inherent to the gestural modality. Manual gestures should thus be considered as an integral part of the given-new contract, further underlining gestures' core role in the pragmatics of dialogue.

The results for words are an independent replication of Fowler and Housum's (1987) and Fowler's (1988) findings using a different task and in a face-to-face dialogue. Fowler and Housum's studies were limited to speech generated in contexts that were unimodal (speech only) and monologic (on radio and television). Fowler (1988) showed the effect in a more social face-to-face context with an addressee present but probably not interacting. Here, the planned comparison between the first and second mention of each item – i.e., mirroring the original analyses by Fowler and colleagues – showed a significant reduction in word duration, replicating the original findings on word duration, this time in a cooperative face-to-face dialogue. The succeeding later comparisons were not significant.

Experiment 2 also showed that given-new effects on words occurred even when participants had more than words at their disposal (i.e., manual gestures). This is significant because these gestures carried a large proportion of semantic and pragmatic information and could thus take over much of the communicative burden (Bavelas & Chovil, 2000; Holler & Beattie, 2003; Kendon, 2004) which might have reduced the effect on words. The fact that the overall effect for words (based on all five repetitions) was not significant suggests that the presence of gestures may have had such an effect. However, we cannot conclude this with certainty since the original experiments by Fowler did not go beyond the second mention of words. Future research may be able to provide more insight into this issue.

Thus, while words and gestures both fulfill referential functions by conveying propositional information, both also share the task of indicating the pragmatic status of this propositional information by varying their duration. In sum, speakers appear to use both modalities, speech and gesture, to mark references as given. In fact, the effect seemed somewhat more pronounced for the duration of gestures than words. This is good reason to argue that the given-new contract should be considered a multimodal phenomenon.

General discussion and conclusions

The present studies investigated the given-new contract (Clark & Haviland, 1977) from a multimodal, dialogic perspective by focusing on the duration of repeated gestures and words in face-to-face dialogues. Specifically, we sought to test two predictions. First, we predicted that this quantitative measure of speakers' co-speech gestures would differentiate between new and given information. This was indeed the case. In both experiments, speakers produced repeated gestures referring to the same entity, and these repetitions were shorter than the original. The shorter durations are consistent with earlier qualitative work by Gerwing (2003, 2004) who observed that repeated gestures depicted given information in a "sloppier"(p. 176), less precisely articulated manner, with the information slowly fading from the gestural depictions with repeated referring. Here, two new and independent datasets differing in a number of respects both showed that speakers mark the given status of repeated gestures with shorter duration, just as they have been shown to mark words (Fowler & Housum, 1987).

Second, we predicted that the given-new effects on words could be extended to truly interactive, dialogic contexts, which are the most common form of human spoken communication. The original findings on the given-new effect by Fowler and Housum (1987) were based on 19 minutes of speech

material from a radio program involving a single narrator telling a scripted story. Although this passage resembled naturalistic speech, the words were not produced spontaneously, and, most importantly, were in a monologue. Fowler (1988) showed that the shortening effect was significantly more likely in spontaneous prose to another person than when reading the same words, transcribed, into a tape recorder. Still, the other person was the experimenter, and there was no dialogue. Unscripted conversational speech produced in face-to-face interaction with another person differs fundamentally from monologic speech produced outside of a social context, for example, in the use of discourse markers, direct quotations, facial displays, co-speech gestures, number of words, and, notably, also word duration (Bavelas et al., 2014, 2008; Fox Tree, 1999; McAllister et al., 1994). The present findings allow us to conclude that the given-new effect on words is also signaled through a shorter duration of words in interactive, dialogic speech. While the main effect on word duration was not significant when considering all five repetitions, the planned comparison mirroring Fowler's original comparison comparing first and second mentions was.

To some extent, our results also speak to the issue of whether the extent to which gestures or words could carry the burden of communication in each task might modulate the given-new effect in the two modalities. When teaching the dance steps in Experiment 1, the Teachers had to use gestures to convey the essential information. In contrast, the Tangram task in Experiment 2 can be completed entirely in words when gestures are not available (e.g., because the participants could not see each other; Clark & Wilkes-Gibbs, 1986). However, interactants doing the Tangram task in face-to-face interaction do make frequent use of gestures in addition to speech (e.g., Holler et al., 2011). Despite this difference in how much of the communicative burden gestures carried in these two tasks, gesture duration showed given-new effects in both tasks, which suggests that the role of gestures in marking information status is pervasive and not just a secondary, compensatory mechanism when words are less effective.

