COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e13428

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13428

Embodying Similarity and Difference: The Effect of Listing and Contrasting Gestures During U.S. Political Speech

Icy (Yunyi) Zhang,[a] Tina Izad,[a] Erica A. Cartmill[a,b]

[a]*Department of Psychology, University of California, Los Angeles*
[b]*Department of Anthropology, University of California, Los Angeles*

Received 7 March 2023; received in revised form 20 February 2024; accepted 27 February 2024

Abstract

Public speakers like politicians carefully craft their words to maximize the clarity, impact, and persuasiveness of their messages. However, these messages can be shaped by more than words. Gestures play an important role in how spoken arguments are perceived, conceptualized, and remembered by audiences. Studies of political speech have explored the ways spoken arguments are used to persuade audiences and cue applause. Studies of politicians' gestures have explored the ways politicians illustrate different concepts with their hands, but have not focused on gesture's potential as a tool of persuasion. Our paper combines these traditions to ask first, how politicians gesture when using spoken rhetorical devices aimed at persuading audiences, and second, whether these gestures influence the ways their arguments are perceived. Study 1 examined two rhetorical devices—*contrasts* and *lists*—used by three politicians during U.S. presidential debates and asked whether the gestures produced during contrasts and lists differ. Gestures produced during contrasts were more likely to involve changes in hand location, and gestures produced during lists were more likely to involve changes in trajectory. Study 2 used footage from the same debates in an experiment to ask whether gesture influenced the way people perceived the politicians' arguments. When participants had access to gestural information, they perceived contrasted items as more different from one another and listed items as more similar to one another than they did when they only had access to speech. This was true even when participants had access to only gesture (in muted videos). We conclude that gesture is effective at communicating

The study design, hypotheses, and analytic plans were not preregistered. The study materials for Study 2 can be accessed through: https://osf.io/2hctd/?view_only=d44686645e4f4a0292c12a20fed3a569

Correspondence should be sent to Icy (Yunyi) Zhang, Department of Psychology, University of California, 502 Portola Plaza, Los Angeles, CA 90095, USA. E-mail: yunyi9847@g.ucla.edu

concepts of similarity and difference and that politicians (and likely other speakers) take advantage of gesture's persuasive potential.

Keywords: Gesture; Political speech; Persuasion; Embodied cognition

1. Introduction

Despite gesture being a ubiquitous human behavior (Morford & Goldin-Meadow, 1992), the impact of gesture on human cognition is still relatively underexplored. Gesture is often studied in conjunction with speech as a way to facilitate the speakers' speech production and support higher-level cognitive processes (Cocks, Sautin, Kita, Morgan, & Zlotowitz, 2009; Eielts et al., 2020; McNeill, 2000). But, of course, because speakers often gesture to communicate with others, gesture also influences how the information speakers convey is perceived (Peters & Hoetjes, 2017). For the recipient, gesture can reinforce, disambiguate, or add to information conveyed in speech (Goldin-Meadow, 2003; McNeill, 1992). And the information conveyed in gestures makes an impact on listeners. People perceive and remember the information in gesture (including information only in gesture) when listening to and watching speakers (Kelly, Özyürek, & Maris, 2010; McNeil, Alibali, & Evans, 2000).

For "professional speakers" like politicians, the communicative potential of gesture is of particular importance. Gesturing can impact perceptions of speaker qualities like trustworthiness (Grahe & Bernieri, 1999; Maricchiolo, Gnisci, Bonaiuto, & Ficca, 2009). Using particular types of gestures can also influence the way audiences see politicians (e.g., Joe Biden was criticized for being too aggressive when he used a forceful index finger point during the vice-presidential debate in 2012). In addition to influencing hard-to-define properties like trustworthiness and likability, gestures can also be crucial in helping politicians communicate clearly, memorably, and persuasively.

Our paper explores the ways politicians gesture when making particular types of arguments. We focus on two political rhetorical devices, contrasts and lists, shown previously to effectively elicit audience applause (Atkinson, 1983, 1984; Heritage & Greatbatch, 1986). We ask first how politicians gesture during these rhetorical devices, and second whether these gestures impact the way audiences perceive the politicians' arguments.

1.1. Literature review

Gesture has been shown to impact speech perception in, at least, three major ways: (i) enhancing listeners' comprehension and learning; (ii) conveying additional semantic information; and (iii) providing information about speakers' cognitive processes and attitudes.

1.1.1. Enhancing listeners' comprehension and learning

Gesture can impact how much people understand, remember, or learn from the information they hear in speech (Lin, 2021). Because hand movements provide another modality for the listener to process information, the multimodal input provided by gesture can support listeners' cognitive processing (Kartalkanat & Göksun, 2020). One possible mechanism explaining

this benefit is that seeing gestures decreases listeners' working memory load, which facilitates information processing (Dargue & Sweller, 2020). However, evidence is mixed as to whether the cognitive load is reduced when listeners receive multimodal input (for evidence that seeing gestures adds to cognitive load, see Moreno & Mayer, 2007; Rohrer, Delais-Roussarie, & Prieto, 2020; for evidence showing that gesture frees up the listeners' cognitive resources see Dargue & Sweller, 2020; Kartalkanat & Göksun, 2020). Another proposed mechanism is that co-speech gestures dynamically direct listeners' attention to specific elements of speech, which can help listeners more effectively notice and process propositional aspects of the speech (Lin, 2017; McNeill, 2005, 2006). For example, pointing gestures direct learners' attention to the object being described in speech (Pi, Hong, & Yang, 2017) and other representational gestures can selectively highlight subsets of the speech (Aussems & Kita, 2019). These gestures capture learners' attention and guide them to particular information in speech, which can then facilitate comprehension, memorization, and learning.

The idea that the speakers' gestures can impact others' perception of their messages is supported by research in embodied cognition. An embodied cognition perspective holds that cognition is deeply grounded in action and perception such that the addressee of the speaker receives information not only from speech but also from the speakers' actions, including gestures (Barsalou, 1999; Foglia & Wilson, 2013). A similar idea is emphasized in the conceptual metaphor theory, which posits that sensorimotor information is foundational for people to understand abstract concepts (Lakoff & Johnson, 1980; Yin, Su, & Ye, 2013). The speaker's gestures can express conceptual metaphors that link abstract ideas with sensorimotor experiences. The addressee leverages information from the speaker's gesture to build a connection with their own physical experience in the past and relies on this easily comprehensible information to parse out more abstract concepts. This is also consistent with the argument put forward by the perceptual symbols system theory: seeing bodily representations or performing actions activates neural connections that create simulations of perceptions and actions that have happened in the past (Barsalou, 1999). These theories are supported by neuroscience research on mirror neurons. When people first perform an action, a class of motor neurons called mirror neurons is activated. Then, seeing the same actions (or gestures symbolizing the actions) reactivates the viewer's same mirror neurons, which can impact the processing of information (Rizzolatti & Craighero, 2004).

Regardless of the exact mechanism(s), it is well-established that seeing gestures can enhance listeners' comprehension, memory, and learning (Dargue, Sweller, & Jones, 2019; Hostetter, 2011). For example, studies have shown that students learn more from speech accompanied by gesture than from speech alone (Perry, Berch, & Singleton, 1995; Wakefield, Novack, Congdon, Franconeri, & Goldin-Meadow, 2018). A meta-analysis done by Dargue et al. (2019) showed a benefit of observing gestures on comprehension (although the benefit is smaller than that of producing gestures oneself).

1.1.2. Conveying additional semantic information

When processing language from speakers who can be seen, listeners immediately process the gesture information through their visual cortex (Hubbard et al., 2012). Neurobiological research suggests that specific cortical regions are sensitive to human movements (e.g.,

Beauchamp, Lee, Haxby, & Martin, 2003), and these regions are more active when gesture and speech appear together than when speech occurs alone (Hubbard, Wilson, Callan, & Dapretto, 2009). Notably, not only does gesture add responses in cortical regions, activation is stronger when the gesture is *semantically unrelated* to speech, suggesting that listeners actively seek out meanings conveyed only in gestures (Dick, Goldin-Meadow, Hasson, Skipper, & Small, 2009).

1.1.3. Providing information about speakers' cognitive processes and attitudes

Co-speech gesture can contain information about speakers' cognitive processes and attitudes (Krauss, Chen, & Chawla, 1996). Gesture can influence the way people perceive the characteristics of a speaker, including traits like trustworthiness, persuasiveness, and competence (Grahe & Bernieri, 1999; Maricchiolo et al., 2009; Mehrabian & Williams, 1969). For example, Maricchiolo et al. (2009) found that participants self-reported paying more attention to speakers when they gestured. They further found that some gestures were more impactful than others; when speakers used deictic or representational gestures, participants perceived their speech to be more persuasive than when hearing speech alone.

1.1.4. Gesture in persuasive and political speech

Because gesture can influence the way people perceive both the gesturer and the content of their speech, it is of particular interest in the realm of political speech, where speakers are particularly concerned with delivering effective, persuasive messages, and how the audience sees them (Streeck, 2008). The gestures of political candidates are usually carefully designed and rehearsed to align with the intonation of their talking points and to metaphorically manipulate political concepts in space in order to convince the addressee of a particular opinion or sentiment (Lempert, 2011; Mendoza-Denton & Jannedy, 2011; Peters & Hoetjes, 2017). Gesture can be used by interlocutors to establish control during this type of argumentative or persuasive speech (Wehling, 2009, 2017). Considerations of gesture's effectiveness in making political arguments date back thousands of years: the earliest recorded writings about gesture are from an ancient Roman text on oration (Quintillian, 95 CE).

