Daniel Alcaraz Carrión* and Javier Valenzuela

Distant time, distant gesture: speech and gesture correlate to express temporal distance

https://doi.org/10.1515/sem-2019-0120
Received November 22, 2019; accepted August 17, 2020; published online June 14, 2021

Abstract: This study investigates whether there is a relation between the semantics of linguistic expressions that indicate temporal distance and the spatial properties of their co-speech gestures. To this date, research on time gestures has focused on features such as gesture axis, direction, and shape. Here we focus on a gesture property that has been overlooked so far: the distance of the gesture in relation to the body. To achieve this, we investigate two types of temporal linguistic expressions are addressed: proximal (e.g., near future, near past) and distal (e.g., distant past, distant future). Data was obtained through the NewsScape library, a multimodal corpus of television news. A total of 121 co-speech gestures were collected and divided into the two categories. The gestures were later annotated in terms of gesture space and classified in three categories: (i) center, (ii) periphery, and (iii) extreme periphery. Our results suggest that gesture and language are coherent in the expression of temporal distance: when speakers locate an event far from them, they tend to gesture further from their body; similarly, when locating an event close to them, they gesture closer to their body. These results thus reveal how co-speech gestures also reflect a space-time mapping in the dimension of distance.

Keywords: gesture; metaphor; multimodality; time

1 Introduction

Time is a universal phenomenon that is experienced by all humans; in the different conceptualization patterns that have been explored, the connection with the spatial domain has received the most detailed focus (e.g., Boroditsky 2000, 2001; Boroditsky and Ramscar 2002; Casasanto and Boroditsky 2008; Kranjec and

***Corresponding author: Daniel Alcaraz Carrión**, University of Wisconsin-Madison, Wisconsin, USA, E-mail: alcarazcarri@wisc.edu. https://orcid.org/0000-0001-5325-431X
Javier Valenzuela, Universidad de Murcia, Murcia, Spain, E-mail: jvalen@um.es

Chatterjee 2010; Kranjec and McDonough 2011; Matlock et al. 2005; Núñez and Sweetser 2006; Santiago et al. 2007; Torralbo et al. 2006).

In English, for example, multiple temporal meanings can be expressed by means of a spatial metaphor (e.g., *winter is coming*; *look back in the past*) which typically makes use of the sagittal axis: future events are placed in front of the speaker and past events behind the speaker (Moore 2006, 2011). The spatialization of time is also reflected in other modalities of communication, like cultural artefacts (e.g., calendars, sundials) or co-speech gesture (Davis 2012; Duffy 2014; Fauconnier and Turner 2008).

Over the last few years, co-speech gestures have become an invaluable source of information for research in language, cognition, and communication. Gesture is often seen as a "window into the mind" (McNeill 1985, 1992, 2005); gestures are a crucial component in many aspects of human communication, helping tasks such as lexical disambiguation or syntactic parsing, and sometimes even supplying additional information not explicitly mentioned in speech (Beattie 2016; Beattie and Shovelton 2011; Kendon 1980, 2000, 2004; Özyrükek 2014). Indeed, most researchers in gesture studies advocate for an integration of speech and gesture in cognition (Beattie 2016; Goldin-Meadow 2003, 2004; Iverson and Goldin-Meadow 1998; McNeill 2005). Cienki (1998, 2008, 2013) presents a wide range of research that highlights the importance of studying both oral and gesture communication when addressing metaphorical thinking and points at the fluid speech-gesture interplay found in metaphorical representation.

When looking at temporal metaphors, and, more precisely, at the TIME IS SPACE metaphor (Lakoff 1987; Lakoff and Johson 1980), the gesture dimension is often used as a tool to spatialize the temporal domain; an ever-growing number of researchers have studied the solid relationship existing between co-speech gesture and time-space metaphors (e.g., Bender and Beller 2014; Bylund et al. 2019; Calbris 2008; Cienki and Müller 2008; Lapaire 2016; Levinson and Majid 2013; Ng et al. 2017; Núñez and Sweetser 2006; Walker and Cooperrider 2016; Valenzuela et al. 2020). Cienki (1998), for example, suggests that, in the case of English speakers, there is a tendency to project events onto imaginary timelines that go from left to right. This was confirmed by Casasanto and Jasmin (2012), who pointed out that English speakers tend to use the lateral axis when spontaneously gesturing about time (see also Pagán Cánovas et al. 2020). On the other hand, Walker and Cooperrider (2016) found that speakers systematically combine left-to-right and back-front spatial axes when gesturing about time. Cooperrider and Núñez (2009) suggest that time spatialization by co-speech gesture might be influenced by culturally shared temporal conceptualization processes (e.g., writing direction, sagittal temporal language, calendars), which could explain why gesture seems to combine both the lateral and the sagittal axis.

There is also a wide range of literature that investigates time, space, and co-speech gesture from a cross-cultural and cross-linguistic perspective. Núñez and Sweetser (2006), for example, indicate that Aymara speakers tend to locate the past in front of them and the future behind them through gesture and linguistic temporal realizations, while Boroditsky (2001) suggests that Mandarin speakers employ linguistic vertical metaphors that are mirrored by co-speech gestures. Recent research by de la Fuente et al. (2014) suggests a possible link between temporal cognition, gesture, and progressive/conservative cultural values, comparing speakers of Spanish and Arabic.

All the aforementioned research has focused on two key characteristics of co-speech gestures: axis and direction. These two features have been identified as maximally informative when establishing cognitive patterns relating language and gesture, since directionality and axial localization are also represented in language (e.g., the expression *back in those days* indicates both a spatial axis, the sagittal, and a directionality, behind the deictic center). However, there is an aspect that has been overlooked so far: there are some cases in which spatial temporal metaphors can also indicate how far (metaphorically) the temporal event is located from the speaker, as in *near past* or *distant future*. These expressions establish an underspecified, often subjective distance between a deictic center and a moment in time, as "seen" from the speaker's viewpoint.