More broadly, we propose that hand gestures show given-new effects because they are part of language use in face-to-face dialogue, which is inherently multimodal. The status of hand gestures as part of language use in face-to-face dialogue has become widely accepted (e.g., Bavelas, 2022; Enfield, 2009; Goldin-Meadow, 2003; Kendon, 2004; McNeill, 1992; Streeck, 2009). What is less often explicit is that gestures have linguistic functions and follow linguistic imperatives. In the present case, because gestures also convey information in the dialogue and form an integral part of language, they have to follow the given-new contract for the same reason words do. If a repeated gesture were not marked as given, then the status of the information would be unclear to the addressee. It follows that words and gestures are following the same given-new contract but doing so independently. When gestures were the best-suited modality for the task at hand, as in Experiment 1, the gestures' duration reflected their new or given status. When both gestures and words were useful, as in Experiment 2, both modalities indicated given-new status by duration (but not necessarily at the same time).

One important question is how gesture and speech interact in marking information as new or given during reference in dialogue. Analyses of the interplay between gesture and speech in marking information status in narrative has provided interesting insights in monologues (e.g., Debreslioska & Gullberg, 2019, 2020a, 2020b; Debreslioska et al., 2013; Gullberg, 2003, 2006; Levy & McNeill, 1992; Levy & Fowler, 2000; McNeill & Levy, 1993; Wilkin & Holler, 2011; Yoshioka, 2008). In contrast, the present study focuses on dialogue. However, the present dataset only allowed for analyses looking at gestures and words separately, irrespective of what was happening in the respective other modality (mainly due to too few repeated gesture-speech ensembles). An important next step will be to compare effects on duration for repeated gestures unaccompanied by repeated words (and vice versa) with repeated speech-gesture ensembles. Also, the particular manner in which information is encoded in one modality may affect the temporal dynamics of the other. For example, the inclusion of hedging devices or other insertions in speech (e.g., uh, um, well) could influence the timing of gestures, as could the particular syntactical construction used. Likewise, the complexity, size or velocity of the gesture could influence the timing of speech. Finding out how the temporal interplay of speech and gesture may (or may not) influence the marking of given information through duration would be an interesting consideration in future analyses.

Another important aspect to investigate in connection with the marking of information status in dialogue is the addressee's contribution. Addressees play an important role in the process of grounding – that is, in accepting and approving their understanding of information as given (e.g., Bavelas et al., 2017; Holler & Wilkin, 2011a; Hömke et al., 2018; Schober & Clark, 1989). Future studies may focus on investigating how the addressees' contributions interact with speakers marking information as given in gesture and speech. Finally, an interesting avenue for future research is how the marking of givenness through shortened duration interacts with other forms of marking information status, such as the use of definite articles, gesture precision, viewpoint, and other changes (e.g., Clark & Wilkes-Gibbs, 1986; Debreslioska & Gullberg, 2020b; Gerwing & Bavelas, 2004). Despite the need to gain a much better understanding of these and related issues through future research, here we have been able to show that two experiments that differed in task, emphasis on gesture, extent to which speakers made use of speech, presence of gesture preparations, and other factors, yielded converging findings.

Thus, the present studies underline the important contribution co-speech gestures make to the pragmatics of human communication, and they extend our knowledge about what those functions are and how they are achieved. Note also that the gestures analyzed in the present studies combined the communication of semantic information with information about pragmatic status, thus underlining the complex multi-functionality of co-speech gestures. Past decades have highlighted the role of co-speech gestures in the process of communicating semantic, propositional information, and their pragmatic contribution has been acknowledged for quite some time (e.g., Bavelas, in press, 2022; Bavelas et al., 1992, 1995; Cooperrider et al., 2018; Graziano, 2014; Holler, 2010; Holler & Wilkin, 2011a, 2011b; Holler & Stevens, 2007; Kendon, 1985, 1995, 2004; Kelly et al., 1999; Müller, 2004; Seyfeddinipur, 2004; Streeck, 2009; Sweetser, 1990). Still, many of the specific functions that gestures fulfill in this domain are just being discovered. These discoveries are made possible by detailed studies of face-to-face dialogue, with the current studies being a further example.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Alibali, M. W., Nathan, M. J., Wolfgram, M. S., Breckinridge Church, R., Jacobs, S. A., Johnson Martinez, C. J., & Knuth, E. J. (2014). How teachers link ideas in mathematics instruction using speech and gesture: A corpus analysis. *Cognition and Instruction*, 32(1), 65–100. https://doi.org/10.1080/07370008.2013.858161