Despite the potential utility of gesture to influence audience perceptions during political speech, most of the studies about the ways politicians influence and persuade audiences have focused on speech. Politicians are thought to use particular kinds of phrasing—called rhetorical devices—to make points, influence listeners, and encourage audience applause (Atkinson, 1983, 1984; Heritage & Greatbatch, 1986). In a study of English politicians, Heritage and Greatbatch (1986) described eight types of rhetorical devices politicians used to structure their arguments and connect with audiences: contrast, list, puzzle-solution, headline-punchline, combination, position-taking, pursuit, and miscellaneous. These rhetorical devices range from relatively simple to highly complex and can occur within a single sentence or over multiple utterances. For example, position-taking involves a straightforward description of events or ideas followed by an evaluative statement that makes the speaker's position about those things clear. In this rhetorical device, the things that the speaker will take a stance for or against could appear within a single sentence, perhaps even a single clause, or they could stretch out across many sentences.

We focused our study on *contrasts* and *lists*, two of the more structurally simple rhetorical devices proposed by Heritage and Greatbatch. Contrasts contain a juxtaposition of ideas, which can be marked through syntax ("A not B"), semantics ("up and down"), or both ("forwards not backwards"). *Contrasts* make the message more salient because the central point is often stated twice, in one positive and one negative form (Heritage & Greatbatch, 1986). For example, in making an argument that not enough resources were given to disabled people, one speaker produced a contrast stating that "too much is spent on the munitions of war and too little is spent on the munitions of peace." The parallelism of the argument highlights the message and drives home the point.

Lists are similar in form to contrasts but are opposite in spirit. They involve several items clustered together and treated similarly in an argument (e.g., "I propose we do A, B, and C" or "People in America want A, B, C, and D"). Lists are similar in structure to contrasts, but involve three or more elements instead of two and conjunctive structures rather than disjunctive ones.

While Heritage and Greatbatch focused on speech in their analysis of applause-generating actions, they discussed gesture as one of five types of actions that could add "stress" to speech. They considered a spoken message stressed if (1) the speaker was looking at the audience at its completion, (2) the message was louder than surrounding speech, (3) the message had greater pitch or stress variation than surrounding speech, (4) it contained marked changes in tempo, or (5) it was accompanied by gesture. They categorized messages as no-stress (lacking any of the five features), intermediate-stress (displaying one of the five), or full-stress (displaying two or more). They found that full-stressed messages were the most likely to receive audience applause (Heritage & Greatbatch, 1986). The ability of gesture to stress political speech suggests that this nonverbal component of political speech has the potential to make a significant contribution to the perception and reception of political messages.

Despite significant interest in politicians' gestures (e.g., Cienki, 2004; Guilbeault, 2017; Hall, Goldstein, & Ingram, 2016; Lempert, 2011, 2017; Mendoza-Denton & Jannedy, 2011; Streeck, 2008) and the persuasiveness of political speech (e.g., Blumenau & Lauderdale, 2024; Charteris-Black, 2011; Hinnell, 2019), surprisingly only a few studies have explicitly examined whether gesture impacts the perception of persuasive arguments, with even fewer in political settings (Wehling, 2017). The few studies about gesture and persuasion are largely correlational (e.g., Bull, 1986; Burgoon, Birk, & Pfau, 1990; Hart & Winter, 2022; Zhang, Sherwin, Dmochowski, Sajda, & Kender, 2014). Hart and Winter (2022) analyzed gestures produced by the UK politician Nigel Farage in the context of anti-immigration political speech, and found that Farage used a range of gestures to purposefully advocate for more restrictive immigration legislation. For example, Farage used metaphoric gestures by opening up his hands horizontally to resemble an open-door immigration policy, while simultaneously conveying a sense of vulnerability to harm from immigration. In an electroencephalogram (EEG) study on viewers' response to politicians' gestures, Zhang et al. (2014) found that certain gesture features correlated with viewer engagement. When politicians used gestures that were particularly salient, unusual, or extreme, EEG measures of viewers' attentional engagement increased. Studies employing correlational methods enable scholars to unpack structural features of gesture that might be important for perception (Bressem, Ladewig, &

Müller, 2018; Ladewig & Bressem, 2013; Stokoe, 1980). However, since gesture viewing is not systematically varied, it is possible that the effect on perception might be driven by other variables.

A few experimental studies have explored the role of gesture on the perception of persuasive speech by manipulating gesture exposure or gesture form (e.g., Jackob et al., 2011; Peters & Hoetjes, 2017). For example, Maricchiolo et al. (2009) found that when people were given access to the actor's gestures when watching a video, they not only paid more attention to the gestures but also found the actor's arguments to be more persuasive. However, while these studies showed that gesture plays a role during persuasive speech, more experimental research is needed in the specific context of political speech. Moreover, given the wide use of spoken rhetorical devices, past research on gesture in persuasive speech does not connect to work on the use of spoken rhetorical devices.

One recent study by Hinnell (2019) looked at gesture in the context of spoken contrasts, though it did not focus on political speech. Using the Red Hen multimedia archive, Hinnell found that when North American English speakers contrasted things in speech, they often embodied the contrast by gesturing first to one side and then to the other. Many of these gestures crossed the sagittal line of the speaker's body (i.e., moving from one side of the body to the other). Gestures were predominantly manual, though contrasts also appeared in other parts of the body (e.g., head tilts, body lean, and hand movements). A study on gesture in political speech (though not on rhetorical devices) revealed the importance of perceived alignment with the politician's position on the role of gesture (Peters & Hoetjes, 2017). When college students did not think a politician's message was relevant to them, the politician's gestures played a bigger role in whether they perceived the speaker as lively, interesting, and sincere. But, when students saw a politician's message as personally relevant, the politician's gestures had less impact on their perception of the speaker.

While few studies have paid attention to the use of gesture in rhetorical devices, even fewer studies have combined observational and experimental designs to first examine where gestures are occurring naturally, and then whether those gestures are effective at communicating ideas with others. We believe that this mixed methods approach is important (1) to understand *how* co-speech gestures are used by speakers when conveying different ideas, and (2) to evaluate *whether* these gestures influence listeners' perception.

Our work aimed to extend Heritage and Greatbatch's study of political rhetorical devices to reveal the use and effect of gesture in political speech. Many of the rhetorical devices Heritage and Greatbatch proposed involved the speaker comparing or juxtaposing different elements. These multipart speech acts seemed perfectly poised for the addition of gesture to underscore, highlight, or offset different conceptual elements.

Among the eight types of rhetorical devices, we chose contrasts to study the impact of gesture on perception because contrasts typically occur within a single sentence and involve comparing two immediately adjacent ideas (unlike some of the other rhetorical devices like headline-punchline or puzzle-solution, which involve more complex linguistic structures and often unfold over several sentences). In contrasts, the core assertion is stressed twice: once in a positive form, and once in a negative form (Heritage & Greatbatch, 1986). This natural contrast is conceptualized as spatial in nature by theories of embodied cognition and

conceptual metaphor (Barsalou, 2008; Lakoff & Johnson, 1980). Past research shows preliminary evidence that people represent items closer together in space if they see them as more associated with each other, and farther apart if they see them as less associated with each other (Zhang, Fu, Xu, & Lu, 2022). Hinnell (2019) highlight the inherently embodied nature of contrasts—speakers often embody the contrast by gesturing first to one side and then to the other. But one must ask whether this spatial relationship is perceived by the addressee and whether it influences their understanding of conceptual ideas and arguments. Because contrasts frequently present two elements in opposition to one another (e.g., "I will do X, not Y" or "X doesn't meet the goal, Y does"), they seem like an ideal place to explore the role of gesture to spatially embody the difference between the contrasted elements.

We chose lists as a secondary device to serve as a foil to contrasts. Lists and contrasts both present the opportunity for gestures to accompany and highlight multiple compared elements adjacent to one another in speech. However, the goal differs between the two devices: contrasts emphasize difference/discontinuity, and lists highlight similarity/continuity.

We designed our study to ask *how* politicians gesture during different rhetorical devices, and *whether* these gestures impact how the politicians' arguments are perceived by listeners. We chose debates for our samples of political speech because debates retain many characteristics of political speech-making (e.g., rehearsed talking points) but also incorporate some elements of spontaneous conversation (e.g., questions from the audience). During televised debates, candidates may employ many means to convince the audience of their argument, which makes these debates excellent resources for studying the occurrence and impact of co-speech gestures.

Study 1 adopted a correlational design to examine the contrasts and lists rhetorical devices used by U.S. presidential candidates during debates and asked whether the candidates gestured differently during spoken contrasts and lists. Following classic work in sign language by Stokoe (1980) and gesture by McNeill (1992), we focused our coding on three features of the presidential candidates' gestures: trajectory, location, and handshape. Our selection of these three features was consistent with the work by Ladewig and Bressem (2013).

We hypothesized that gestures produced during contrasts and lists would differ from one another in their use of these three features. We expected that gesture would spatially represent the metaphors of dissimilarity and similarity present in contrasts and lists. Specifically, given that politicians employ contrasts to signal difference/discontinuity and lists to highlight similarity/continuity, we predicted that gestures during contrasts would have more discontinuous features (i.e., changes in trajectory, handshape, or location) than those made during lists.

Study 2 asked whether the gestures politicians produced during contrasts and lists influence the way people perceive the spoken arguments. We conducted an experiment in which clips of politicians producing contrasts and lists were presented in three conditions: *speech-only* (audio), *gesture-only* (muted video), and *speech+gesture* (video with sound). Based on past research showing that gesture can influence the perception of political speech (e.g., Maricchiolo et al., 2009; Peters & Hoetjes, 2017), we predicted that gestures would enhance people's awareness of the differences or similarities between items. In other words, we expected that participants in the *speech+gesture* group would rate contrasted items as more different and listed items as more similar than the *speech-only* group. We also predicted that without access

to the semantic information in speech, the *gesture-only* group would not show as strong an effect of condition (contrast vs. list) than the speech+gesture group.