So far, research has not explored how the dimension of distance is represented in temporal gestures. A previous study has explored the relation between spatial distance and temporal duration (Cai et al. 2013), suggesting that long-distance gestures could lengthen the subjective duration of time. This relation was established without using linguistic inputs, and refers to a slightly different aspect of time that the one that we will address in this manuscript (temporal *duration* refers to how long an event lasts, while temporal *distance* refers to when an event happened in relation to the speaker). Previous studies have suggested that the location the gesture influences the visual attention of the listeners as well as the meaning extracted from the gesture, with gestures performed in the periphery being more fixated than gestures performed in the center (Beattie, 2016; Gullberg and Holmqvist 1999, 2002, 2006; Gullberg and Kita 2009). In the case of temporal expressions that refer to temporal distance, we hypothesize that speakers will be more likely to gesture close to them when using a proximal linguistic expression and far from them when using a distal temporal expression. The results would add much weight to the idea that the domain of metaphorical distance is represented not only through language, but also through co-speech gestures. If the temporal meaning being represented affects this aspect of the shape of gestural patterns across a significant number of instances, this will provide relevant evidence for the cross-modal expression of yet another aspect of spatialized time, besides

directionality. Such results would support the idea that time is spatialized through a rich mental scene or simulation, where different spatial properties can be foregrounded opportunistically for meaning construction.

Thus, this paper explores the relation between temporal linguistic expressions that use spatial metaphors to indicate temporal proximity or temporal distance (e.g., *near past, distant past*) and the location of the gestures that co-occur with these linguistic expressions. Qualitative and quantitative methods are employed in order to explore the relation between gesture location and linguistic expressions of temporal distance.

2 Methodology

2.1 Dataset

The data was extracted from the NewsScape Library of Television News (http://newsscape.library.ucla.edu/), a repository of television news with both textual and audio-visual data from more than 250,000 television programs, through the collaboration with the Distributed Little Red Hen Lab, co-directed by Francis Steen and Mark Turner (http://www.redhenlab.org/home). The database stores around 350,000 h of television, and a textual database of more than two billion words. The NewsScape Archive includes broadcasts in a large number of languages, including English, Spanish, German, Italian, and Arabic, among other languages, in the form of subtitled television news programs that have been recorded since 2004 until the present day. Data is collected from a wide range of channels such as BBC, CNBN, or Fox News. The user accesses all the information provided by the database through the Edge search engine, which has customized tools for qualitative and quantitative analysis. The NewsScape Library allows researchers to look for a particular linguistic expression and obtain textual information from the subtitles, as well as a video clip of each moment in which the expression was uttered across the TV collection (Figure 1).

2.2 Procedure

The first step in the data collection process was to elaborate a list of different spatial temporal linguistic expressions that explicitly state temporal distance in relation to the speaker. The type of language targeted employs the spatial domain to talk about time, but with no reference to directionality, only mentioning the proximity or distance of the event (e.g., *distant/near future*). It could be argued that

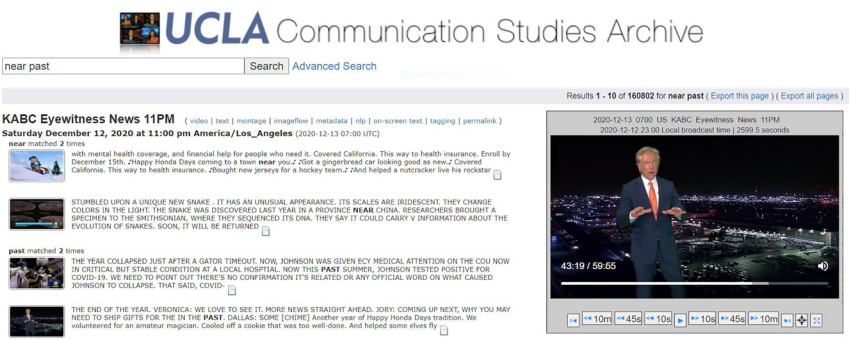

Figure 1: NewsScape Edge search engine.

the temporal distance of an event can also be gleaned from other types of spatial temporal expressions. For example, expressions such as *back in those days* could be said to have an implicit temporal distance, but such information is not included in the temporal expression itself. This information needs to be extracted from the accompanying discourse or from contextual factors. Furthermore, the notion of *temporal* distance is a property that depends on the perception of the speaker, and is thus highly subjective. One person can, for example, talk about a memory from their childhood as if it were something that happened yesterday, or as something that happened long time ago. The use of spatial, non-directional language offers a more objective way of knowing whether the speaker is conceptualizing the event as something temporally close or far, since this information is explicitly encoded in the language used. This allows us to link the temporal co-speech gesture to the speaker's perception of the temporal event, rather than to the actual moment when the event took place, which is often signaled by additional linguistic cues.

When elaborating a list of linguistic items, searches were restricted to simple grammatical patterns involving a noun phrase with the nouns *past* and *future* modified by a distance adjective (*distant, near, far, long, remote*), or and adverbial phrase with a distance adjective followed by the prepositional phrase *in the past* or *in the future*. The resulting list was thus formed by linguistic expressions that employed the spatial domain to locate temporal events both in the past and in the future at a certain distance from the speaker (see Table 1).

It is important to highlight that English presents a reduced number of expressions to indicate temporal proximity. *Near past* and *near future* were the only expressions that exclusively referred to temporal proximity that could be found in the multimodal database. The rest of the set of expressions that express

Table 1: Temporal linguistic expressions searched in NewsScape.

Temporal period	Spatial location	Expression	Search time span on NewsScape
Proximal	Past	Not so/too distant past	01/01/2004–28/02/2016
		Not (that/too) far in the past	01/01/2004–28/02/2016
		Not (that/too) long in the past	01/01/2004–28/02/2016
		Near past	01/01/2004–28/02/2016
	Future	Not so/too distant future	04/06/2014–29/02/2016
		Not (that/too) far in the future	01/01/2004–28/02/2016
		Not (that/too) long in the future	01/01/2004–28/02/2016
		Near future	30/09/2015–29/02/2016
Distal	Past	Distant past	01/01/2004–28/02/2016
		Far in the past	01/01/2004–28/02/2016
		Long in the past	01/01/2004–28/02/2016
		Remote past	01/01/2004–28/02/2016
	Future	Distant future	04/06/2014–29/02/2016
		Far in the future	01/01/2004–28/02/2016
		Long in the future	01/01/2004–28/02/2016
		Not in the near future	30/09/2015–29/02/2016

temporal proximity had to be formed by the negation of temporal distance expressions (e.g., *not so distant future*). Despite being infrequent, relatively numerous tokens of these expressions could be found, thanks to the large size of the corpus. This shows the need for sizeable audiovisual datasets to attest the full scope of language use. The opposite pattern, negated temporal proximity expressions to indicate temporal distance, was less common, although some marginal examples were also found, which was again made possible by the size of the dataset. In general terms, the amount of temporal expressions that make exclusive reference to the proximity of a temporal event is quite reduced in English, with *near* + TEMPORAL NOUN PHRASE being the most common construction.

