## Title
Do gestures really facilitate speech production?

**Permalink**
https://escholarship.org/uc/item/67z729c5

**Journal**
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

**ISSN**
1069-7977

**Authors**
Kısa, Yağmur Deniz
Goldin-Meadow, Susan
Casasanto, Daniel

**Publication Date**
2021

Peer reviewed

# Do Gestures Really Facilitate Speech Production?

**Yağmur Deniz Kısa (ykisa@uchicago.edu)**
Department of Psychology, University of Chicago
5848 S. University Avenue, Chicago, IL 60637 USA

**Susan Goldin-Meadow (sgm@uchicago.edu)**
Department of Psychology, University of Chicago
5848 S. University Avenue, Chicago, IL 60637 USA

**Daniel Casasanto (casasanto@cornell.edu)**
Department of Human Development, Cornell University
Martha Van Rensselaer Hall, Ithaca, NY 14853 USA

## Abstract

Why do people gesture when they speak? According to one influential proposal, the Lexical Retrieval Hypothesis (LRH), gestures facilitate speech production by helping people find the right spatial words. Do gestures also help speakers find the right words when they talk about abstract concepts that are spatialized metaphorically? If so, gesture prevention should increase disfluencies during speech about both literal and metaphorical space. We sought to conceptually replicate the finding that gesture prevention increases disfluencies in speech about literal space, which has been interpreted as evidence for the LRH, and to extend this pattern to speech about metaphorical space. Our large dataset provided no evidence that gestures facilitate speech production, even for speech about literal space. Upon reexamining past research, we conclude that there is, in fact, no reliable evidence that preventing gestures makes speech more disfluent. These findings challenge long-held beliefs about why people gesture when they speak.

## Introduction

Why do we gesture when we speak? For some gestures, it is self-evident that the speaker intends them to be communicative, as when we wave "hello" or give a "thumbs up" gesture to signal approval. Do speakers also gesture because gesturing serves a cognitive function in the speaker's mind, helping them to think or to talk?

For decades, researchers have posited that gestures facilitate speech production (Butterworth & Hadar, 1989; Krauss, 1998; Krauss & Hadar, 1999). According to an influential version of this proposal, the Lexical Retrieval Hypothesis (LRH), gestures facilitate speech production by helping speakers find the right words; however, gestures are only posited to affect words of a particular kind. Krauss and colleagues (1996; 1998) noted that people gesture far more frequently during phrases with spatial content than during phrases without it, and hypothesized that gesturing helps speakers find the right *spatial words*. Krauss and colleagues hypothesized that some gestures derive from knowledge encoded in a spatial format; the spatial features of gesture (e.g., upward trajectory) facilitate production of spatial words by priming the spatial features (e.g., upwardness) that enter into the search for that word (e.g., for the word "up"; Krauss & Hadar, 1999).

What is the evidence that gesturing helps speakers find the right spatial words? According to Krauss and colleagues (1996), if people have difficulty finding words, their speech should be more disfluent. If gesturing helps speakers find the right words, then preventing speakers from gesturing should make their speech production more disfluent, compared to when they are free to gesture. In an influential study, Krauss and colleagues claimed that preventing people from gesturing increased disfluencies (pauses, repairs, etc.) selectively for the production of spatial phrases (Rauscher, Krauss, & Chen, 1996). This study is frequently cited as evidence that gesturing helps speakers find the right spatial words and is discussed in reviews summarizing the state of knowledge about speaker-internal functions of gesture (e.g., Goldin-Meadow, 1999; Hostetter & Alibali, 2008). Largely based on this one study, one speaker-internal cognitive function played by gesture appears to be helping speakers find the right words and thus facilitating speech production (e.g., Krauss, 1998).