Together, these studies help to deepen our understanding of the ways speakers use co-speech gestures to illustrate the relationships between ideas in speech and whether these gestures influence the ways these relationships are communicated to others. This work builds on broad academic interest in the hand movements of politicians, connecting studies of gesture to interdisciplinary interest in fields, such as multimodality, speech perception, embodied cognition, and language and thought.

2. Study 1

2.1. Methods

2.1.1. Participants

We analyzed speech and gesture from Hillary Clinton, Barack Obama, and Mitt Romney speaking in nationally televised presidential debates from the 2008, 2012, and 2016 general elections. Two debates were used for each politician: Obama 10/7/2008, Obama 10/16/2012, Clinton 9/26/2016, Clinton 10/9/2016, Romney 10/3/2012, Romney 10/9/2012. As Barack Obama ran for office twice, one debate from each election was used, while two debates from the same election cycle were used for Hillary Clinton (2016) and Mitt Romney (2012). Transcripts for each of the debates were obtained from the Commission on Presidential Debates (debates.org), a free and public resource. The original broadcast sources of the debates included the New York Times, C-SPAN, NBC, and PBS news. Gesture was coded from videos of the broadcast debates, downloaded from YouTube. Examples of politicians' gestures can be seen in the video stimuli we prepared for Study 2, which were cut from the longer debate videos: https://osf.io/2hctd/?view_only=d44686645e4f4a0292c12a20fed3a569.

2.1.2. Speech coding

Spoken *contrasts* and *lists* were identified from transcripts of presidential debates downloaded from debates.org. Prior to coding, all researchers read "Generating Applause: A Study of Rhetoric and Response at Party Political Conferences" by Heritage and Greatbatch (1986) to become familiar with the definitions of different political rhetorical devices. The coding team consisted of three undergraduate research assistants and the principal investigator. The coding was conducted using only the transcript, without listening to the audio or viewing the video. Only the principal investigator had seen the video before the coding and all other coders were accessing the content for the first time. Coders used criteria and examples from Heritage and Greatbatch to identify *contrasts* and *lists* from the transcripts (as defined below). Each transcript was coded independently by two coders. All contrasts and lists were marked as clear/strong or marginal examples, according to how closely they matched the definition and examples. If both coders working on a transcript agreed that an example was clear/strong, it was included in the dataset and not discussed further. If either coder had marked an example as marginal (or it was identified by only one coder), it was presented and discussed at weekly

group meetings to decide if the utterance should be included as a contrast or list. These group meetings consisted of all three coders and the PI (EC). An example of a marginal case might involve an utterance containing the structural elements of a contrast (described below), but spread out across several sentences. After discussing the example as a group, marginal cases were included in the data only if all three coders and the PI agreed they should be included.

Contrasts were defined as two juxtaposed ideas whose differences were highlighted through either syntax or, less frequently, semantics. For example, "not last week but today" would be a contrast marked through syntax, while "up and down" or "apples and oranges" would be marked through semantics. These distinctions followed examples given in Heritage and Greatbatch (1986) of the different ways contrasts could be built. *Lists* were defined as three or more related ideas, words, or phrases listed consecutively where the speaker had the same stance to all three (e.g., liking or disliking all). Lists could have more than three elements but had to have at least three. Table 1 gives examples of spoken contrasts and lists included in our dataset. The contrasted or listed elements in the utterances were underlined to aid subsequent gesture coding. Unusually long utterances and those with embedded clauses were typically excluded from the dataset since the rhetorical device of interest was not particularly salient. For example, an utterance that presented two contrasting ideas stretching over several sentences would have been excluded since the contrast was unlikely to be salient to listeners. We aimed to create a dataset consisting only of clear, salient contrasts and lists, rather than the most inclusive one possible.

The final dataset contained 123 *contrasts* (two contrasting things, ideas, or topics) and 109 *lists* (three or more noncontrasting things, ideas, or topics) drawn from transcripts of the six debates. List and contrast categories were mutually exclusive in our dataset. If an utterance contained both a list and a contrast, it was not included.

2.1.3. Gesture coding

Once the utterances containing contrasts and lists were identified, the accompanying video was studied in order to code any gestures that accompanied the spoken rhetorical devices. We coded gestures that temporally overlapped with the contrasted and listed elements of the media clips (i.e., those underlined in the speech coding). Temporal overlap was defined as having at least one gesture occur during a contrasted or listed item (underlined text in Fig. 1). If a gesture was made earlier or later in a sentence but did not overlap with a contrasted or listed item, it was not coded. Fig. 1 shows examples of some of the gestures made during contrasts and lists.

For each spoken element in a contrast or list, we coded first whether it was accompanied by a gesture. There are many definitions of gesture, which can include movements of the entire upper body, such as head tilts and hand gestures (Parrill & Stec, 2017; Wehling, 2009). In Study 1, we focused specifically on hand gestures, defined as symbolic movement of the hands and arms in order to accompany the speech (Peters & Hoetjes, 2017). We did not exclude any types of gesture, in order to cast a wide net on gestures that might accompany contrasts and lists, and to maximize the ecological validity of the findings. Then, we coded whether the gesture displayed any changes in *handshape*, *location*, and *trajectory* from the preceding gesture. In determining whether there was a change, we used the first gesture during

Table 1
Examples of spoken contrasts and lists from our corpus of presidential debates

Rhetorical device	Speaker	Utterance	Contrasted/listed items
Contrast	Barack Obama	…you've got to have somebody in Washington thinking about the middle class and not just those who can afford to hire lobbyists	1. middle class 2. those who can afford to hire lobbyists
	Mitt Romney	I want regulators to see their job as encouraging small enterprise, not crushing it.	1. encouraging small enterprise 2. crushing enterprise
	Hillary Clinton	Broad-based, inclusive growth is what we need in America, not more advantages for people at the very top	1. broad-based inclusive growth 2. more advantage for people at the very top
List	Barack Obama	…folks who are criminals, gang bangers, people who are hurting the community	1. criminals 2. gang bangers 3. people who are hurting the community
	Mitt Romney	…the revenue I get is by more people working, getting higher pay, paying more taxes	1. more people working 2. getting higher pay 3. paying more taxes
	Hillary Clinton	…because those coal miners and their fathers and their grandfathers…	1. coal miners 2. fathers 3. grandfathers

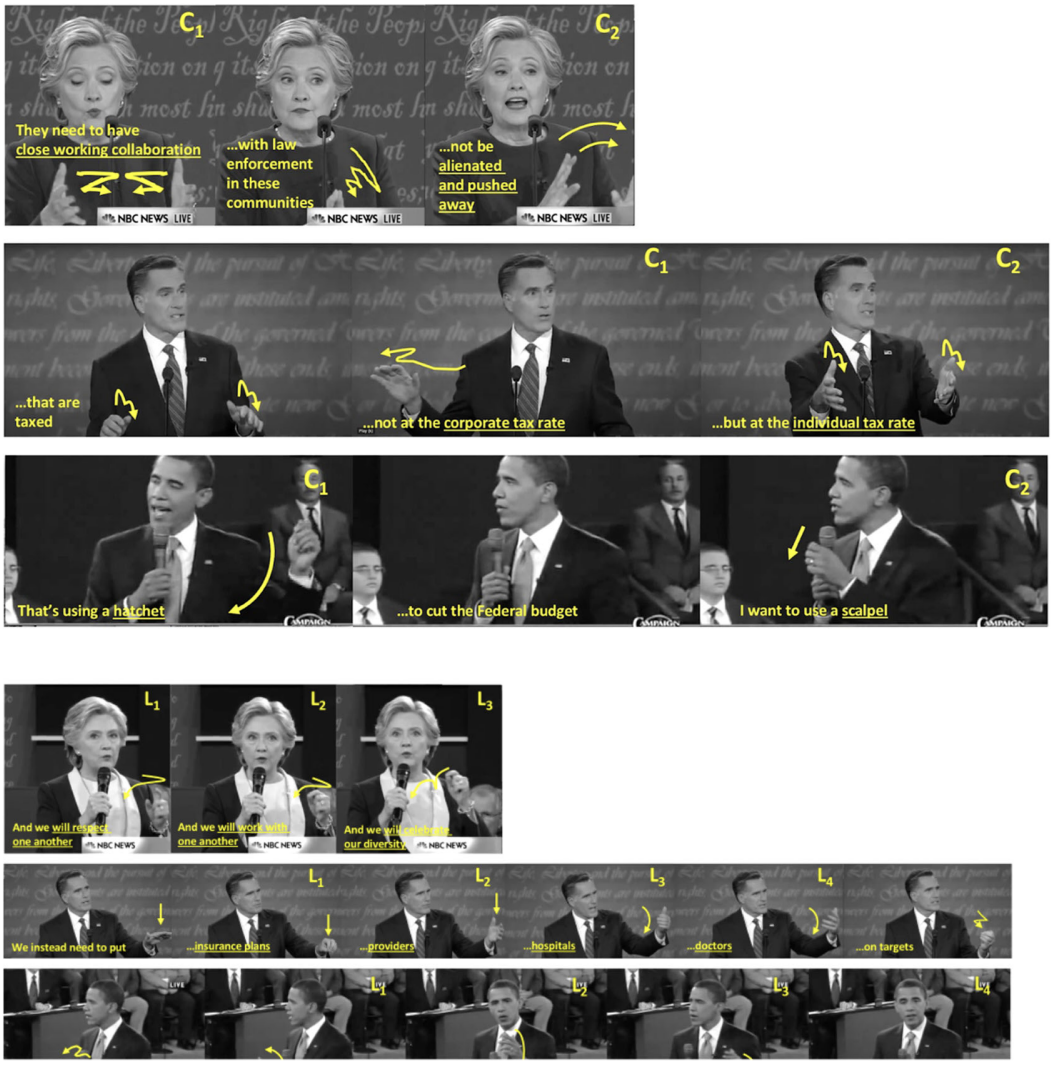

Fig. 1. Examples of gestures made during contrast and list rhetorical devices.