Arnold, J. E., Kaiser, E., Kahn, J. M., & Kim, L. K. (2013). Information structure: Linguistic, cognitive, and processing approaches. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(4), 403–413. https://doi.org/10.1002/wcs.1234

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bavelas, J. B. (2022). *Face-to-face dialogue. Theory, research, and applications*. Oxford University Press.

Bavelas, J. B., & Chovil, N. (2000). Visible acts of meaning. An integrated message model of language use in face-to-face dialogue. *Journal of Language and Social Psychology*, 19(2), 163–194. https://doi.org/10.1177/0261927X00019002001

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, 21(4), 394–405. https://doi.org/10.1177/0146167295214010

Bavelas, J. B., Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse Processes*, 15(4), 469–489. https://doi.org/10.1080/01638539209544823

Bavelas, J. B. (in press). Gesturing for the addressee. In A. Cienki (Ed.), *The Cambridge handbook of gesture studies*. Cambridge University Press.

Bavelas, J. B., Gerwing, J., & Healing, S. (2014). Effects of dialogue on demonstrations: Direct quotations, facial portrayals, hand gestures, and figurative references. *Discourse Processes*, 51(8), 619–655. https://doi.org/10.1080/0163853X.2014.883730

Bavelas, J. B., Gerwing, J., & Healing, S. (2017). Doing mutual understanding. Calibrating with micro-sequences in face-to-face dialogue. *Journal of Pragmatics*, 121, 91–112. https://doi.org/10.1016/j.pragma.2017.09.006

Bavelas, J. B., Gerwing, J., Sutton, C., & Prevost, D. (2008). Gesturing on the telephone: Independent effects of dialogue and visibility. *Journal of Memory and Language*, 58(2), 495–520. https://doi.org/10.1016/j.jml.2007.02.004

Bavelas, J. B., & Healing, S. (2013). Reconciling the effects of mutual visibility on gesturing. A review. *Gesture*, 13(1), 63–92. https://doi.org/10.1075/gest.13.1.03bav

Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* [Computer program]. Version 6.1. http://www.praat.org/

Boersma, P., & Weenink, D. (2020). *Praat: Doing phonetics by computer* [Computer program]. Version 6.1.16. http://www.praat.org/

Bögels, S., Barr, D. J., Garrod, S., & Kessler, K. (2015). Conversational interaction in the scanner: Mentalizing during language processing as revealed by MEG. *Cerebral Cortex*, 25(9), 3219–3234. https://doi.org/10.1093/cercor/bhu116

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482. https://doi.org/10.1037/0278-7393.22.6.1482

Chafe, W. L. (1974). Language and consciousness. *Language*, 50(1), 111–133. https://doi.org/10.2307/412014

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

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). APA Books.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Ablex.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge University Press.

Clark, H. H., & Marshall, C. R. (1978). Reference diaries. In D. L. Waltz (Ed.), *Theoretical issues in natural language processing* (pp. 57–63). Association for Computing Machinery.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7

Cooperrider, K., Abner, N., & Goldin-Meadow, S. (2018). The palm-up puzzle: Meanings and origins of a widespread form in gesture and sign. *Frontiers in Communication*, 3, 23. https://doi.org/10.3389/fcomm.2018.00023

Crystal, D. (2001). *A dictionary of language*. University of Chicago Press.