Once the linguistic items were chosen, searches were performed using the tools available in the NewsScape library. Each of the hits obtained from this search was individually categorized using the following coding protocol:

– Firstly, features related to the organization and identification of the dataset were annotated. Each entry was assigned a number, a hyperlink to the NewsScape library in order to facilitate its retrieval for viewing, the date and the name of the program, and also some textual context (10 words before and after the search word(s)).

– Secondly, each clip was individually analyzed and classified to distinguish between noise data (categories 1, 2, and 3, explained below), data in which the speaker was not fully visible (4 and 5) and clips in which the speaker was clearly visible (6, 7, and 8). These are the categories in which the clips were divided:

(1) *Repeated clip*: cases in which the same clip appeared more than once; this happened quite frequently with expressions that were used in advertisements or videos that were broadcast several times on television.
(2) *False hit*: cases in which the close caption was different from the real speech, that is, the speech did not correspond to the textual data. Additionally, this category included cases in which the search item was part of two different sentences, i.e., *Jenner has not been charged with any crime so* ***far. In the past****, she has pledged …* (KABC Eyewitness News, Thursday July 9, 2015).
(3) *Broken link*: cases in which technical problems in the recording of the program in the database did not allow for its observation.
(4) *Voice-over*: cases in which the speaker was not visible; a voice can be heard uttering the relevant expression but the image shows something else, i.e., a video, a landscape or an image, the audience …
(5) *Non-visible hands*: cases in which the speaker is partially visible but the camera angle did not allow for a clear observation of their hand(s), e.g., a close-up of the speaker's face.
(6) *No gesture*: this category included all the cases in which the hand(s) of the speakers were clearly visible but no co-speech gesture was performed.
(7) *Unrelated*: There was a number of cases in which beat gestures (repeated realizations of vertical strokes; McNeill, 1992) were performed while the temporal linguistic expression was being uttered; the speaker performed this gesture frequently before, during, and after the linguistic expression.
(8) *Relevant co-speech gesture*: this last category included all the temporal gestures that were relevant to the analysis.

Finally, we also included a brief description or the name of the speaker that was performing the gesture (e.g., KCAL weatherman; Jimmy Fallon) in order to easily recognizes who was the gesturer. This was important, since the data from the NewsScape library is obtained from television news programs, there are certain gesturers such as journalists and television personalities that have a higher amount of screen time. This could affect the results by including a higher amount of gesture realizations of some of these more frequent speakers. In order to avoid our dataset to be biased in this way and thus keep the independence of our measurements, we only included in our final dataset gesture realizations that were not repeated by the same speaker. For example, if there were three instances in

which the same speaker performed a gesture in the extreme periphery, only one of those instances would count as valid while the other two would be discarded.

All the gestures that fell under the category of *relevant co-speech gesture* subsequently underwent a fine-grained analysis based on Bressem et al. proposal for co-verbal gesture annotation (2013), although some modifications were made to this annotation protocol. Our analysis focused on the gestural space used by the speaker when producing a co-speech gesture. We started from the usual partitioning found in the literature, based in McNeill's gesture space diagram (McNeill 1992). Most of the categorizations used in the literature are based on his four-way distinction (extreme-periphery, periphery, center, and center-center), which has been directly used in many studies (e.g., Cohen et al. 2010), or adapted depending on the purpose of the study (e.g., Gullberg and Holmqvist 1999 and Gullberg and Kita 2009 only distinguish between central and peripheral space). In our case, three categories were established for lateral gestures: gestures were performed in the *center* when the hands were found between both shoulders; gestures were performed in the *periphery* when the end of the stroke was found in the space located slightly further from the shoulders; and finally, gestures were performed the *extreme periphery* when the end of the stroke was located further beyond the *periphery*. The application of this classification for gestures in the lateral axis is rather straightforward. However, the analysis of the gesture space when the gesture is performed in other axes is more complicated. Leaving aside the vertical axis, which is not used much in English (and actually, no examples were found in our analysis), there is no mention of how to address the gestures performed in the sagittal axis in previous literature. In this case, the body cannot be taken as a clear reference to establish their location. What was done, thus, was to take into account physiological features when analyzing the gesture space (Cienki et al. 2016), namely, looking at the movement of hand or the wrist as well as the level of extension of the arm.

More concretely, when a sagittal gesture was performed forwards, three different locations were distinguished. Firstly, the gesture would be located in the center if the gesture was performed just by using the hands and the wrist, with little or no arm motion, in the gesture space of the chest (Figure 2).

A sagittal forward gesture could also be located in the periphery. In this case, the level of extension and movement of the arm were taken into account. So, a gesture would be located in the periphery if the speaker extended the forearm partially when gesticulating (Figure 3).

Finally, a sagittal forward gesture could also be located in the extreme periphery if the speaker fully extended the arm. This could sometimes be accompanied by the movement of the shoulder or the body (Figure 4).

Figure 2: Sagittal forward gesture located in the center.

Figure 3: Sagittal forward gesture located in the periphery.

Figure 4: Sagittal forward gesture located in the extreme periphery.

Sagittal gestures with a backwards motion turned up to be a problematic case. It is difficult to establish how far from the body is a gesture that is performed backwards, since there are anatomical limitations in the backward motion we can perform. Hence, only two different gesture locations in backwards gestures were established: in the center or in the extreme periphery.

First, a gesture would be located in the *center* if the backward movement was again performed with a small movement of the hands or the wrists in the gesture space surrounding the chest, with little or no arm motion (Figure 5).

Figure 5: Sagittal backward gesture located in the center.

The rest of the sagittal backward gestures that were not located in the center presented a motion in which the speaker generally rose the arm backwards, with the hand at the same level as the shoulder and performed some type of movement with the hand to signal behind them (Figure 6).

Figure 6: Backwards sagittal located in the extreme periphery.

The consistency of the coding scheme was verified by checking the level of agreement between two coders when analyzing the gesture data in terms of gesture location. Both coders presented high levels of agreement when locating the gestures in the gestural space (N = 25, 88% congruence, Cohen's κ = 0.76). Both coders agreed in those cases in which the gesture was located either in the *center* or in the *extreme periphery*, while there were some discrepancies when coding a gesture as being located in the *periphery*.

2.3 Results

The initial searches of the linguistic items in the NewsScape Library provided a total of 1,138 hits (see Table 2), comprised by 517 hits of distance-related language and 621 hits of proximity-related language. Almost 20% of the dataset consisted of noise data, so the total amount of valid hits was reduced to 902. This amount was further reduced when eliminating the clips that presented a voice over (38%) and the cases in which the hands of the speakers were not visible due to the camera shot or to on-screen images (36%).