Note. The contrasted or listed concepts are underlined in the corresponding text and are marked as C_1 and C_2 or $L_{1\text{-}n}$ on the top right of the image panels. Arrows show the movement path of gestures.

a contrasts/listed element as the baseline and did not consider gestures made before the start of the rhetorical device. *Handshape* was defined as the position of the fingers, wrist, and palm, so a change in handshape could involve an index finger point changing to a whole hand point. *Location* was defined as the relative position of the hand in relation to the body, so a change in location could involve a fist held close to the body and then extended away from it. *Trajectory*

Table 2
Number of contrast versus list gestures that showed a change versus no change in terms of trajectory, location, or handshape

	Trajectory		Location		Handshape	
	Contrast	List	Contrast	List	Contrast	List
Change	28 (22.6%)	55 (50.5%)	74 (60%)	30 (27.5%)	41 (33.1%)	35 (32.1%)
No change	96 (77.4%)	54 (49.5%)	50 (40%)	79 (72.5%)	83 (66.9%)	74 (67.9%)

was defined as the path traced through the air by a gesture, so a change in trajectory could involve a flat hand circling and then changing to a linear path.

Changes in handshape, location, and trajectory were not mutually exclusive and many gestures accompanying contrasts and lists displayed changes in two or more variables. Because the clips included situations both when the speaker was holding a microphone (typical in town hall debates) and when the speaker was not, we did not include the change from unimanual to bimanual gesture or the switch of the gesturing hand as their own variables. However, if a speaker were to switch from one hand to another, it would be captured either by a change in location or by changes in both location and trajectory. This decision was consistent with the coding approach used by Ladewig and Bressem when they coded naturalistic clips (2013). Changes in gesture variables were not quantitatively determined by a measurement like estimated distance traveled or number of pixels changed. Instead, they were categorized by two or more gesture coders working independently. This human-driven categorization process allowed for the incredible diversity of postures, framing, and distance to the speaker in the video clips.

The same three trained research assistants who had coded the speech also conducted the gesture coding. The coders were given a coding manual, video examples, and in-person training before beginning to code. Coders were assigned to politicians in pairs so that each clip was coded by two research assistants. As with the speech coding, any discrepancies in coding were presented in a weekly group meeting for discussion. If the coders could not come to an agreement on how to categorize an example during the group meeting, the example was excluded from the dataset. Because there was no statistical calculation of reliability for the initial three coders, one of the authors (TI) coded 20 randomly selected video clips to assess the accuracy of the coding. This fourth coder coded the clips independently, with no opportunity to discuss the video with the original coders, and agreed with the initial gesture coding 95% of the time.

2.2. Results

Table 2 shows the number of contrast and list gestures that displayed changes in trajectory, location, and handshape. Because there were three speakers (Obama, Clinton, and Romney), we first adopted a random intercept model with a random intercept for the speaker to test whether there was a significant random effect. The model found an intraclass correlation coefficient of 0 ($\tau_{00\ \text{speaker}\cdot} = 0$), which informed our decision to use a fixed effects model

instead. Thus, we ran a logistic regression with a fixed effect of rhetorical device and speaker to analyze the relationship between change in gesture trajectory and the type rhetorical device associated with the gesture (contrast vs. list). We found that, controlling for the speaker, the odds of having a change in gestural trajectory *increased* by 322% for the list type compared to the contrast type (odds ratio = 4.22, 95% CI = [2.32, 7.89]). The result indicated that the list gestures showed significantly more changes in trajectory than contrast gestures ($p < .001$).

Then, to explore whether list and contrast differed in change in location, we performed a logistic regression with a fixed effect for the rhetorical devices associated with gesture and a fixed effect for the speaker. The analysis showed that controlling for the speaker, the odds of having a change in gesture location *decreased* by 72% for the list gesture compared to the contrast gesture (odds ratio = 0.28, 95% CI = [0.15, 0.49]). The result indicated that gestures that accompany contrast showed significantly more changes in location than gestures that accompany list ($p < .001$).

Lastly, a logistic regression with fixed effects for the rhetorical device and speaker showed no significant correlation between the type of gesture and change in handshape (odds ratio = 1.17, 95% CI = [.65, 2.12], $p = .607$).

Given that both trajectory and location can change in one gesture, we performed the same logistic regression, but using a binary outcome of whether there was a change in both trajectory and location or not (which included neither changing as well as only one changing). We found that, controlling for the speaker, the odds of having a change in both gesture trajectory and location *increased* by 102% for the list gesture compared to the contrast gesture (odds ratio = 2.02, 95% CI = [1.02, 4.03]). The results indicated a significant correlation between the type of rhetorical gesture and change in a combination of trajectory and location ($p = .044$) such that the list type was correlated with higher odds of change in this combination.

Lastly, we also pooled the gesture variables (trajectory, location, and handshape) and asked whether list and contrast gestures differed in their likelihood of (1) changing at least one variable; (2) and changing all three variables. A logistic regression showed that, controlling for the speaker, the odds of having at least one variable change *decreased* by 58% for list gestures compared to contrast gestures (odds ratio = 0.42, 95% CI = [0.23, 0.77], $p = .005$). However, controlling for the speaker, the odds of having all three variables change *increased* by 209% for list compared to contrast (odds ratio = 3.09, 95% CI = [1.26, 8.15], $p = .017$).

2.3. Discussion

We found that gestures made during contrasts and lists differed in how likely they were to display changes in trajectory and location, but not handshape. Gestures made during lists showed more changes in trajectory and gestures made during contrasts changed more often in location. The findings of Study 1 support our hypothesis that gestures produced during contrasts would have different features than gestures produced during lists.

One limitation is that while we observed differences in gesture features, we cannot draw conclusions about the mechanisms underlying these differences. When we grouped all gesture variables together and looked for an overall tendency to change, we found that gestures made during contrasts were more likely to change in at least one variable than gestures made during

lists were, but that gestures during lists were more likely to change in all three variables. Our analysis did not explicitly explore the variation in which hand was utilized or the shift from unimanual to bimanual gestures during the speech (though these changes would have been captured in our coding). Considering that past literature suggests variation in hand use is meaningful in signaling the speaker's preferences (e.g., the speaker tends to use the dominant hand things they feel positive about; Casasanto & Jasmin, 2010), future studies might consider using clips where both hands of the speaker are available to explore such variation and shift in which hand was utilized.

To our knowledge, this is the first observational study that examines how politicians' co-speech gestures differ when they are discussing contrasting ideas versus similar ideas. However, we did not know whether the differences in gestures during contrasts and lists that trained coders perceived would be noticed by naïve observers, and whether they would impact the way observers perceived the concepts the politicians were discussing. We designed Study 2 to address those questions, asking whether gestures produced during contrasts highlighted difference/discontinuity and gestures produced during lists highlighted similarity/continuity for people viewing the gestures. Specifically, we asked whether people would perceive concepts as more similar if list gestures were seen and more different if contrast gestures were seen.

3. Study 2

In order to investigate whether the co-speech gestures produced during spoken contrasts and lists influenced the qualities people attribute to contrasted or listed items, we designed an experiment in which people were asked to rate similarities or differences while listening to speech and/or viewing gestures. Because people gesture with many body parts—including hands, head, and even eyebrows—in Study 2, we followed norms in gesture studies and adopted a broad definition of gesture that includes all upper body movements (e.g., Neff et al., 2010; Wehling, 2009). Adopting this comprehensive definition of gesture allowed us to examine speech perception in a more ecologically valid setting and explore the effect of this multimodal input on speech perception.

Gesture can have a significant impact on the processing and interpretation of accompanying speech (Kelly, Barr, Church, & Lynch, 1999; McNeill, Cassell, & McCullough, 1994; Neff et al., 2010; Özyürek, Willems, Kita, & Hagoort, 2007; So, Sim Chen-Hui, & Low Wei-Shan, 2012). For example, addressees automatically use speakers' gestures to process information in the discourse (Sekine & Kita, 2017). However, many past studies have focused on how gestural input impacts addressees' perception of concrete entities (e.g., where characters in a story were or how they acted), which already have physical spatial features that can map onto gesture. More recently, scholars started to consider the ways in which gesture can impact the perception of abstract entities, such as arguments or conceptual ideas (Hinnell & Parrill, 2020; Parrill & Stec, 2017; Parrill et al., 2022). For example, Hinnell and Parrill (2020) presented participants with videos of speakers contrasting two ideas each accompanied by a gesture to the right or the left (e.g., "My little brother's not on Facebook because he thinks it's a waste

of time" + gesture to right "but my other brother says he can't do job networking without it" + gesture to left). The speakers then expressed a preference or alignment with one of the ideas using an ambiguous pronoun (e.g., "I agree with him"). One group of participants saw a gesture to the right or left along with the statement, while the other group did not. Participants who saw gesture used that information to disambiguate the referent, using the gesture to refer back to the side where one argument was presented.

Research from the area of mathematics learning suggests a closely intertwined relationship between gesture exposure, perception, and comprehension (e.g., Cook & Goldin-Meadow, 2006; Cook, Duffy, & Fenn, 2013). Exposure to gesture can facilitate the processing of abstract concepts and help people grasp information that they would not otherwise comprehend without the gesture (Cook & Goldin-Meadow, 2006). For example, Cook and Goldin-Meadow (2006) found that when children were learning the concept of mathematical equivalence, adding gesture that highlighted both sides of the equal sign to spoken instruction improved their understanding of mathematical equivalence. Children were able to pick up on the nonverbal information conveyed in the instructor's gesture and leverage that information to help themselves produce more correct solution strategies in gesture, which in turn benefited learning.