If LRH only explains how gestures help people talk about space, however, it provides a limited account of how gesturing helps speaking. People spend a lot of time speaking about non-spatial ideas, including highly abstract concepts: entities like *time* and *value* that have no spatial magnitude, direction, or location. Yet, there is abundant evidence that people use space metaphorically to speak, think, and gesture about abstract concepts (Lakoff & Johnson, 1980; McNeill, 1992; for a review, see Casasanto & Bottini, 2014). In one study, on which the present study builds, people spontaneously produced gestures whose form reflected the spatial direction implied in their speech (e.g., upward), regardless of whether they talked about concrete space (e.g., "the rocket went up") or metaphoric space (e.g., "my grades went up"; Yap, Brookshire, & Casasanto, 2018). People's gestures reflected the predicted spatial directions (e.g., better is metaphorically upward), even when they talked about abstract concepts without using any spatial words (e.g., "my grades got better"). According to this study, which analyzed over 5000 gestures, people were just as likely to gesture spontaneously in the predicted directions for metaphorically spatialized concepts (e.g., grades are not the kind of entity that can literally rise in space) as for literal spatial concepts. These results suggest that, like words for literal spatial concepts, words for metaphorical spatial concepts correspond to particular kinds of spatial information in speakers' minds (e.g., schematic representations of upward, downward, rightward, or leftward space), even when these words have no literal spatial uses (e.g., the word "better" cannot be used sensibly to denote literal spatial locations or paths).

If spatial information is activated in memory not only when people produce literal spatial language, but also when they produce metaphorical spatial language, then the same gestural mechanism should help people find words for both literal spatial scenarios and metaphorically spatialized ideas. Gesturing upward, for example, should help speakers not only to produce words or phrases like "my rocket went up," but also to produce words or phrases like "my grades went up" and perhaps even "my grades got better." If so, this discovery would substantially expand the scope of gesture's role in speech production, particularly since spatial schemas appear to be part of people's mental representations in many non-spatial conceptual domains that become spatialized metaphorically in language and thought, including time, number, and emotional valence, among others (Lakoff & Johnson, 1980).

Here, we tested whether gestures serve a speaker-internal cognitive function by helping people find the right words with literal or metaphorical spatial content. We sought first to conceptually replicate Rauscher and colleagues' (1996) study testing whether gesturing helps speakers produce words for literal spatial scenarios, and then to determine whether this benefit extends to producing speech about metaphorically spatialized ideas. We compared rates of speech disfluencies when people were allowed to gesture freely and when they were prevented from gesturing as they told stories with either literal or metaphorical spatial content. If gesturing only helps people find the right concrete spatial words, as suggested by Rauscher and colleagues (1996), then preventing people from gesturing should increase the number of disfluencies only for speech with literal spatial content. Alternatively, if gesturing can also help speakers find the right abstract words, then preventing people from gesturing should increase the number of disfluencies for speech not only with literal spatial content, but also with metaphorical spatial content.

To preview our findings, contrary to our expectations based on earlier claims, preventing gesture had no significant effect on disfluency rate— not for speech with metaphorical spatial content, and not even for speech with literal spatial content. Due to the large number of data points in our study, it is not likely that the absence of these effects was the result of low statistical power. In response to this unexpected outcome, we first scrutinized our own data to confirm that there was no effect of gesture prevention beyond our planned analyses, in subsets of the data (i.e., disfluency rates for different types of disfluencies). We then scrutinized the results of Rauscher et al's (1996) study and other studies that sought to extend Rauscher et al.'s results in the two decades since their publication. Upon reexamining these studies, we conclude that (as in the present study), preventing gesture had no interpretable effect on speech production, even for literal spatial words, motivating a reexamination of widely-held beliefs about why people gesture when they speak.

## Method

### Participants

56 Stanford University undergraduates (28 male) were recruited in pairs, and participated for course credit after giving informed consent.

### Materials

There were 12 brief stories in total, each 50-100 words, implying motion or extension in one of four spatial directions: upward, downward, right, or left. Four of the stories had literal spatial content, describing actual spatial scenarios in the physical world using concrete spatial directions (e.g., "the rocket went *higher*"; "the scuba diver went *down*"). Eight of the stories had metaphorical spatial content, describing abstract non-spatial phenomena that are nevertheless commonly talked and thought about using spatial directions metaphorically (e.g., "my grades went *higher*"; "the price went *down*"). Each of the eight metaphorical stories had two versions: metaphorical stories with spatial language and metaphorical stories without spatial language. Metaphorical stories with spatial language described non-spatial phenomena using spatial words or phrases in their abstract metaphoric senses (e.g., "my grades went *higher*"). Metaphorical stories *without* spatial language were identical to metaphorical stories with spatial language, except that spatial words or phrases that are used metaphorically were replaced with non-spatial paraphrases conveying nearly the same meaning and implying the same spatial directions (e.g., "my grades got *better*").