The remaining data was formed by 262 clips (23% of the original dataset) in which the hands of the speakers were clearly visible; 136 in proximity-related language and 126 in distance-related language. From all these cases, there were 76 instances in which the speakers did not produce any type of hand gesture. An additional 65 hits were categorized as unrelated, non-temporal gestures.

This left the dataset with a total of 121 temporal semantic co-speech gestures. These instances were lastly reviewed in terms of the gesturer to ensure the independence of the data, with only three cases in which the same speaker performed a similar gesture in a different occasion. The first one was a KCAL weatherman performing a sagittal gesture in the extreme periphery, which was repeated in two other clips that were removed from the dataset. The last case that was eliminated shows the lead presenter of the *KABC Today in LA Weekend* performing a lateral gesture located in the periphery in two different video clips, and accordingly one of them was eliminated from the dataset. All the cases were instances that employed the distal temporal structure *distant future*. By removing these three instances in which the same speakers performed the same type of gesticulations, we ensure that we have a total of 118 unique gesture realizations performed by 118 different speakers. These gesture realizations are distributed among both distal and proximal language as follows.

When examining proximal language, 42 out of the 136 clearly visible hits were analyzed as semantic gestures (30%), 44 cases were included in the category of non-semantic gestures (32%) and the remaining 50 gestures (38%) were cases in

which speakers did not perform a co-speech gesture. Some of the temporal linguistic structures show some interesting tendencies: there were a total of 59 cases in which the hands were visible when the speakers used the expression *near future,* which is the most frequent expression when referring to temporal proximity. However, it seems that speakers do not often perform a semantic gesture with this expression: only 13 of the cases (22%) were semantic gestures; the rest were either non-semantic (44%) or there was no gesture (44%). In fact, most of the expressions presented a low ratio of semantic gestures, with *not so/too distant past* with a 37% of semantic gestures and *not so distant future* (25% of semantic gestures). The rest of the proximity expressions (i.e., *not that/too far in the past, not that/too long in the past, not too far in the future*) presented a very low amount of hits in which the hands were visible; the dataset of these expressions was not large enough to make strong generalizations. These 42 gestures were performed in different communicative situations in television, including face-to-face interviews between two people (31%), presenters talking to the camera (38%), reporters talking to the camera (21.4%), political speeches (7.1%) and weather forecasts (2.3%).

Concerning distal language, the number of semantic co-speech gestures is higher than in proximal language. Out of the total 126 cases in which the hands were visible, three semantic gestures were discarded since they were produced by the same speaker, as it was previously stated. From the remaining 123 video clips, 76 of them were semantic gestures (61.7%), 21 of them were non-semantic (17%) and 26 cases did not include a co-speech gesture (21.13%). Examining the gesture frequency of specific linguistic expressions, *distant future* and *far in the future* show a very high amount of semantic gestures (68 and 74%, respectively); this tendency is slightly weaker when speakers employ expressions to refer to the past, such as *distant past* (51% of semantic gestures). There are some expressions that had a very low frequency in the database searches: *long in the past*, *remote past, long in the future* and *nor in the near future* presented a very reduced number of cases in which the hands were visible, so no reliable generalizations about these specific expressions can be made. Nevertheless, it can be stated that the rest of the expressions present an average of semantic gestures well above the 60%. Similarly to proximal gestures, distal gestures were employed in different communicative situations, as follows: 35.5% were face-to-face interviews, 42.2% were presenters talking to the camera, 13.15% were reporters speaking to a camera, 6.6% were weather reports and finally the remaining 2.6% were political speeches. The data shows that speakers are more likely to produce a temporal semantic gesture when they are talking about distant events than when they talk about proximal events. A preliminary chi-square test shows that these differences are statistically significant, $\chi^2(2) = 26.69$, $p < 0.001$ (see Figure 7).

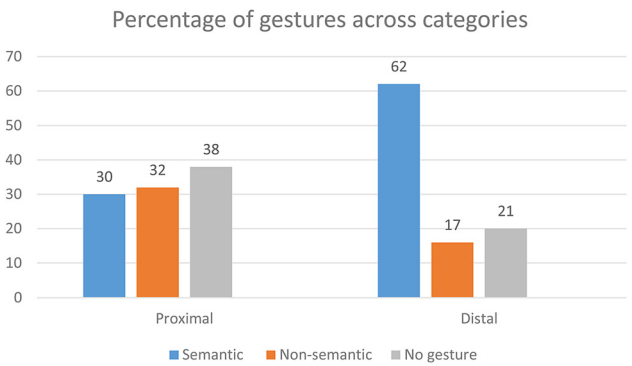

Figure 7: Percentage of gesture frequency across categories.

Distal and proximal co-speech gestures were further analyzed in terms of their location in gesture space. From the 76 semantic gestures that co-occurred with distal language, four were discarded because they signaled two different distances simultaneously (e.g., "in the **near future**, much less a **distant future**, that we can actually turn the tide"; in 2013–08–28, KOCE Tavis Smiley). Of the remaining 72, 14 of them were located in the center (19%); 22 gestures were located in the periphery (32%); and the remaining 36 gestures were located in the extreme periphery (49%). Concerning proximal language, out of the 42 co-speech gestures, 20 gestures were located in the center (48%); 16 gestures were located in the periphery (38%); and six gestures were located in the extreme periphery (14%; Figure 8).

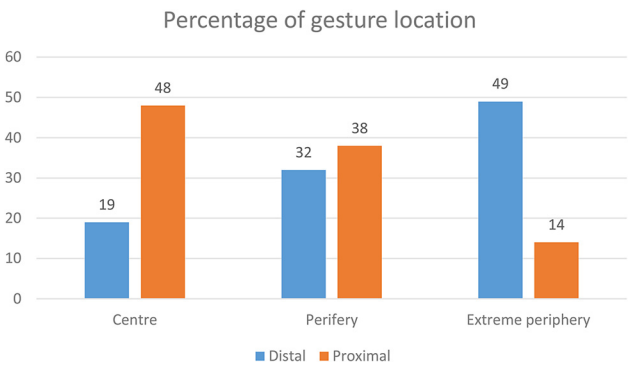

Figure 8: Percentage of gesture location in distal and proximal language.