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

Debreslioska, S., & Gullberg, M. (2019). Discourse reference is bimodal: How information status in speech interacts with presence and viewpoint of gestures. *Discourse Processes*, 56(1), 41–60. https://doi.org/10.1080/0163853X.2017.1351909

Debreslioska, S., & Gullberg, M. (2020a). The semantic content of gestures varies with definiteness, information status and clause structure. *Journal of Pragmatics*, 168, 36–52. https://doi.org/10.1016/j.pragma.2020.06.005

Debreslioska, S., & Gullberg, M. (2020b). What's new? Gestures accompany inferable rather than brand-new referents in discourse. *Frontiers in Psychology*, 11, 1935. https://doi.org/10.3389/fpsyg.2020.01935

Debreslioska, S., Özyürek, A., Gullberg, M., & Perniss, P. (2013). Gestural viewpoint signals referent accessibility. *Discourse Processes*, 50(7), 431–456. https://doi.org/10.1080/0163853X.2013.824286

Elffers, J. (1976). *Tangram. The ancient Chinese shapes game*. McGraw-Hill.

Enfield, N. J. (2009). *The anatomy of meaning: Speech, gesture, and composite utterances* (No. 8). Cambridge University Press.

Foraker, S. (2011). Gesture and discourse: How we use our hands to introduce and refer back. In G. Stam, M. Ishino, & R. Ashley (Eds.), *Integrating gestures* (pp. 279–292). John Benjamins.

Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, 31(4), 307–319. https://doi.org/10.1177/002383098803100401

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26(5), 489–504. https://doi.org/10.1016/0749-596X(87)90136-7

Fowler, C. A., Levy, E. (1991). Some ways in which forms arise from functions in linguistic communications, and A. Cutler (Ed.), *Prosody in situations of communication. Symposium conducted at the meeting of the 12th International Congress of Phonetic Scientists* (vol. 2, pp. 279–282). Aix-en-Provence.

Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, 27(1), 35–53. https://doi.org/10.1080/01638539909545049

Galati, A., & Brennan, S. E. (2014). Speakers adapt gestures to addressees' knowledge: Implications for models of co-speech gesture. *Language, Cognition and Neuroscience*, 29(4), 435–451. (Also available as *Language and Cognitive Processes*, 2013, because of a change in journal title). https://doi.org/10.1080/01690965.2013.796397

Gerwing, J. (2003). *The effect of immediate communicative function on the physical form of conversational hand gestures.* Unpublished Masters thesis, University of Victoria

Gerwing, J., & Bavelas, J. B. (2004). Linguistic influences on gesture's form. *Gesture*, 4(2), 157–195. https://doi.org/10.1075/gest.4.2.04ger

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross language study.* John Benjamins Publishing Company.

Glucksberg, S., Krauss, R. M., & Weisberg, R. (1966). Referential communication in nursery school children: Method and some preliminary findings. *Journal of Experimental Child Psychology*, 3(4), 333–342. https://doi.org/10.1016/0022-0965(66)90077-4

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

Graziano, M. (2014). The development of two pragmatic gestures of the so-called open hand Supine family in Italian children. In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance* (pp. 311–330). John Benjamins.

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

Gullberg, M. (2003). Gestures, referents, and anaphoric linkage in learner varieties. In C. Dimroth & M. Starren (Eds.), *Information structure and the dynamics of language acquisition* (pp. 311–328). John Benjamins.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(2), 199–244. https://doi.org/10.1017/S0022226700016613

Hickmann, M., & Hendriks, H. (1999). Cohesion and anaphora in children's narratives: A comparison of English, French, German, and Mandarin Chinese. *Journal of Child Language*, 26(2), 419–452. https://doi.org/10.1017/S0305000999003785

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

Holler, J., & Bavelas, J. (2017). Multi-modal communication of common ground: A review of social functions. In R. B. Church, M. W. Alibali, & S. D. Kelly (Eds.), *Why gesture? How the hands function in speaking, thinking and communicating* (pp. 213–240). Benjamins.

Holler, J., & Beattie, G. (2003). How iconic gestures and speech interact in the representation of meaning: Are both aspects really integral to the process? *Semiotica*, 146(1), 81–116. https://doi.org/10.1515/semi.2003.083

Holler, J. (2010). Speakers' use of interactive gestures to mark common ground. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction* (pp. 11–22). Springer. 8th International Gesture Workshop, Bielefeld, Germany, 2009; Selected Revised Papers.