We build on work showing that the instructor's gesture can impact students' learning and that the speaker's gesture can impact learners' interpretation of speech, to ask whether gesture influences the way statements comparing or contrasting abstract entities during political speech are perceived. In particular, we ask whether gesture is effective at conveying the concepts of similarity (lists) or difference (contrasts) on its own, or whether it plays a more supporting role to the information conveyed in speech.

3.1. Methods

We conducted an experiment with one between-subjects variable, *condition* (speech-only, gesture-only, or speech+gesture) and one within-subjects variable, *rhetorical device* (contrast or list). Participants were randomly assigned to one of the three conditions. They watched or listened to a combination of list and contrast media clips presented as audio only (speech-only), video with no sound (gesture-only), or video with sound (speech+gesture). After each clip, participants were asked to rate how similar or different the ideas discussed in the clip were to one another. This may have seemed like a strange task in the gesture-only condition when they could not hear what concepts the politician was discussing, but all participants were warned that they might see images without sound, and all participants completed the task.

3.1.1. Participants

One hundred and sixty-eight undergraduate students from a large public research institution were recruited to participate in Study 2 via their enrollment in psychology courses. They were compensated with extra credit for their course. There were no other forms of compensation. Participants who did not pass a validation question at the beginning of the survey or an attention check question placed in the middle of the study were removed from

Table 3
Descriptive summary of video stimuli

	Trajectory		Location		Handshape		Headshake	
	Contrast	List	Contrast	List	Contrast	List	Contrast	List
Change (presence for headshake)	9 (32.1%)	5 (41.7%)	25 (89.3%)	1 (8.3%)	11 (39.3%)	3 (25.0%)	6 (21.4%)	6 (50.0%)
No change (no presence for headshake)	19 (67.9%)	7 (58.3%)	3 (10.7%)	11 (91.7%)	17 (60.7%)	9 (75.0%)	22 (78.6%)	6 (50%)

the analysis. The final sample consisted of 148 participants. Of these, 27 self-identified as male, 116 as female, 4 as other, and 1 preferred not to say. Participants formed a relatively diverse group, with 67 participants identifying as Asian (45%), 41 White (28%), 7 Black or African American (5%), 12 multiracial (8%), and 21 who preferred not to say (14%). In terms of ethnicity, 32 participants (22%) indicated they were of Hispanic or Latino origin.

All participants were fluent in English (1 identified as somewhat weak, 10 as neither strong nor weak, 30 as somewhat strong, and 107 as very strong). We asked participants about their political orientation in relation to U.S. political parties: 97 identified as Democrat, 5 as Republican, 5 with other parties, 23 did not identify with a party, and 18 said that they were not familiar with the political parties in the United States. We also asked participants about their handedness: 123 identified as right-handed, 23 as left-handed, and 2 as ambidextrous.

3.1.2. Materials

We selected 40 media clips from the video corpus in Study 1 as stimuli for Study 2. These stimuli consisted of 28 contrasts and 12 lists ranging between 2 and 15 s (mean 6.7). Videos were selected randomly from videos in which (1) the hands of the politician were clearly visible and (2) the camera did not zoom suddenly or cut away during the utterance. Whenever possible, video clips were cropped to remove potentially distracting elements like other politicians, audience members, and news banners. Most video clips depicted the politician from the waist up with their hands visible throughout the clip. All of the videos contained gestures. Although it would be nice to isolate the effect of manual gesture from other upper body movement and facial expressions, we chose to not blur the face of the speaker because we wanted the stimuli to appear as naturalistic as possible and we did not want to draw their attention to the fact that we were interested in gestures specifically. To provide a descriptive summary of the videos, we documented whether each clip showed a change in handshape, location, and trajectory, and whether a head shake occurred to express negation (see Table 3). The gesture variables in the experimental list and contrast videos roughly matched the trend in Study 1, with contrast videos having more changes in location and list videos having more changes in trajectory. Because the videos were chosen randomly from those in Study 1, the contrast videos also appeared to have more changes in handshape. Lastly, we also coded whether the speaker was holding a microphone in one hand. Out of the 28 contrast videos, 18 speakers (64.3%) were holding a microphone. Out of the 12 lists, seven speakers (58.3%) were holding a microphone.

3.1.3. Design/procedure

Participants were recruited online via the institution's Psychology Department subject pool. Participants who wished to participate clicked a link to complete a Qualtrics survey (Qualtrics, 2022). Upon clicking on the link, participants were randomly assigned to one of three conditions: speech-only, gesture-only, or speech+gesture condition.

First, participants were presented with an introduction including instructions to make sure that they were within an environment with few distractions, in which they could listen to the media clips. Participants were told that they would be viewing 40 clips of politicians making arguments from past political debates and were told, based on their condition, "these clips will include video without audio / video with audio / audio only." They were also told they could play each clip as many times as they liked and that they would then be asked to mark how similar or different "the two or more things, concepts, or ideas" each politician was talking about were on a 4-point scale.

Participants were presented with the comprehension question "What will you be asked to do in this study?" along with four possible answers. If they did not choose the correct answer, they were asked to read the instructions more carefully and answer again. To ensure that participants were using browsers and devices that would display videos in their entirety, participants were shown a still image with a dot in each corner and asked how many dots were visible. All participants reported viewing four dots.

Participants were then shown 40 media clips one after another along with the following instructions: "Please rate how similar the topics that the politician discussed are. If you are not sure, please make your best guess." Participants made a selection on a 4-point scale: "Very Similar," "Somewhat Similar," "Somewhat Different," or "Very Different." In the middle of the sequence of 40 clips, participants were presented with an attention check instructing participants to select a particular answer. Nineteen participants failed this check and were excluded from the analysis.

At the end of the study, participants completed a self-administered demographic questionnaire. Their responses did not affect their participation credit. Participants were asked about their age, gender, race, ethnicity, political affiliation, handedness, and English proficiency. They were also asked to indicate if they were distracted at any point during the study.

3.1.4. Measures

Our dependent measure was *similarity rating*. Participants rated how similar or different they thought the concepts in the clip they listened to or watched were. Similarity ranged from very different (0) to very similar (3). Our independent variables were *condition* (speech-only, gesture-only, speech+gesture) and *rhetorical device* (list vs. contrast). The distribution of the three conditions' similarity ratings is shown in Fig. 2. We also measured on a binary scale whether the speaker signaled negation through headshake in each of the video clips.

3.1.5. Statistical analysis

To investigate the difference in similarity ratings between the three conditions by rhetorical device type (contrast vs. list), we planned to fit and compare two linear mixed effects models

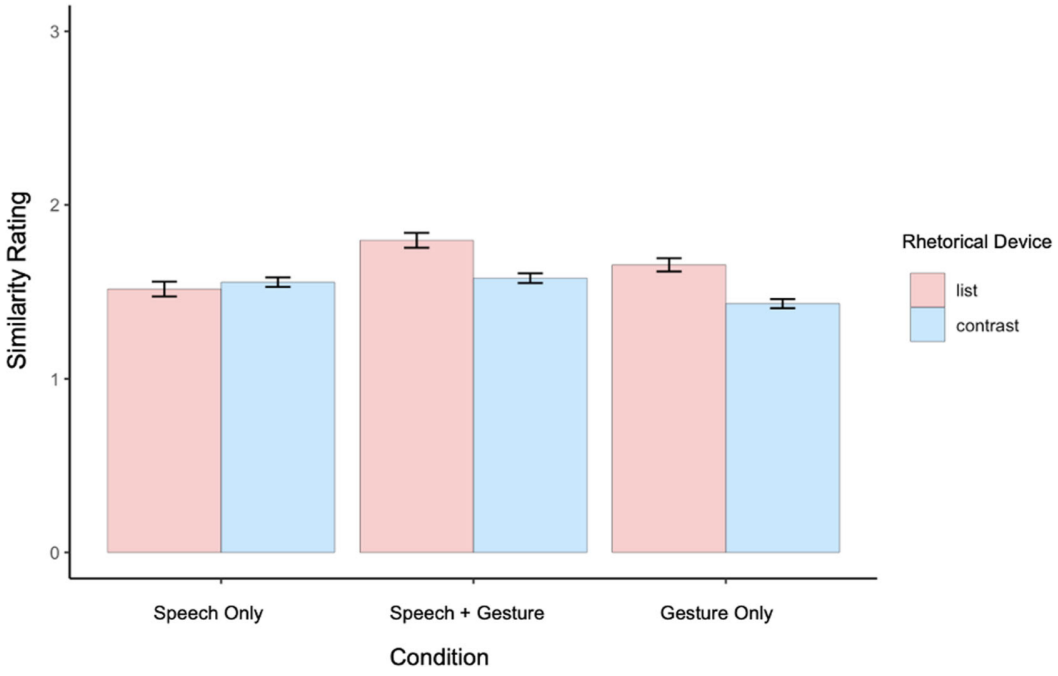

Fig. 2. Average similarity rating of contrasts and lists by condition.

Note. The 4-point rating scale ranged from 0 (very different) to 3 (very similar). A value of 1.5 would indicate the items were not rated as either similar or different (though this was not an option on the rating scale).

in order to decide which model would perform better in explaining the effect of condition and rhetorical device on people's perception of similarity.

The first model, named the random-intercept model, was a linear mixed effects model with similarity rating as the outcome variable and fixed effects of condition, rhetorical device, and their two-way interaction. We also included random intercepts for participants and items. The pseudocode for this model is: similarity ~ condition* rhetorical_device + (1|subject_id) +(1|item).

The second model, named the random-intercept-and-slope model, was the same as the first but included random slopes for rhetorical devices by participants, and a random intercept for items. The pseudocode for this model is similarity ~ condition* rhetorical_device + (1+rhetorical_device|subject_id) +(1|itemquestion_id).