A binomial logistic regression analysis was carried with the Type of Linguistic Expression (distal or proximal) as a dependent variable and Gesture Distance (center, periphery and extreme periphery) as a predictor. The overall model was shown to be significant (Cox and Snell's $R^2 = 0.146$; Nagelkerke's $R^2 = 0.201$), with an adequate fit, Wald $\chi^2(2) = 15.473$, $p < 0.001$.

When the gesture was made in the extreme periphery, an odds ratio of 6.333 was found, indicating an 86.4% probability of the linguistic expression being distal. For gestures in the periphery, the odds ratio was 1.5, that is, a 60% probability that the expression was distal; finally, when the gesture was central, the odds ratio was 0.698, linking this gesture to distal language in a mere 41.1% (that is, most of the times, it was linked to proximal language). Additionally, we also made sure that the distribution of these data was not influenced by factors related to the physical position of the speakers. Most of the co-speech gestures that conform our database were performed while seated (68% of distal expressions and 62% of proximal), but no significant differences in gesture distance were found when comparing gesture realizations while standing or seated, $\chi^2(2) = 0.014$, $p = 0.993$.

3 Discussion

3.1 Gesture frequency

Initial research suggested that, when talking about physical distance, speakers contrast peripersonal or near space, which extends to the arm's reach, and extrapersonal or distal space, which covers all the spatial configurations outward from that boundary (Kemmerer 1999). However, more recent empirical studies have demonstrated that the situation is more complex: distal and proximal linguistic distinctions also take into account highly abstract semantic notions that are often combined with several pragmatic, communicative and contextual features (Coventry et al. 2008; Diessel 1999, 2006, 2013)

One of the possibilities to explain the higher frequency of distance-related co-speech gestures could thus be related to the wider scope associated to concept of distance. Since the amount of spatial configurations in distal expressions is greater, speakers may be prompted to perform more gestures in order to express more clearly where the temporal event is located. This can be appreciated when looking at English demonstratives. In English, we find one demonstrative making reference to proximity (*this*) and another one to distance (*that*). The amount of items that we can refer to with *this* are reduced to the ones that are located in the immediate, peripersonal space of the speaker, while *that* comprises the rest of the

Table 2: Data analysis by categories.

Spatial location	Temporal period	Ling. expression	Hits rendered	Repeated clips	False hits	Broken link	Real hits	Voice over	No hands	Visible hands	No gesture	No time gesture	Time gesture
Proximity	Past	Not so/too distant past	59	3	0	0	56	21	16	19	7	5	7
		Not (that/too) far in the past	6	2	0	0	4	0	0	4	0	0	4
		Not (that/too) long in the past	2	1	0	0	1	0	0	1	0	0	1
		Near past	17	0	3	0	14	6	6	2	0	1	1
	Future	Not so/too distant future	171	30	0	1	140	47	54	39	18	11	10
		Not (that/too) far in the future	25	3	0	0	22	5	6	11	2	3	6
		Not (that/too) long in the future	2	0	0	0	2	0	1	1	0	1	0
		Near future	339	33	3	5	298	116	123	59	23	23	13
Distance	Past	Distant past	180	23	4	1	152	64	41	47	15	8	24
		Far in the past	107	17	52	0	38	5	19	14	2	4	8
		Long in the past	20	0	10	0	10	4	4	2	1	0	1
		Remote past	3	0	0	0	3	1	1	1	0	0	1
	Future	Distant future	105	25	2	1	77	22	30	25	5	3	17
		Far in the future	97	12	2	2	81	21	25	35	3	6	26

Table 2: (continued)

Spatial location	Temporal period	Ling. expression	Hits rendered	Repeated clips	False hits	Broken link	Real hits	Voice over	No hands	Visible hands	No gesture	No time gesture	Time gesture
		Long in the future	4	0	1	0	3	0	2	1	0	0	1
		Not in the near future	1	0	0	0	1	0	0	1	0	0	1
TOTAL	–	–	1138	149	77	10	902	312	328	262	76	65	121

spatial configurations (Diessel 1999). In a similar fashion, temporal proximity could be considered to refer to the immediate "temporal" space; there is thus no need to specify further the location of the temporal event, since it is located close to the deictic center. This familiarity of speakers with their immediate space would require a lesser level of specification; on the other hand, distance, covering a wide search space, requires further specification, which is provided by co-speech gesture. The gesture would then be employed to provide a better localization of the temporal period when speakers are talking about temporal distance.

This possibility is also supported by the difference in the variability of the linguistic structures available to express metaphorical distance While distal language presents some variability in the type of linguistic configurations employed, using ADJ + N (e.g., distant future, remote past) and ADJ + PP (e.g., long in the past, far in the future), proximal expressions show less variation; most of them consisting of negations of temporal distal events (e.g., not so distant past) and only two instances of direct ADJ + N containing no negation (e.g., near future). The lack of variation to exclusively refer to temporal proximity in English is supported by the high amount of hits registered by *near future*, while in distal language the hits are distributed more evenly across several expressions (e.g., distant past, far in the past). Since the concept of distance refers to a wider scope of spatial configurations, it could also present more linguistic variation in how to express that distal spatial location. On the contrary, the more familiar proximal space can be expressed with a smaller set of linguistic structures. It is not clear, however, whether the use of a "positive" proximal expression (e.g., near past) is in any way different from the negation of distance to indicate proximity (e.g., *in the not so distant past*). The data sample obtained in this study is not big enough to perform such comparison, although it strongly points to the convenience of performing future studies that look at the interplay of language and gesture when expressing temporal proximity through positive or negative linguistic structures.

An additional, complementary, explanation could be related to aspects of discourse. Utterances integrating a moment in time with a far-away position in space may be more likely to involve emphasis, to which gesture would contribute by providing an additional mark of distance. That is, expressions that indicate that something is far-away in the past or the future may attract more gesture because highlighting this long temporal distance is central to the function of the expression.

3.2 Temporal and gestural distance

The data therefore strongly suggests that there is an analog relation between metaphorical distance, encoded linguistically, and gestural distance, encoded by

the location of the gesture in the gestural space. When we talk about the metaphorical distance of a temporal event and linguistically locate it closer to us, we gesture closer to our body. Similarly, while when we talk about the metaphorical distance of a temporal event and locate it further from us, we tend to gesture further from our body. The space-time mapping does not take place only at conceptual and linguistic levels, but it is also present in the majority of cases at the gestural level too. Consider the following examples:

(1) Here future cast gives us an idea, **high resolution, a better look into the distant future**. You can see here on Saturday … (KCAL 9 News, 26–2–2015).

Figure 9: Both-hand sagittal gesture located in the extreme periphery.