### Procedure

Participants were told that the experiment was about storytelling. They took turns studying written stories, each for 60 seconds, and then retelling the stories to their partners. They were told to retell the stories as accurately as possible because their partner would be quizzed on the content of the stories. All stories were written in the second person (e.g., "You're testing some new model rockets"), but participants were asked to retell the stories in the first person (e.g., "I'm testing some new model rockets") as if retelling their own experiences.

After starting with a warm-up story, each participant retold 6 stories in randomized order: 2 stories with literal spatial content and 4 stories with metaphorical spatial content. Each pair of participants received only one version of each metaphorical story: either with spatial language or without spatial language (i.e., one pair of participants would receive either the story about "grades going higher" or the story about "grades getting better").

Each pair of participants was assigned to one of two gesture conditions: gesture prevented or gesture allowed. In the gesture prevented condition, participants were instructed to hold down keys on a computer keyboard, one key with each hand, during the entire time they were retelling the stories. They were told that the keys activated the microphones mounted on top of the computer monitor in front of them; in fact, the microphones were nonfunctional. In the gesture allowed condition, participants simply told the stories without being instructed to hold down keys on a keyboard; they were not told to gesture. Testing lasted 20-30 minutes.

## Coding

Analyses of the gestures from the Gesture allowed condition were reported in Yap et al. (2018), but no analysis of speech disfluencies was reported, and no data from the Gesture prevented condition have been reported previously. In the Gesture allowed condition, participants produced a total of 2249 gestures including 1609 beats, 328 iconic gestures, 252 deictics, 48 metaphoric gestures, 10 adaptors and 2 emblems. Beats were categorized solely based on form, following McNeill's (1992) beat filter. When categorized based on meaning with respect to accompanying speech, 629 of the 1609 beats reflected the spatial ideas expressed in the accompanying speech. So, the overall rate of gestures that would be predicted to facilitate spatial speech (i.e. beats reflecting spatial semantics, iconic gestures, deictics and metaphoric gestures) in the Gestures allowed condition was 56 percent (1257 out of 2249 gestures).

**Speech content coding** Participants' audio recordings of the stories were transcribed verbatim; 22 of the 336 stories were excluded because the speech was inaudible. The transcriptions of participants' audio recordings of the stories were parsed into clauses and phrases. We determined whether each phrase had spatial content, literal or metaphorical. A given phrase was classified as having spatial content if it contained language that implies literal or metaphorical motion, extent, or position along either the lateral or vertical axis (e.g. "went higher"). Alternatively, phrases were classified as having no spatial content if they did not imply a literal or metaphorical spatial schema.

**Spatial content type coding** Phrases with spatial content were classified as literal or metaphorical, and phrases with metaphorical spatial content were further classified as having or not having spatial language, using the same criteria used to construct the stories. Participants produced a total of 7969 spoken phrases (Gesture allowed 4000; Gesture prevented: 3969). Overall, 2801 phrases included spatial content, with 962 literal and 1839 metaphorical spatial content.

**Speech disfluency coding** We recorded the location and type of speech disfluency for each story. Speech disfluencies included repeats, repairs, filled pauses (uh, um, etc.), and unfilled pauses. Participants produced a total of 2075 disfluencies (Gesture allowed: 1041; Gesture prevented: 1034). Speech disfluencies included 905 filled pauses, 492 repairs, 446 unfilled pauses, and 232 repeats. For each story, we calculated the total number of disfluencies that occurred during phrases with spatial content (both overall and for each spatial content type separately) and during phrases without spatial content. Finally, for each story, we calculated the total number of words in phrases with and without spatial content, which we used as a baseline in our analysis.

### Analysis

We conducted all analyses by fitting generalized linear mixed-effect models, using R. We used "maximal" random effect structures justified by our design (Barr, Levy, Scheepers, & Tily, 2013), including not only random intercepts for Subject (N = 56) and Story (N = 12), but also random slopes for our fixed factors that are within-subject or within-story. We used likelihood ratio tests (LRTs) to test for fixed effects, with post-hoc contrasts performed on subsets of the data.