Holler, J., & Stevens, R. (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. https://doi.org/10.1177/0261927X06296428

Holler, J., Tutton, M., & Wilkin, K. (2011, September 5–7). Co-speech gestures in the process of meaning coordination. In *Proceedings of the 2nd GESPIN- Gesture in Speech and Interaction Conference*, Bielefeld. http://hdl.handle.net/11858/00-001M-0000-0012-1BAF-A

Holler, J., & Wilkin, K. (2009). Communicating common ground: How mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24(2), 267–289. https://doi.org/10.1080/01690960802095545

Holler, J., & Wilkin, K. (2011a). Co-speech gesture mimicry in the process of collaborative referring during face-to-face dialogue. *Journal of Nonverbal Behavior*, 35(2), 133–153. https://doi.org/10.1007/s10919-011-0105-6

Holler, J., & Wilkin, K. (2011b). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics*, 43(14), 3522–3536. https://doi.org/10.1016/j.pragma.2011.08.002

Hömke, P., Holler, J., & Levinson, S. C. (2018). Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS One*, 13(12), e0208030. https://doi.org/10.1371/journal.pone.0208030

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

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

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics*, 23(3), 247–279. https://doi.org/10.1016/0378-2166(94)00037-F

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

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship between verbal and nonverbal communication* (pp. 207–227). De Gruyter Mouton.

Kendon, A. (1985). Some uses of gesture. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp. 215–234). Ablex Publishing Corporation.

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction, Lecture notes in computer science volume 1371* (pp. 23–35). Springer.

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*(1), 54–72. https://doi.org/10.3758/s13423-012-0341-8

Levy, E. T., & Fowler, C. A. (2000). The role of gesture and other graded language forms in the grounding of reference in perception. In D. McNeill (Ed.), *Language and gesture* (pp. 215–234). Cambridge University Press.

Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes*, *15*(3), 277–301. https://doi.org/10.1080/01638539209544813

Marslen-Wilson, W., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action* (pp. 339–378). Wiley.

McAllister, J., Potts, A., Mason, K., & Marchant, G. (1994). Word duration in monologue and dialogue speech. *Language and Speech*, *37*(4), 393–405. https://doi.org/10.1177/002383099403700404

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

McNeill, D., & Levy, E. T. (1993). Cohesion and gesture. *Discourse Processes*, *16*(4), 363–386. https://doi.org/10.1080/01638539309544845

Müller, C. (1998). *Redebegleitende Gesten: Kulturgeschichte, Theorie, Sprachvergleich*. Arno Spitz Verlag.

Müller, C. (2004). Forms and uses of the palm up open hand: A case of a gesture family. In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gestures: Proceedings of the Berlin conference April 1998* (Vol. 9; pp. 233–256). Weidler. Körper, Zeichen, Kultur.

Perniss, P., & Özyürek, A. (2015). Visible cohesion: A comparison of reference tracking in sign, speech, and co-speech gesture. *Topics in Cognitive Science*, *7*(1), 36–60. https://doi.org/10.1111/tops.12122

R Core Team. (2020). *R: A language and environment for statistical computing*. https://www.R-project.org/

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology*, *21*(2), 211–232. https://doi.org/10.1016/0010-0285(89)90008-X

Seyfeddinipur, M. (2004). Meta-discursive gestures from Iran: Some uses of the 'Pistol Hand.' In C. Müller & R. Posner (Eds.), *The semantics and pragmatics of everyday gestures: Proceedings of the Berlin conference April 1998* (Vol. 9; pp. 205–216). Weidler. Körper, Zeichen, Kultur.

Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning*. John Benjamins Publishing.

Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure* (Vol. 54). Cambridge University Press.

Vajrabhaya, P., & Pederson, E. (2018). Teasing apart listener-sensitivity. The role of interaction. *Gesture*, *17*(1), 65–97. https://doi.org/10.1075/gest.00011.vaj

Wilkin, K., & Holler, J. (2011). Speakers' use of 'action' and 'entity' gestures with definite and indefinite references. In G. Stam & M. Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 293–308). John Benjamins.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006, May 22–26). ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation* (pp. 1556–1559).

Woods, J. (2005). *New vs. given information: Do gestures dance to the same tune as words?* Unpublished honours thesis, Department of Psychology, University of Victoria.

Yoshioka, K. (2008). Gesture and information structure in first and second language. *Gesture*, *8*(2), 236–255. https://doi.org/10.1075/gest.8.2.07yos