Example 1 shows a distance-related gesture that is performed in the extreme periphery. The preparation of the gesture starts when the speaker says "high resolution" (left). Here he positions both hands laterally to his left, with the left palm facing forwards and the right palm backwards. The gesture starts when the speaker says "better look"; here he moves his left hand forwards, until he performs a total extension of his left arm while saying "distant future" (right). Additionally, the gesture is also accompanied by the motion of the whole body. The speaker moves his whole body forward, as if he were pushing something, while performing the extension of the left arm (Figure 9).

(2) Hey folks, I went out and saw the number one movie in America this weekend, "The Martian," It's the sci-fi thriller that **takes place in the near future** where … (KCBS Late Show with Stephen Colbert, 6–10–2015).

Example 2 presents a case of a gesture that was performed in the center and co-occurred with a proximity-related linguistic expression. Here the speaker prepares the gesture while saying "takes place." He positions both hands in front of him, with the palms facing downwards and performing an extension of some of the

Figure 10: Both-hands sagittal gesture located in the center.

fingers. Then, when he says "near future," he performs a very subtle forward movement with both hands. He slightly rises his left hands and rotates it in order to move it forward, but the final position of the gesture is very close to the initial one (Figure 10).

The most popular categorization scheme of gestures (McNeill 1992, 2005; Kendon 2004) was inspired by Peirce's sign distinction between icon-index-and symbol. However, this partitioning involved a certain simplification of Peirce's original ideas, and as a result, many gesture scholars routinely comment on overlapping problems. Due to their semiotic complexity, different aspects of gestures has been considered to cut across these categories; gestures combine in varying degrees iconicity, indexicality and the conventionality associated with symbols; overlaps have also been reported within the finer distinctions recognized within iconicity: image, diagrammatic and metaphor iconicity (CP 2.277; Farias and Queiroz 2006). In this sense, authors such as McNeill (2005) have resorted to speaking about different mixtures of "dimensions" of iconicity and indexicality (e.g., McNeill 2005; see also Iriskhanova and Cienki 2018 for a similar idea). This is the reason behind proposals such as Mittelberg (2019), who has suggested looking back at the original ideas, taking stock of concepts such as Peirce's Universal Concepts. As Mittelberg (2019: 193) claims: "compared to the basic sign-object relations icon, index, symbol mainly used in previous gesture research, the more fundamental UCs allow one to discern additional facets of how coverbal gestures act as signs." This author shows how using Peirce's UCs (Firstness – possibility, Secondness – actuality, and Thirdness – law, habit) instead of the more basic distinctions of icon-index-and-symbol, can help get at some "additional facets" of how coverbal gestures act as signs (see Colapietro 2001 or Mittelberg 2019 for a detailed discussion). This is especially relevant for the type of gestures this study focuses on, which involve a direct relationship between the distance of the gesture and the distance of the temporal event, thus pointing at a "parallelism" between

facets of these two domains, thus exemplifying a case of "metaphorical iconicity" (Nöth 1990, 1999).

While the different components of the gestures are normally not assigned to any specific functionality (e.g., Bressem and Ladewig 2011), it is clear that this "distance" component clearly has a semiotic function. The distance parameter is part of a metaphorical gesture, after all; as mentioned above, in Peircean terms, this metaphoric gesture represents a parallelism between an abstract object and a physical representamen. It could be argued that in all these cases, the gesture is actually making an iconic reference to the adjective included in these expressions; however, it should be noted that these gestures are not signaling a real point in space, since they co-occur with a *temporal*, metaphorical expressions. The distances that are being expressed by the linguistic expression are not real; they are metaphorical representations of the concept of temporal distance. Time gestures have been found to represent facets of metaphorical motion such as direction or axis (Casasanto and Jasmin 2012; Núñez and Sweetser 2006), and here they would be representing another facet: distance. Common grounds between Peirce's semiotics and cognitive metaphor theory have been discussed by several authors, especially Danaher (1998; see also Mittelberg 2008, 2019). Danaher (1998), for example, was the first author to discuss "the semiotic reality of image-schemas" (Danaher 1998: 190). He argued that image schemata (like the PATH image schema which structures the TIME IS SPACE metaphor on which these gestures are based) can be considered "iconic symbols"; they acquire symbolic status by the formation of habits of interpretation, which gradually lead to the formation of similar interpretants. According to our data, it seems that the space-time conceptual patterns that are found in verbal language are often also present (and mirrored) in gesture: metaphorical distance in time corresponds with spatial gesture distance, while metaphorical proximity in time corresponds with spatial gesture proximity.

Previous literature has successfully investigated the relation between gesture features such as direction or axial position and temporal cognition (see Casasanto and Jasmin 2012; Cooperrider and Núñez 2009; Cienki 1998; Núñez and Sweetser 2006), but this is the first time that a study shows how another gesture feature, distance, is also involved in the representation of temporal metaphorical motion. The predominance of the lateral axis in time conceptualization and the tendency to use a left-to-right direction, for example, are some of the phenomena that have been detected. However, it seems that these are not the only gesture features that play a role in the embodied representation of time through gesture. Together with previous studies that have investigated the relation of gesture distance and temporal duration (Cai et al. 2013), the present dataset shows how the extent or distance covered by gesture also plays an important role in the representation of spatial metaphorical information through co-speech gestures. Together with other

gesture features such as axis and direction (Casasanto and Jasmin 2012), temporal distance is another spatial feature that is (metaphorically) represented through co-speech gestures. Even though this correlation between temporal distance and gesture distance has proved to be remarkably high, there are some cases that present the opposite pattern: sometimes, gestures are performed in the center when talking about temporal distance or in the extreme periphery when talking about temporal proximity. In these cases, the location of the co-speech gesture does not correspond to the temporal distance of their linguistic expression. If we were to think that the gesture is exclusively representing the conceptual relation between time and space, we would expect an almost 100% of the cases reflecting the time-space association. However, we find that around 20% of the gestures in distal language and 15% of the gestures in proximal language are performed following the opposite pattern, making the gesture signaling something completely opposite from the linguistic structure. The exact reasons behind this deviation from the main trend are still unknown, but it is likely that other communicative and informative factors (e.g., discursive or contextual) may have an effect of the type of co-speech gesture performed by the speaker. Most of the cases present speakers in very specific communicative settings (e.g., press conference, discussion in a round table, one-to-one conversations) that may have an impact on the type of gesture that is produced, while in other cases this might be due to the communicative intention of the gesturer (e.g., emphasis, discursive prosody) or contextual factors (e.g., physical landmarks, position of the gesturer).