## Results

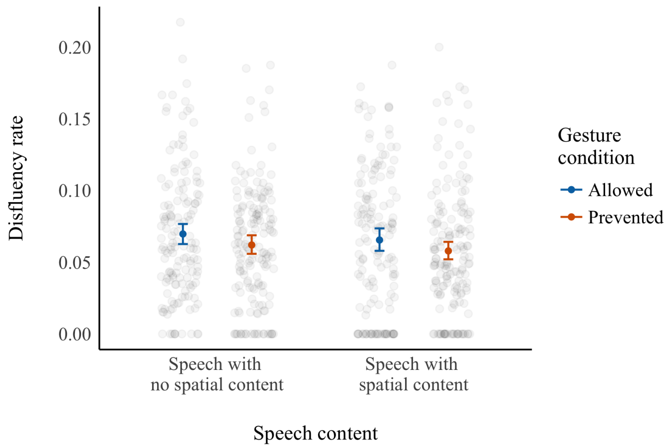
To compare speech disfluency rate across experimental conditions, we used mixed-effects Poisson regressions. We incorporated the number of words as an offset term into the model so that we modeled speech disfluency rate (number of speech disfluencies per word), rather than raw count data of number of speech disfluencies.

### Overall effect of gesture prevention on disfluency rate

Did people produce a higher rate of disfluencies when they were prevented from gesturing, compared to when they were allowed to gesture? In a first analysis including all speech content types (with and without spatial content), we found no evidence of an effect of gesture prevention on rate of speech disfluencies. Disfluency rates when people were prevented from gesturing ($M = 0.07, SD = 0.04, Median = 0.07$) were statistically indistinguishable from disfluency rates when people were allowed to gesture ($M = 0.06, SD = 0.03, Median = 0.06; \chi^2(1) = 1.07, p = .30$).

### Effect of gesture prevention during speech with vs. without spatial content

A second analysis tested the effect of gesture prevention on disfluency rates in speech with spatial content and in speech with no spatial content. Results showed that preventing people from gesturing had no significant effect on disfluency rates during speech with spatial content ($\chi^2(1) = 1.71, p = .19$) or during speech with no spatial content ($\chi^2(1) = 0.56, p = .45$). Notably, the nonsignificant trends went in the opposite direction of what the LRH would predict: People were slightly less disfluent when prevented from gesturing compared to when they are allowed to gesture, both during speech with spatial content and during speech with no spatial content (see Figure 1). The (non-)effect of gesture prevention did not differ significantly between speech with and without spatial content, as indicated by a non-significant interaction between Gesture condition and Speech content ($\chi^2(1) = 0.30, p = .58$).



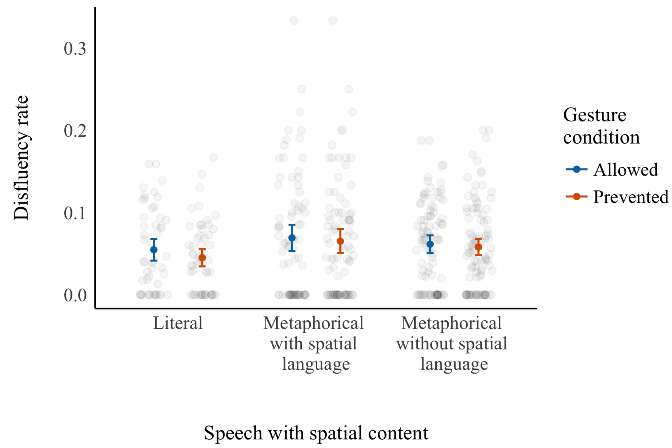




Figure 1: Disfluency rates during speech with spatial content (right) and speech with no spatial content (left). Grey dots show individual data points. Error bars show bootstrapped 95 percent CI around the group means.

Figure 2: Disfluency rates during speech with literal and metaphorical spatial content. Grey dots show individual data points. Error bars show bootstrapped 95 percent CI around the group means.