Analyses of the two models were done using the lme4 package (v1.1-34; Bates, Mächler, Bolker, & Walker, 2015). We planned to compare the two models and adopt the model that performed better in explaining the variation of the data. After analyzing the fixed effects of the selected model, if there was not a significant interaction between condition and rhetorical device, we would analyze the main effects by comparing the marginal means; if there were a

Table 4
Fixed effects of modality and rhetorical device using random intercept and random slope

	Estimate	Standard error	df	*t*-value	Pr(>|t|)
(Intercept)	1.51	0.09	92.42	16.00	< .001***
ConditionSpeech + Gesture	0.29	0.09	148.08	3.28	.001**
ConditionGesture Only	0.15	0.09	148.05	1.67	.096
Rhetorical Device (List as Reference)	0.05	0.11	87.43	0.44	.660
Speech + Gesture:Contrast	−0.27	0.10	147.74	−2.63	.009**
Gesture Only:Contrast	−0.27	0.10	147.72	−2.70	.008**

significant interaction, we would perform simple effect analysis by comparing the cell means (Graham, 2000).

3.2. Results

One hundred and forty-eight participants completed the experiment. They were more or less evenly distributed across the three experimental conditions: speech-only ($N = 51$), gesture-only ($N = 48$), and speech+gesture ($N = 49$). A chi-square test of independence showed that there was no significant difference between the three conditions in terms of participants' race, X^2 (12, $N = 148$) = 10.25, $p = .594$, English proficiency, X^2 (6, $N = 148$) = 7.69, $p = .262$, or handedness, X^2 (4, $N = 148$) = 2.75, $p = .600$.

3.2.1. Video stimulus check

To investigate whether the use of negation through head shake might be different in contrast and list videos, we conducted a logistic regression controlling for the fixed effect for the speakers and found that the odds of having headshake increased by 308% for list compared to contrast (odds ratio = 4.08, 95% CI = [0.93, 19.94]), indicating that this increase was not statistically significant ($p = .067$).

Mean values for Contrast items were: Speech-only = 1.55(±1.04), Speech+Gesture = 1.58(±1.04), Gesture-only = 1.43(±0.97). Mean values for List items were: Speech-only = 1.51(±1.04), Speech+Gesture = 1.79(±1.04), Gesture-only = 1.65(±0.92).

Fig. 2 shows the mean similarity rating for participants in each condition for contrast and list stimuli. To investigate the difference in similarity ratings between the three conditions by rhetorical device type (contrast vs. list), we fitted two linear mixed effects models.

We conducted a model comparison to determine which of the two models was a better fit. The random-intercept-and-slope model, which included a random slope of the rhetorical device for each participant, explained significantly more variation than the random-intercept model with only two random intercepts (X^2 (2, 5898) = 100.88, $p < .001$). We report only the output of the random-intercept-and-slope model; the random-intercept model can be found in Appendix A. (Note: The results of both models are consistent with each other.)

Table 4 shows the results of the random-intercept-and-slope model. Compared to the Speech Only group's rating of List items, similarity ratings of List items increased by 0.29 for the Speech+Gesture group ($p = .001$) and by 0.15 for the Gesture Only group ($p = .096$).

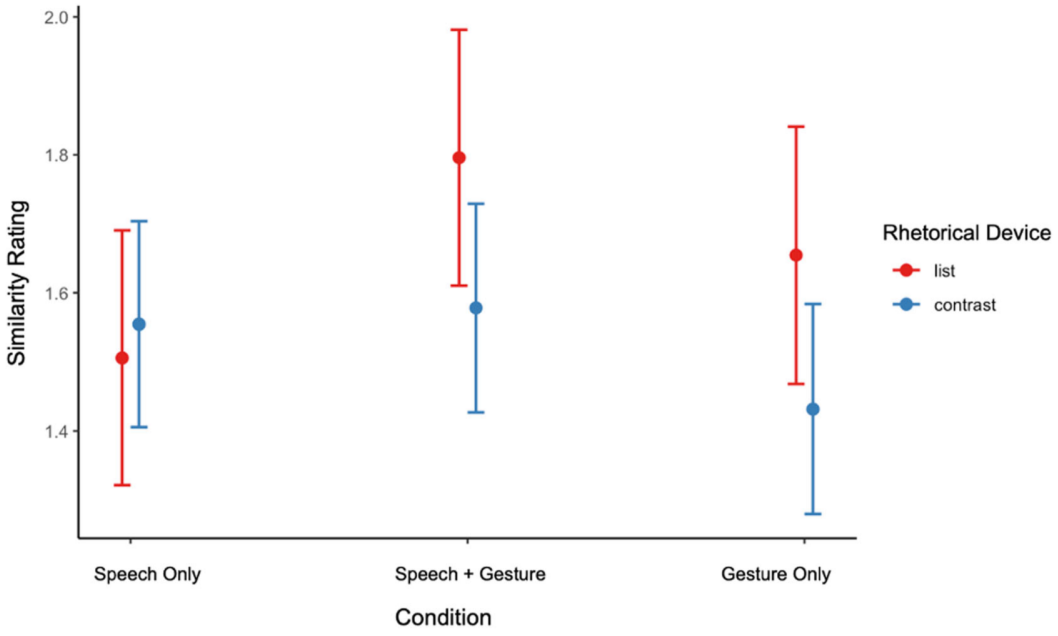

Fig. 3. Predicted values of similarity rating of contrasts and lists by condition.

There was no significant difference between List and Contrasts in terms of their similarity rating ($p = .660$). However, there was a significant interaction between condition and rhetorical device: the effect of Contrast on similarity ratings under the Speech+Gesture condition decreased by 0.27 compared to the effect of Contrast on similarity ratings under the Speech Only condition ($p = .009$); the effect of Contrast on similarity ratings under the Gesture Only condition decreased by 0.27 compared to the effect of Contrast on similarity ratings under the Speech Only condition ($p = .008$). Fig. 3 shows the predicted values of similarity rating for each condition by the type of rhetorical device.

3.2.2. *Simple effects*

Because there was a significant interaction, we conducted an analysis of simple effects with the participant number and question ID as random intercepts within each type of rhetorical device. Within the *list* rhetorical device questions, the effect of condition was significant ($F(2,147) = 5.41$, $p = .005$), such that the rating from the gesture+speech condition on the list questions was significantly different than that from the speech-only condition ($t(148) = 3.29$, $p = .001$). This effect remained significant after adjusting for multiple comparisons using Bonferroni. No other conditions were significantly different from each other. For the *contrast* rhetorical device questions, the effect of condition was not significant ($F(2,147) = 1.61$, $p = .203$).

Similarly, we conducted an analysis of simple effects with participant number and question ID as random intercepts within each condition. Within the speech-only condition, the effect

of rhetorical device was not significant (t (39) = .55, p = .587). Within the gesture-only condition, the effect of rhetorical device was significant (t (40) = 2.26, p = .030), with contrast questions being rated significantly more different (less similar) than list questions. Within the speech+gesture condition, the effect of rhetorical device was also significant (t (40) = 2.14, p = .039), with contrast questions being rated significantly more different than list questions.

3.2.3. Handedness

We found that controlling for the random effect of participants, there was no significant correlation between the match of handedness with the candidate and participants' similarity rating (t (5831) = −1.34, p = .181).

3.3. Study 2 Discussion

In Study 2, we found a significant interaction between condition and rhetorical device such that participants rated lists and contrasts differently in terms of item similarity when they received both gesture and speech, or gesture alone, but not when they listened only to the audio track of speech. Interestingly, although the effect of condition was significant for list items, this effect was not significant for contrast items.

Despite the significance of the results, Study 2 was limited in several ways that suggest important directions for future research. First, although adopting a broad definition of gesture including head and upper body movements enhanced the study's ecological validity, we could not separate the effect of hand gesture from other bodily movements. In the gesture-only and speech + gesture condition, participants saw the politician's facial expressions and other bodily movements and postures in addition to manual gestures. It is possible that the politician's entire bodily movements and facial expressions contributed to the differences in similarity ratings we observed (e.g., Masters, Sullivan, Lanzetta, McHugo, & Englis, 1986 studies of politicians' facial expressions). We tested for the effect of head shake as a marker of negation, and found that head shakes did not differ between conditions. However, further disentangling the nuances between information conveyed by the body, face, and hand would be a worthy topic for future studies. Second, participants in our study were presented with more examples of contrasts than lists. There was no evidence to suggest that viewing more contrasts might influence the perception of similarity, but future work could vary stimuli numbers to determine whether the composition of the set impacts similarity ratings. Third, we focused on only two rhetorical devices; it might be interesting to explore the use of gesture with other rhetorical devices to gain a fuller picture of how gesture might influence audience perception of political arguments. Lastly, as a note to future studies collecting participants' gender information, it is beneficial to adopt an inclusive measure of gender/sex such that participants will have the opportunity to self-describe (Cameron & Stinson, 2019).

Our findings shed light on how communicative messages are perceived with or without the accompaniment of gesture. Past research on rhetorical devices has largely focused on speech. Our study connects the study of rhetorical devices with research on multimodal communication, gesture, and the embodiment of abstract ideas. By demonstrating that co-speech gestures can influence the perception of similarity and difference between ideas spontaneously

compared in speech, this study encourages a holistic and integrated approach to understanding human communication and cognition. It also underscores the importance of considering both the verbal and nonverbal features of interaction to gain a richer understanding of how humans process information and make sense of the world.

4. General discussion

Our study explored the production and perception of co-speech gestures during presidential speech. It found that politicians produce different kinds of gestures when contrasting and comparing concepts and that these gestures are effective at conveying the concepts of difference and similarity to an audience. Importantly, we found that gesture alone but not speech alone was effective at conveying these concepts. Our findings provide important insights into people's perception of gestures with and without speech and into the multimodal mechanisms underlying the persuasive power of political discourse.

In Study 1, we found differences in the gestures politicians produced when listing versus comparing items. When listing items, politicians were less likely to change the gestures they produced from item to item than they were when contrasting items. Changes during lists were most likely to be changes in trajectory pattern. When contrasting items in speech, politicians were likely to change features of their gestures between the two contrasted items, and these changes were most often changes in location.