4 Conclusion

For the first time, we have observational evidence, extracted from ecologically valid data with sufficient quantitative power that shows the relation between the linguistic expression of the concept of "temporal distance" and the distance of co-speech gestures. This is the first time that gesture distance has been analyzed as a key component of time gestures. Just a small amount of data shows incongruence between gesture and speech when addressing distance, with some gestures located in the center when using distal language and some gestures located in the extreme periphery when using proximal language. The causes for the emergence of these incongruencies between gesture and language might be due to external communicative and contextual factors from the time-space mappings, which may modify some gestural properties (such as distance) in order to fulfill other communicative, non-conceptual aims.

Future research on the relation between temporal and gesture distance is required to establish whether the patterns found in this study are also encountered

in other types of temporal language that do not explicitly state the distance of the temporal event (e.g., *back in those days, later than*). However, the present dataset clearly indicates that there is a great coherence between speech and gesture regarding distal and proximal deixis, suggesting that both modalities may be triggers of the same underlying mental simulation, and that enacting this simulation in a coherent way is a core aspect of the process of meaning construction. Additionally, future gesture research on (time) gesture should consider including other understudied features of co-speech gestures, such as gesture distance, as part of the features to analyze.

Research funding: This work was supported by the Spanish Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación & FEDER/UE funds (grant number PGC2018-1551 097658- B-100).

References

Beattie, Geoffrey & Heather Shovelton. 2011. An exploration of the other side of semantic communication: How the spontaneous movements of the human hand add crucial meaning to narrative. *Semiotica* 184. 33–51.

Beattie, Geoffrey. 2016. *Rethinking body language*. London: Routledge.

Bender, Andrea & Sieghard Beller. 2014. Mapping spatial frames of reference onto time: A review of theoretical accounts and empirical findings. *Cognition* 132(3). 342–382.

Boroditsky, Lera. 2000. Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* 75. 1–28.

Boroditsky, Lera. 2001. Does language shape thought?: English and Mandarin speaker's conceptions of time. *Cognitive Psychology* 43. 1–22.

Boroditsky, Lera & Michael Ramscar. 2002. The roles of body and mind in abstract thought. *Psychological Science* 13(2). 185–189.

Bressem, Jana & Silva Ladewig. 2011. Rethinking gesture phases: Articulatory features of gestural movement? *Semiotica* 184(1/4). 53–91.

Bressem, Jana, Silva Ladewig & Cornelia Müller. 2013. Linguistic annotation system for gestures (LASG). In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Jana Bressem (eds.), *Body – language – communication: An international handbook on multimodality in human interaction*, 2nd ed. vol. 38, 1098–1124. Berlin: De Gruyter Mouton.

Bylund, Emanuel, Pascal Gygax, Steven Samuel & Panos Athanasopoulos. 2019. Back to the future? The role of temporal focus for mapping time onto space. *Quarterly Journal of Experimental Psychology* 73(2). 174–182.

Cai, Zhenguang G., Louise Connell & Judith Holler. 2013. Time does not flow without language: Spatial distance affects temporal duration regardless of movement or direction. *Psychonomic Bulletin & Review* 20. 973–980.

Calbris, Geneviève. 2008. From left to right … Coverbal gestures and their symbolic use of space. In Cornelia Müller & Alan Cienki (eds.), *Metaphor and gesture*, 27–55. Amsterdam & Philadelphia: John Benjamin.

Casasanto, Daniel & Lera Boroditsky. 2008. Time in the mind: Using space to think about time. *Cognition* 106. 579–593.
Casasanto, Daniel & Kyle Jasmin. 2012. The hands of time: Temporal gestures in English speakers. *Cognitive Linguistics* 23(4). 643–674.
Cienki, Alan & Cornelia Müller. 2008. *Metaphor and gesture*. Amsterdam, Philadelphia: John Benjamins Pub. Co.
Cienki, Alan, Aliyah Morgenstern, Cornelia Müller & Dominique Boutet. 2016. *Linguistic aspect, tense and gestural movement quality in French, German, and Russian utterances*, vol. 7. Paris: International Society of Gestures Studies.
Cienki, Alan. 1998. Metaphoric gestures and some of their relations to verbal metaphoric expressions. In Jean-Pierre Koenig (ed.), *Discourse and cognition: Bridging the gap*, 189–204. Stanford, CA: CSLI Publications.
Cienki, Alan. 2008. Why study metaphor and gesture? In Alan Cienki & Cornelia Müller (eds.), *Metaphor and gesture*, 5–25. Cambridge: Cambridge University Press. https://doi.org/10.1075/gs.3.04cie.
Cienki, Alan. 2013. Cognitive linguistics: Language and gestures as expressions of conceptualisation. In Cornelia Müller, Alan Cienki, Ellen Fricke, Silva Ladewig, David McNeill & Sedinha Tessendorf (eds.), *Body-language-communication: An international handbook on multimodality in human interaction*, 182–201. Berlin: De Gruyter Mouton.
Cooperrider, Kensy & Rafael Núñez. 2009. Across time, across the body: Transversal temporal gestures. *Gesture* 9(2). 181–206.
Cohen, Doron, Geoffrey Beattie & Heather Shovelton. 2010. Nonverbal indicators of deception: How iconic gestures reveal thoughts that cannot be suppressed. *Semiotica* 182. 133–174.
Colapietro, Vincent. 2001. A lantern for the feet of inquirers: The heuristic function of the Peircean categories. *Semiotica* 136(1/4). 201–216.
Coventry, Kenny R., Berenice Valdés, Alejandro Castillo & Pedro Guijarro-Fuentes. 2008. Language within your reach: Near–far perceptual space and spatial demonstratives. *Cognition* 108. 889–895.
Davis, Stephen Boyd. 2012. History on the line: Time as dimension. *Design Issues* 26(4). 4–17.
Danaher, David. 1998. Peirce's semiotic and cognitive metaphor theory. *Semiotica* 119(1/2). 171–207.
de la Fuente, Juanma, Julio Santiago, Antonio Román, Cristina Dumitrache & Daniel Casasanto. 2014. When you think about it, your past is in front of you: How culture shapes spatial conceptions of time. *Psychological Science* 29. 1682–1690.
Diessel, Holger. 1999. *Demonstratives: Form, function, and grammaticalization*. Amsterdam/Philadelphia: John Benjamins.
Diessel, Holger. 2006. Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics* 17(4). 463–489.
Diessel, Holger. 2013. Distance contrasts in demonstratives. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
Duffy, Sarah. 2014. The role of cultural artifacts in the interpretation of metaphorical expressions about time. *Metaphor and Symbol* 29(2). 94–112.
Farias, Priscila & Joao Queiroz. 2006. Images, diagrams, and metaphors: Hypoicons in the context of Peirce's sixty-six-fold classification of signs. *Semiotica* 162(1/4). 287–307.
Fauconnier, Gilles & Mark Turner. 2008. *Rethinking metaphor*. New York: Cambridge University Press.