### Effect of gesture prevention for literal vs. metaphorical spatial content

A third set of analyses tested for effects of preventing gesture during speech with literal spatial content and with metaphorical spatial content. Results showed no significant effect of gesture prevention on disfluency rate for any type of spatial content (Literal: $\chi^2(1) = 0.97, p = .32$; Metaphorical with spatial language: $\chi^2(1) = 1.56, p = .21$; Metaphorical without spatial language: $\chi^2(1) = 1.09, p = .30$), and the (non-)effect of gesture prevention on disfluency rate did not differ across these conditions ($\chi^2(2) = 0.07, p = .97$, see Figure 2). The nonsignificant trends went again in the opposite direction of what the LRH would predict (see Figure 2).

### Effect of gesture prevention for different disfluency types

Did people produce a higher rate of any kind of disfluency (i.e. repairs, repeats, filled pauses, or unfilled pauses) when they were prevented from gesturing, compared to when they were allowed to gesture? There was no effect of preventing gesture on the rate of any type of disfluency. Similarly, there was no effect of preventing gesture on disfluency rates in speech with spatial content across different disfluency types. Notably, the nonsignificant trends went in the opposite direction of what the LRH would predict: For any kind of disfluency, people were slightly less disfluent when prevented from gesturing compared to when they are allowed to gesture during speech with spatial content.

Disfluency rate has been the measure that received the most attention in testing whether gesture prevention impairs speaking (e.g. disfluency rate results were reproduced in influential reviews such as Krauss, 1998). However, disfluency rate was not the only dependent measure that Rauscher et al (1996) reported. Rauscher et al (1996) also tested effects of gesture prevention on two other measures of disfluency, speech rate (number of words per minute) and non-juncture filled pause rate (rate of filled pauses to be within clauses rather than at the juncture of clauses), and interpreted their results on these two measures as support for the LRH. We also tested effects of gesture prevention on speech rate and non-juncture filled pause rate. Space precludes detailing our analyses here. However, in summary, we found no statistically significant difference in speech rate or non-juncture filled pause rate between speakers who were allowed to gesture freely and speakers who were not allowed to gesture for any category of speech (literal spatial content, metaphorical spatial content, no spatial content).

## Discussion

Does gesturing facilitate speech production by helping people find the right spatial words? We found no evidence that speakers speak less fluently when they are prevented from gesturing, compared to when they are allowed to gesture freely. We failed to find an effect of preventing gesture during speech with metaphorical spatial content, a finding that would have expanded the scope of LRH to encompass speech about abstract concepts. More fundamentally, we also found no significant effect of preventing gesture during speech with literal spatial content. We thus failed to find support for Rauscher and colleagues' (1996) influential claim that preventing gesture increases disfluencies for spatial language. More broadly, our data provide no support for the idea that gesturing facilitates speech production by helping people find the right words.

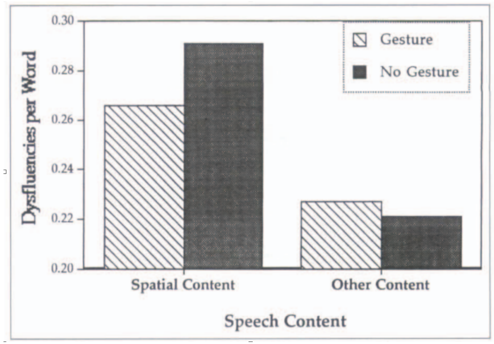



Figure 3: Disfluency rates for speech with spatial content and with non spatial content and when subjects were and were not allowed to gesture.
*Note*. Reprinted from "Why do we gesture when we speak?", by Krauss, R.M., 1998, *Current Directions in Psychological Science*, 7(2), p. 58.

## Why did we find no evidence that gestures help speakers find the right spatial words?

Why did preventing gestures have no effect on our participants' speech, even when speakers were using words like "up" and "down" to describe concrete spatial scenarios? A first possible explanation to consider for any null result may be lack of statistical power. This explanation is unlikely for our study, however, given both the size of our data set (7969 phrases containing 2075 disfluencies) and the qualitative patterns in the data. If a lack of power were responsible for our null effects then, overall, we would expect to find trends in the predicted direction that failed to reach statistical significance. This was not the case. On the contrary, as noted above, all of the trends for spatial speech went opposite from the predicted direction.