In Study 2, we asked whether these gestural differences (along with other upper-body movements) contributed to people's perception of the items in speech as being similar or different. We found that the impact of gesture differed when the participant was rating a list question than when the participant was rating a contrast question. Participants in the speech-only condition did not rate concepts in lists as more similar than concepts in contrasts. However, participants in the gesture-only condition and the speech+gesture condition rated concepts in lists as more similar than concepts in contrasts. The fact that the model was improved by adding a random slope of the rhetorical device for each participant suggests that the ways different people perceive list and contrast gestures may vary. Despite this individual variation, our findings suggest that gestures are effective at conveying concepts of same and different—even in the absence of speech—and may carry much of the communicative weight in at least the listing rhetorical device.

It was striking that the effect of rhetorical devices on similarity rating was significant even when participants only had access to gesture (and could not hear the concepts being discussed), but not when people only had access to speech. It seems that when presenting contrasting ideas, merely laying out the two ideas in speech does not cause listeners to perceive a clear difference between the ideas. When gesture accompanies speech, it can be used to visibly mark the ideas as different from one another, by placing them apart in space or by changing the form or movement of one's hands when describing the ideas. Separating ideas in space (i.e., gesture location) was used significantly more during contrasts than lists. This is consistent with research showing that speakers use gesture space to establish locations for ideas that they can then refer back to later in conversation (So, Kita, &

Goldin-Meadow, 2009). Movement, on the other hand, was significantly more likely to change between listed items than contrasted ones. This may be because many list gestures involved largely repetitive circular or peaked movements that stayed in the same location relative to the body but displayed slight path variation during the sequence. List gestures more closely resembled *beat* gestures, while contrast gestures were more likely to resemble *metaphoric* gestures (McNeill, 1992).

Our research aligns with existing literature on the role of gesture in conveying conceptual ideas, as evidenced in studies by Cook and Goldin-Meadow (2006), Cook et al. (2013), and Macoun and Sweller (2016). However, our findings extend this understanding by illustrating how gestures vary across different rhetorical devices, and importantly, how even gestures on their own are perceptible and interpretable by audiences in a manner consistent with the speaker's intended message. This insight contributes to a more nuanced understanding of the communicative power of gesture in persuasive speech.

It is worth pointing out one important aspect of gesture produced by politicians. While these may appear to be spontaneous co-speech gestures, the structured nature of debates means that answers are likely rehearsed. Given that politicians typically undergo extensive media training and preparation for such high-stakes events, their gestures, though often seemingly impromptu, may differ from spontaneous gestures observed in typical everyday conversation. Politicians' gestures might have been strategically crafted for persuasion. A useful question for future research would be to explore the differences between these more "rehearsed" gestures and more spontaneous ones, to examine whether training enhances their persuasiveness. And, of course, rehearsal might not increase persuasiveness: audiences might pick up on the rehearsed nature of the gesture and this might contribute to them feeling less connected or perceiving the message differently.

Beyond the potentially rehearsed nature of the politician's gestures, there are four broad limitations of these two studies that should be acknowledged. First, our dataset of gestures accompanying contrasts versus lists only came from three politicians and six presidential debates. The relatively small sample size limits the variability and range of gestures, topics, and debate contexts that could be examined, which inherently constrains the generalizability of the findings. Second, in a similar vein, our studies only included debate clips of American politicians. Given that there are cross-cultural variations in the production and perception of co-speech gestures (Kita, 2020), caution must be exercised when extrapolating these findings to non-American settings. Future studies might be interested in examining this question using a larger and more diverse sample size. Third, conducting the experiment online in Study 2 introduces variability in the testing environment. Although we put in attention checks, participants might still have differed in their engagement level and the quality of their environment. Contextual factors such as distractions in the participants' surroundings, varying screen sizes, and even differing internet connectivity can impact the way participants perceive and interpret the gestures. Fourth, our participants in Study 2 were all undergraduate students at a highly selective public research institution, who were participating in research for course credit. We did not have a representative sample of U.S. political affiliations represented, and indeed, not all students were familiar with the politicians or U.S. political parties.

It is also important to remember that the multimodal communications politicians make are not confined to laboratory experiments on college campuses. The ways political candidates present their arguments are critical to winning votes and inspiring their supporters. Clearly differentiating one's positions from those of political opponents or listing the key elements in one's platform is crucial, particularly in crowded primary elections. In the era of televised debates and video-based social media, how a politician uses their body during a persuasive speech could help catch the eye and capture the vote of a prospective supporter. This is not meant to imply that there is a simple formula to win votes with the right gestures. However, understanding how politicians' hands can be used to clarify or confuse their intended messages is critical to ensuring that their ideas are heard and understood.

These two studies together underscore the importance of considering multiple modes of communication when we examine how people perceive and process abstract information. The results indicate that the combination of speech and gesture is particularly potent in shaping individuals' perceptions of similarity when dealing with lists of similar items. This finding aligns with the theory of embodied cognition, which implies that bodily actions and representations can impact perception. While prior research has primarily concentrated on tangible entities, the current study extends this understanding by revealing that the way people perceive abstract concepts can be significantly shaped by the speaker's gestures. This revelation not only prompts further exploration of the role and mechanisms of bodily representations in the use of political speech, but also stimulates future research to investigate the application of bodily actions and representations in communication, especially within contexts where there is a deliberate intent to sway others' opinions or preferences through persuasive speech.

Author contributions

EAC designed Study 1 and Study 2, and supervised Study 1. TI prepared the stimuli and collected the data for Study 2, and conducted reliability coding for Study 1. IZ supervised and helped design Study 2, conducted all statistical analyses, and wrote the first draft of the manuscript. All authors edited the manuscript.

Conflict of interest

We have no conflict of interest to disclose.

References

Atkinson, J. M. (1983). Two devices for generating audience approval: A comparative study of public discourse and texts.

Atkinson, M. (1984). *Our masters' voices: The language and body language of politics*. London: Routledge.

Aussems, S., & Kita, S. (2019). Seeing iconic gestures while encoding events facilitates children's memory of these events. *Child Development*, *90*(4), 1123–1137.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.
Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01
Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). FMRI responses to video and point-light displays of moving humans and manipulable objects. *Journal of Cognitive Neuroscience*, *15*(7), 991–1001.
Blumenau, J., & Lauderdale, B. E. (2024). The variable persuasiveness of political rhetoric. *American Journal of Political Science*, *68*(1), 255–270.
Bressem, J., Ladewig, S. H., & Müller, C. (2018). Ways of expressing action in multimodal narrations–The semiotic complexity of character viewpoint depictions. In A. Hübl & M. Steinbach (Eds.), *Linguistic foundations of narration in spoken and sign languages* (pp. 223–249). Amsterdam: Benjamins. https://doi.org/10.1075/la.247.10bre
Bull, P. (1986). The use of hand gesture in political speeches: Some case studies. *Journal of Language and Social Psychology*, *5*(2), 103–118.
Burgoon, J. K., Birk, T., & Pfau, M. (1990). Nonverbal behaviors, persuasion, and credibility. *Human Communication Research*, *17*(1), 140–169.
Cameron, J. J., & Stinson, D. A. (2019). Gender (mis) measurement: Guidelines for respecting gender diversity in psychological research. *Social and Personality Psychology Compass*, *13*(11), e12506.
Casasanto, D., & Jasmin, K. (2010). Good and bad in the hands of politicians: Spontaneous gestures during positive and negative speech. *Plos one*, *5*(7), e11805.
Charteris-Black, J. (2011). *Politicians and rhetoric: The persuasive power of metaphor*. Springer.
Cienki, A. (2004). Bush's and Gore's language and gestures in the 2000 US presidential debates: A test case for two models of metaphors. *Journal of Language and Politics*, *3*(3), 409–440.
Cocks, N., Sautin, L., Kita, S., Morgan, G., & Zlotowitz, S. (2009). Gesture and speech integration: An exploratory study of a man with aphasia. *International Journal of Language & Communication Disorders*, *44*(5), 795–804.
Cook, S. W., Duffy, R. G., & Fenn, K. M. (2013). Consolidation and transfer of learning after observing hand gesture. *Child Development*, *84*(6), 1863–1871.
Cook, S. W., & Goldin-Meadow, S. (2006). The role of gesture in learning: Do children use their hands to change their minds? *Journal of Cognition and Development*, *7*(2), 211–232.
Dargue, N., & Sweller, N. (2020). Learning stories through gesture: Gesture's effects on child and adult narrative comprehension. *Educational Psychology Review*, *32*(1), 249–276.
Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765.
Dick, A. S., Goldin-Meadow, S., Hasson, U., Skipper, J. I., & Small, S. L. (2009). Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping*, *30*(11), 3509–3526.
Eielts, C., Pouw, W., Ouwehand, K., Van Gog, T., Zwaan, R. A., & Paas, F. (2020). Co-thought gesturing supports more complex problem solving in subjects with lower visual working-memory capacity. *Psychological Research*, *84*(2), 502–513.
Foglia, L., & Wilson, R. A. (2013). Embodied cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 319–325.
Goldin-Meadow, S. (2003). Thought before language: Do we think ergative. In D. Gentner & S. Goldin-Meadow, (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 493–522).
Graham, J. M. (2000). Interaction effects: Their nature and some post hoc exploration strategies.
Grahe, J. E., & Bernieri, F. J. (1999). The importance of nonverbal cues in judging rapport. *Journal of Nonverbal Behavior*, *23*(4), 253–269.
Guilbeault, D. (2017). How politicians express different viewpoints in gesture and speech simultaneously. *Cognitive Linguistics*, *28*(3), 417–447.
Hall, K., Goldstein, D. M., & Ingram, M. B. (2016). The hands of Donald Trump: Entertainment, gesture, spectacle. *HAU: Journal of Ethnographic Theory*, *6*(2), 71–100.