Goldin-Meadow, Susan. 2003. *Hearing gesture: How our hands help us think*. Cambridge: Belknap Press of Harvard University Press.
Goldin-Meadow, Susan. 2004. Gestures role in the learning process. *Theory Into Practice* 43. 314–321.
Gullberg, Marianne & Sotaro Kita. 2009. Attention to speech-accompanying gestures: Eye movements and information uptake. *Journal of Nonverbal Behavior* 33. 251–277.
Gullberg, Marianne & Kenneth Holmqvist. 1999. Keeping an eye on gestures: Visual perception of gestures in face-to-face communication. *Pragmatics and Cognition* 7. 35–63.
Gullberg, Marianne & Kenneth Holmqvist. 2002. Visual attention towards gestures in face-to-face interaction vs. on screen. In Ipke Wachsmuth & Timo Sowa (eds.), *Gesture and sign language in human-computer interaction. GW 2001*. Berlin, Heidelberg: Springer.
Gullberg, Marianne & Kenneth Holmqvist. 2006. What speakers do and what listeners look at: Visual attention to gestures in human interaction live and on video. *Pragmatics and Cognition* 14. 53–82.
Iriskhanova, Olga & Alan Cienki. 2018. The semiotics of gestures in cognitive linguistics: Contribution and challenges. *Voprosy Kognitivnoy Lingvistiki* 4. 25–36.
Iverson, Jana & Susan Goldin-Meadow. 1998. Why people gesture as they speak. *Nature* 396. 228.
Kemmerer, David. 1999. "Near" and "far" in language and perception. *Cognition* 73(1). 35–63.
Kendon, Adam. 1980. Gesture and speech: Two aspects of the process of utterance. In Mary R. Key (ed.), *Nonverbal communication and language*, 207–227. The Hague: Mouton.
Kendon, Adam. 2000. Language and gesture: Unity or duality. In David McNeill (ed.), *Language and gesture: Window into thought and action*, 47–63. Cambridge: Cambridge University Press.
Kendon, Adam. 2004. *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
Kranjec, Alexander & Anjan Chatterjee. 2010. Are temporal concepts embodied? A challenge for cognitive neuroscience. *Frontiers in Psychology* 1. 240.
Kranjec, Alexander & Laraine McDonough. 2011. The implicit and explicit embodiment of time. *Journal of Pragmatics* 43(3). 735–748.
Lakoff, George & Mark Johson. 1980. *Metaphors we live by*. Chicago: Chicago University Press.
Lakoff, George. 1987. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
Lapaire, Jean-Remi. 2016. From ontological metaphor to semiotic make-believe: Giving shape and substance to fictive objects of conception with the globe gesture. *Signo* 41. 70.
Levinson, Stephen & Asifa Majid. 2013. The island of time: Yélî Dnye, the language of Rossel island. *Frontiers in Psychology* 4. 61.
Matlock, Teenie, Michael Ramscar & Lera Boroditsky. 2005. On the experiential link between spatial and temporal language. *Cognitive Science* 29(4). 655–664.
McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92. 350–371.
McNeill, David. 1992. *Hand and mind*. Chicago: University of Chicago Press.
McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press.
Mittelberg, Irene. 2019. Peirce's universal categories: On their potential for gesture theory and multimodal analysis. *Semiotica* 228. 193–222.
Mittelberg, Irene. 2008. Peircean semiotics meets conceptual metaphor: Iconic modes in gestural representations of grammar. In Alan Cienki & Cornelia Müller (eds.), *Metaphor and gesture*, 115–154. Amsterdam & Philadelphia: John Benjamins.

Moore, Kevin. 2006. Space-to-time mappings and temporal concepts. *Cognitive Linguistics* 17. 199–244.
Moore, Kevin. 2011. Ego-Perspective and field-based frames of reference: Temporal meanings of FRONT in Japanese, Wolof and Aymara. *Journal of Pragmatics* 43(3). 759–776.
Ng, Melvin R., Winston D. Goh, Melvin J. Yap, Chi-Sing Tse & Wing-Chee So. 2017. We think about temporal words: A gestural priming study in English and Chinese. *Friontiers in Psychology* 8. 974.
Nöth, Winfried. 1990. *Handbook of semiotics*. Bloomington: Indiana University Press.
Nöth, Winfried. 1999. Peircean semiotics in the study of iconicity in language. *Transactions of the Charles S. Peirce Society* 35(3). 613–619.
Núñez, Rafael & Eve Sweetser. 2006. With the future behind them: Converge evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30. 1–49.
Özyrükek, Asli. 2014. Hearing and seeing meaning in speech and gesture: Insights from brain and behaviour. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences* 369. 1651.
Pagán Cánovas, Cristóbal, Javier Valenzuela, Alcaraz Carrión Daniel, Olza Inés & Ramscar Michael. 2020. Quantifying the speech-gesture relation with massive multimodal datasets: Informativity in time expressions. *PloS One* 15(6). e0233892.
Peirce, Charles S. 1931–1966. *The collected papers of Charles S. Peirce*. In Charles Hartshorne, Paul Weiss & Arthur W. Burks (eds.), vol. 8. Cambridge: Harvard University Press [Reference to Peirce's papers will be designated CP followed by volume and paragraph number].
Santiago, Julio, Juan Lupiáñez, Elvira Pérez & Maria José Funes. 2007. Time also flies from left to right. *Psychonomic Bulletin & Review* 14(3). 512–516.
Torralbo, Ana, Julio Santiago & Juan Lupiañez. 2006. Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science* 30. 745–757.
Valenzuela, Javier, Cristobal Pagán-Cánovas, Inés Olza & Daniel Alcaraz. 2020. Gesturing in the wild: Spontaneous gestures co-occurring with temporal demarcative expressions provide evidence for a flexible mental timeline. *Review on Cognitive Linguistics* 18(2). 289–316.
Walker, Esther & Kensy Cooperrider. 2016. The continuity of metaphor: Evidence from temporal gestures. *Cognitive Science* 40(2). 481–495.