A second possible explanation for a null effect could rest in having the predicted effects 'hidden' in subsets of the data and obscured by aggregating over conditions or trial types. To ensure that this was not the case, we report planned analyses for all subsets of the data, broken out not only by spatial vs. non-spatial content, but also by multiple types of language (literal, metaphorical spatial, nonspatial language). To ensure that effects of any particular type of disfluency were not being masked by non-effects for other types, we also conducted post-hoc analyses of each disfluency type, individually, preventing gestures did not have the predicted effect on any type of disfluency.

The most fruitful explanation for the null results in the present study, we believe, rests in a re-examination of Rauscher et al.'s (1996) results, and of other studies testing effects of gesture prevention on disfluency. Much to our surprise, a careful examination of these studies yields no clear evidence that preventing gesture increases disfluencies in speech.

## Is there any evidence that gestures help speakers find the right spatial words?

Rauscher and colleagues' (1996) study has been widely cited as evidence for the idea that gesturing helps speakers find the right words and, more broadly, as some of the first evidence that gesturing serves a cognitive function for speakers (e.g., Krauss, 1998). However, a careful reexamination of Rauscher et al. (1996)'s disfluency rate results shows that they did not find statistically significant results supporting the idea that gesturing helps speakers find the right spatial words.

Rauscher et al.'s (1996) most influential claim is that preventing gesture causes higher disfluency rates *only* during speech with spatial content. The claim that gesture prevention makes spatial speech more dysfluent, but does not affect non-spatial speech, relies on showing two effects: First, preventing gesture should make spatial speech more disfluent, resulting in a simple effect of gesture prevention in speech with spatial content. Second, preventing gesture should increase disfluency *selectively* during speech with spatial content, as opposed to speech with no spatial content, resulting in a 2-way interaction of gesture condition (gestures allowed, gestures prevented) and speech content (spatial content, no spatial content).

Yet, the simple effect required to support this claim was not statistically significant, and the required 2-way interaction was never reported, neither in Rauscher et al. (1996), nor in subsequent review articles and chapters highlighting these results (see Figure 4 in Krauss, 1998, reprinted here as Figure 3). Regarding the required simple effect, the LRH predicted that participants should produce more disfluencies when they are prevented from gesturing, compared to when they are allowed to gesture, during speech with spatial content (i.e., the two bars on the left in Figure 3 should be significantly different from each other). However, this critical simple effect was only marginally significant, as reported (i.e., $p < .066$). Notably, even this reported value is anti-conservative, in at least two ways. First, the statistical test did not account for item-wise variance (Clark, 1973), leading to an increased probability of Type I error (i.e., a false positive result). Second, the alpha value for pairwise comparisons was not corrected for the multiple statistical comparisons reported, and the even greater number of comparisons that could have been conducted in this 2 x 2 x 3 design. After correcting the alpha value appropriately, the reported marginal p-value would no longer approach significance. It appears that Rauscher et al. (1996) did not, in fact, have any statistically significant results to support their influential claim that "speech with spatial content was less fluent when speakers could not gesture than when they could gesture" (ibid, p. 226).[1]

[1]Rauscher et al. (1996) also reported the effects of gesture prevention on two other measures of disfluency, speech rate and non-juncture filled pause rate, and interpreted their results on these two measures as support for the LRH. Space precludes a detailed re-examination of the speech rate and non-juncture filled pause rate results. However, our reexamination revealed that none of the de-

Several studies in the last five decades, before and after Rauscher et al. (1996), have also failed to find higher disfluency rates when speakers are prevented from gesturing than when they are allowed to gesture (Graham & Heywood, 1975; Rimé, Schiaratura, Hupet, & Ghysselinckx, 1984; Finlayson, Forrest, Lickley, & Beck, 2003; Hostetter, Alibali, & Kita, 2007; Hoetjes, Krahmer, & Swerts, 2014; Cravotta, Busà, & Prieto, 2018). Notably, none of the studies reporting null effects of gesture prevention distinguished between disfluencies during spatial and nonspatial speech; thus, arguably, these studies did not attempt to validate Rauscher et al.'s (1996) claim that gesture prevention selectively affects spatial speech. By contrast, our study tested this claim explicitly, but still found no evidence that gesture prevention increases disfluency rate—in either spatial or non-spatial speech. So, there appears to be no statistically significant evidence that preventing gesture makes speech more disfluent. Overall, we conclude that there is no clear support for the long-standing, influential claim that people gesture when they speak, in part, because gesturing helps speakers produce the right spatial words.