Hart, C., & Winter, B. (2022). Gesture and legitimation in the anti-immigration discourse of Nigel Farage. *Discourse & Society*, *33*(1), 34–55.

Heritage, J., & Greatbatch, D. (1986). Generating applause: A study of rhetoric and response at party political conferences. *American Journal of Sociology*, *92*(1), 110–157.

Hinnell, J. (2019). The verbal-kinesic enactment of contrast in North American English. *American Journal of Semiotics*, *35*(1–2), 55–92.

Hinnell, J., & Parrill, F. (2020). Gesture influences resolution of ambiguous statements of neutral and moral preferences. *Frontiers in Psychology*, *11*, 587129.

Hostetter, A. B. (2011). When do gestures communicate? A meta-analysis. *Psychological Bulletin*, *137*(2), 297.

Hubbard, A. L., McNealy, K., Scott-Van Zeeland, A. A., Callan, D. E., Bookheimer, S. Y., & Dapretto, M. (2012). Altered integration of speech and gesture in children with autism spectrum disorders. *Brain and Behavior*, *2*(5), 606–619.

Hubbard, A. L., Wilson, S. M., Callan, D. E., & Dapretto, M. (2009). Giving speech a hand: Gesture modulates activity in auditory cortex during speech perception. *Human Brain Mapping*, *30*(3), 1028–1037.

Jackob, N., Roessing, T., & Petersen, T. (2011). The effects of verbal and nonverbal elements in persuasive communication: Findings from two multi-method experiments.

Kartalkanat, H., & Göksun, T. (2020). The effects of observing different gestures during storytelling on the recall of path and event information in 5-year-olds and adults. *Journal of Experimental Child Psychology*, *189*, 104725.

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.

Kelly, S. D., Özyürek, A., & Maris, E. (2010). Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension. *Psychological Science*, *21*(2), 260–267.

Kita, S. (2020). Cross-cultural variation of speech-accompanying gesture: A review. In S. Kita (Ed.), *Speech accompanying-gesture* (pp. 145–167).

Krauss, R. M., Chen, Y., & Chawla, P. (1996). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? *Advances in Experimental Social Psychology*, *28*, 389–450.

Ladewig, S. H., & Bressem, J. (2013). New insights into the medium hand: Discovering recurrent structures in gestures. *Semiotica*, *2013*(197), 203–231.

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

Lempert, M. (2011). Barack Obama, being sharp: Indexical order in the pragmatics of precision-grip gesture. *Gesture*, *11*(3), 241–270.

Lempert, M. (2017). Uncommon resemblance: Pragmatic affinity in political gesture. *Gesture*, *16*(1), 35–67.

Lin, Y. L. (2017). Co-occurrence of speech and gestures: A multimodal corpus linguistic approach to intercultural interaction. *Journal of Pragmatics*, *117*, 155–167.

Lin, Y. L. (2021). Gestures as scaffolding for L2 narrative recall: The role of gesture type, task complexity, and working memory. *Language Teaching Research*, 13621688211044584.

Maricchiolo, F., Gnisci, A., Bonaiuto, M., & Ficca, G. (2009). Effects of different types of hand gestures in persuasive speech on receivers' evaluations. *Language and Cognitive Processes*, *24*(2), 239–266.

Masters, R. D., Sullivan, D. G., Lanzetta, J. T., McHugo, G. J., & Englis, B. G. (1986). The facial displays of leaders: Toward an ethology of human politics. *Journal of Social and Biological Structures*, *9*(4), 319–343.

Macoun, A., & Sweller, N. (2016). Listening and watching: The effects of observing gesture on preschoolers' narrative comprehension. *Cognitive Development*, *40*, 68–81.

McNeil, N. M., Alibali, M. W., & Evans, J. L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior*, *24*, 131–150.

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

McNeill, D. (2000). *Language and gesture*. Cambridge: Cambridge University Press.

McNeill, D. (2006). *Gesture and communication. Concise Encyclopedia of Pragmatics*. New York: Elsevier.

McNeill, D. (2005). *Gesture and thought*. University of Chicago Press.

McNeill, D., Cassell, J., & McCullough, K. E. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction*, *27*(3), 223–237.

Mehrabian, A., & Williams, M. (1969). Nonverbal concomitants of perceived and intended persuasiveness. *Journal of Personality and Social Psychology*, *13*(1), 37.

Mendoza-Denton, N., & Jannedy, S. (2011). Semiotic layering through gesture and intonation: A case study of complementary and supplementary multimodality in political speech. *Journal of English Linguistics*, *39*(3), 265–299.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, *19*, 309–326.

Morford, M., & Goldin-Meadow, S. (1992). Comprehension and production of gesture in combination with speech in one-word speakers. *Journal of Child Language*, *19*(3), 559–580.

Neff, M., Wang, Y., Abbott, R., & Walker, M. (2010). Evaluating the effect of gesture and language on personality perception in conversational agents. In Intelligent Virtual Agents: 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings 10 (pp. 222-235). Springer Berlin Heidelberg.

Özyürek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related brain potentials. *Journal of Cognitive Neuroscience*, *19*(4), 605–616.

Parrill, F., & Stec, K. (2017). Gestures of the abstract: Do speakers use space consistently and contrastively when gesturing about abstract concepts? *Pragmatics & Cognition*, *24*(1), 33–61.

Parrill, F., Hinnell, J., Moran, G., Boylan, H., Gupta, I., & Zamir, A. (2022). Observers use gesture to disambiguate contrastive expressions of preference. *Language and Cognition*, *14*(1), 19–46.

Perry, M., Berch, D., & Singleton, J. (1995). Constructing shared understanding: The role of nonverbal input in learning contexts. *Journal of Contemporary Legal Issues*, *6*, 213.

Peters, J., & Hoetjes, M. W. (2017). The effect of gesture on persuasive speech. In *Proceedings of Interspeech* (pp. 659–663).

Pi, Z., Hong, J., & Yang, J. (2017). Effects of the instructor's pointing gestures on learning performance in video lectures. *British Journal of Educational Technology*, *48*(4), 1020–1029.

Qualtrics. (2022). Provo, Utah, USA. Available at: https://www.qualtrics.com

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.

Rohrer, P. L., Delais-Roussarie, E., & Prieto, P. (2020). Beat gestures for comprehension and recall: Differential effects of language learners and native listeners. *Frontiers in Psychology*, *11*, 575929.

Sekine, K., & Kita, S. (2017). The listener automatically uses spatial story representations from the speaker's cohesive gestures when processing subsequent sentences without gestures. *Acta Psychologica*, *179*, 89–95.

So, W. C., Kita, S., & Goldin-Meadow, S. (2009). Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science*, *33*(1), 115–125.

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665–681.

Stokoe, W. C. (1980). Sign language structure. *Annual Review of Anthropology*, *9*(1), 365–390.

Streeck, J. (2008). Gesture in political communication: A case study of the democratic presidential candidates during the 2004 primary campaign. *Research on Language and Social Interaction*, *41*(2), 154–186.

Wakefield, E., Novack, M. A., Congdon, E. L., Franconeri, S., & Goldin-Meadow, S. (2018). Gesture helps learners learn, but not merely by guiding their visual attention. *Developmental Science*, *21*(6), e12664.

Wehling, E. (2009). Argument is gesture war: Function, form, and prosody of discourse structuring gestures in political argument. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 35, No. 2, pp. 54–65).

Wehling, E. (2017). Discourse management gestures. *Gesture*, *16*(2), 245–276.

Yin, R., Su, D., & Ye, H. (2013). Conceptual metaphor theory: Basing on theories of embodied cognition. *Advances in Psychological Science*, *21*(2), 220.

Zhang, I. Y., Fu, S., Xu, A. Z., & Lu, H. (2022). Causal versus associative relations: Do humans perceive and represent them differently? In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44).

Zhang, J. R., Sherwin, J., Dmochowski, J., Sajda, P., & Kender, J. R. (2014). Correlating speaker gestures in political debates with audience engagement measured via EEG. In *Proceedings of the 22nd ACM International Conference on Multimedia* (pp. 387–396).

Appendix A

Compared to the Speech Only group's rating of List items, similarity ratings of List items increased by 0.29 for the Speech+Gesture group ($p < .001$) and by 0.15 for the Gesture Only group ($p = .088$). There was no significant difference between List and Contrasts in terms of their similarity rating ($p = .613$). However, there was a significant interaction between condition and rhetorical device: the effect of Contrast on similarity ratings under the Speech+Gesture condition decreased by 0.27 compared to the effect of Contrast on similarity ratings under the Speech Only condition ($p < .001$); the effect of Contrast on similarity ratings under the Gesture Only condition decreased by 0.27 compared to the effect of Contrast on similarity ratings under the Speech Only condition ($p < .001$).

Fixed effects of modality and rhetorical device

	Estimate	SE	df	*t*-value	Pr(>|t|)
(Intercept)	1.51	0.09	94.19	16.17	< .001***
ConditionSpeech + Gesture	0.29	0.09	278.37	3.37	< .001***
ConditionGesture Only	0.15	0.09	278.32	1.72	0.086
Rhetorical Device (List as Reference)	0.05	0.10	54.89	0.52	0.607
Speech + Gesture:Contrast	−0.27	0.06	5713.66	−4.18	< .001***
Gesture Only:Contrast	−0.27	0.06	5713.63	−4.23	< .001***

Type III analysis of variance using Satterthwaite's method for model 1

	Sum of squares	Mean square	NumDF	DenDF	*F*-value	Pr(>F)
Modality	4.5546	2.2773	2	156.8	2.6841	0.07142
Rhetorical device	1.8515	1.8515	1	39.9	2.1822	0.14746
Modality: Rhetorical device	20.0668	10.0334	2	5712.7	11.8255	7.496e-06***