### Why do people gesture when they speak?

The LRH provided the first functional explanation for what causes a speaker to gesture that did not restrict co-speech gesture's function to communicative benefits for the listener, and instead suggested that gesturing may serve a cognitive function in the speaker's mind. Even though gesturing does not appear to help people find the right spatial words, gestures do serve a variety of other speaker-internal cognitive functions. For example, gesturing lightens speakers' cognitive load and also facilitates learning (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Cook, Mitchell, & Goldin-Meadow, 2008). As Krauss (1998) argued, gesture's function cannot be limited to communicative purposes. However, in light of the evidence we present here, an account of the cognitive functions of gesture will need to look beyond the role of gesture in helping people find the right spatial words.

## Conclusions

Do gestures facilitate speech production? Typically, reviews listing the cognitive functions of gesture start this list with the assertion that gesture helps people find the right words (e.g., Goldin-Meadow, 1999; Kita, Alibali, & Chu, 2017). Here, however, we showed that there is no compelling evidence to support this influential hypothesis. Rauscher et al.'s (1996) claim that gesture prevention increases disfluency rates for spatial speech is among the most widely cited empirical results in the gesture literature; yet, upon reexamining this study, we found that the data do not support the LRH. Our further reexamination of five decades of research testing effects of gesture prevention on disfluency, before and after Rauscher et al.'s (1996) influential study, revealed that there is no reliable evidence that gesture prevention makes speech more disfluent. Accordingly, the results from our study showed no statistically significant effects of gesture prevention on speech disfluency, for speech about literal or metaphorical space. Gestures do *not* appear to facilitate speech production by helping people find the right spatial words, challenging long-held beliefs about why people gesture when they speak.

pendent measures Rauscher et al.'s (1996) showed the full pattern of results predicted by the LRH.

## References


Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, *68*(3), 255–278.

Butterworth, B., & Hadar, U. (1989). Gesture, speech, and computational stages: A reply to mcneill.

Casasanto, D., & Bottini, R. (2014). Spatial language and abstract concepts. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(2), 139–149.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of verbal learning and verbal behavior*, *12*(4), 335–359.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058.

Cravotta, A., Busà, M. G., & Prieto, P. (2018). Restraining and encouraging the use of hand gestures: Effects on speech. In *Proceedings 9th international conference on speech prosody* (pp. 206–210).

Finlayson, S., Forrest, V., Lickley, R., & Beck, J. M. (2003). Effects of the restriction of hand gestures on disfluency. *Proceedings of Diss, Gothenburg Papers in Theoretical Linguistics*.

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

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological science*, *12*(6), 516–522.

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

Hoetjes, M., Krahmer, E., & Swerts, M. (2014). Does our speech change when we cannot gesture? *Speech Communication*, *57*, 257–267.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic bulletin & review*, *15*(3), 495–514.

Hostetter, A. B., Alibali, M. W., & Kita, S. (2007). Does sitting on your hands make you bite your tongue? the effects of gesture prohibition on speech during motor descriptions. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 29).

Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? the gesture-for-conceptualization hypothesis. *Psychological review*, *124*(3), 245.

Krauss, R. M. (1998). Why do we gesture when we speak? *Current directions in psychological science*, *7*(2), 54–54.

Krauss, R. M., & Hadar, U. (1999). The role of speech-related arm/hand gestures in word retrieval. *Gesture, speech, and sign*, *93*.

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive science*, *4*(2), 195–208.

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

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological science*, *7*(4), 226–231.

Rimé, B., Schiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion*, *8*(4), 311–325.

Yap, D., Brookshire, G., & Casasanto. (2018). Beat gestures encode spatial semantics. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*.