Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

Gesturing in the wild: evidence for a flexible mental timeline

Abstract

Psycholinguistic evidence shows that spatial domains are automatically activated when processing temporal expressions. Speakers conceptualize time as a straight line deployed along different axes (mostly sagittal, though also vertical). The use of the lateral axis, which cannot be lexicalized in any language, has nonetheless been attested in temporal tasks in laboratories using a variety of experiments. This leads to the question of what axes are actually at work when conceptualizating time in oral communication.

The present study examines a great number of temporal expressions, taken from television shows, noting their associated co-speech gestures. Our results show that (1) speakers overwhelmingly use the lateral axis; (2) they are not performing simple space-to-time mappings, but are using instead a "timeline", a material anchor which is a far more complex construct and that can explain some of the intricacies and contextual variations shown in the pattern of results.

Keywords

- Space-to-time metaphors; co-speech gestures; time conceptualization; timelines; material anchors

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

22

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

Gesturing in the wild: evidence for a flexible mental timeline

Introduction

Of the many conceptual metaphors proposed across different languages (see Dirven & Pöring, 2002; Kovecses, 2005), the TIME IS SPACE metaphor is probably the most thoroughly studied, having been termed the "fruit fly" of metaphor research (Casasanto, 2009). The literature on this specific topic is quite vast and a full review of all the aspects that have been examined is beyond the scope of the present work. In the TIME IS SPACE metaphor, the domain of time is structured by means of information derived from spatial domains, specifically motion. Linguistic patterns in which temporal information is expressed using spatial terms (e.g. *the end of the world is **coming***, *we have a great future **ahead** of us*, and *we have left those sad days **behind***) are found in a great deal of languages (Radden, 2004), though perhaps not universally (Sinha, Silva Sinha, Zinken & Sampaio, 2011; Le Guen & Balam, 2012). There is now abundant psycholinguistic evidence for the (mostly automatic) activation of spatial

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

domains when processing temporal expressions (Casasanto & Boroditsky, 2008; Ulrich, et al., 2012; Weger & Pratt, 2008).

In the cases studied, the TIME IS SPACE metaphor is realized by means of a line that represents time and can be deployed along different axes, with different directionalities and shapes. In most of the languages analyzed, this line is anchored deictically, i.e., with respect to the speaker, and runs along the sagittal axis. The future is preferentially mapped onto the front of the speaker, (e.g. *the days **ahead of** us*) and the past is mapped onto his/her back (e.g. ***back** in those days*) (Sell & Kaschak, 2011; Torralbo, Santiago & Lupiáñez, 2006; Ulrich et al., 2012). This is not the only possibility, though; in cultures such as the Aymara, the position of future and past on this sagittal axis is reversed, with the future located on the back and the past in front (Nuñez & Sweetser, 2006). Though the sagittal axis is the most frequent, there are also languages, such as Mandarin Chinese, that use a vertical axis; in that case, the past is mapped to the upper positions, while the future is located in the lower part (Boroditsky, Fuhrman & McCormick, 2011; Scott, 1989). It should be noted that the use of non-

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

deictic frames of references has also been attested: Boroditsky & Gaby (2010) show that cultures that use a “geocentric” frame of reference (Levinson, 1996) can align their mental timeline along geocentric coordinates, such as East-West.

The lateral axis

The use of the third axis, the lateral one, is currently the focus of much research, for a number of reasons. The first reason is that no attested language uses this axis, that is, there is no attested use of the spatial labels “right” and “left” to indicate temporal anteriority or posteriority. This tendency is also present in signed languages, which overwhelmingly favor the sagittal axis over the lateral one (Emmorey, 2001). In spite of this, psycholinguistic research has revealed that people do use this axis in their conceptualization of time in laboratory tasks (Ishihara, Keller, Rossetti & Prinz, 2008; Santiago, Lupiáñez, Perez & Funes, 2007). This leads to the question of what axes are actually at work for the conceptualization of time in real-life oral communication. A second point of interest regards the directionality of the flow of time along this lateral axis, which has been found to depend on one specific cultural practice: the direction of reading. Thus, in cultures with left-to-right reading direction (as English), the future is mapped onto the right, while the past is located on the left (Ulrich & Maienborn, 2010); in cultures with a right-to-left reading direction, such as Hebrew or Arab, the opposite pattern is found (Casasanto & Bottini, 2014; Fuhrman & Boroditsky, 2010; Ouellet, Santiago, Israeli & Gabay, 2010).

While standard psycholinguistic research is both extremely useful and indeed quite necessary, it has also been subjected to criticisms such as the ecological validity of its results (Brewer, 2000). One possibility that is hard to rule out from many of these experiments is that the use of a lateral axis is a task artifact; that is, subjects do organize time laterally in those circumstances, as required by the specific experimental task they are asked to carry out in the laboratory, but the use of this pattern would not be found in spontaneous speech. Naturalistic observational studies, on the other hand, tend to lack the compactness and the discriminative power of an experimental setting, but allow instead for a more ecologically valid gamut of results. In the case discussed here, however, there have been so far obvious difficulties in using this type of approach, since it is unclear how observational studies of this type could be informative about the conceptualization patterns used by speakers while referring to time.

The localisation of temporal concepts has also been studied in the gesture dimension. Temporal co-speech gestures, as well as other body parts such as the fingers, the head and the gaze, are often employed to locate the future in front of us or on our right, while the past is behind us and on our left. (Figure 1) This localisation of temporal concepts might be mirroring writing direction in the case of the lateral axis and walking direction in the case of the sagittal axis. (Calbris, 1990; 2011). Several studies have confirmed this tendency (Casasanto & Jasmin, 2012; Cienki, 1998; Cienki & Müller, 2008; Kita, 2009) as well as reporting the use of gestures that combined both the lateral and sagittal axis (Walker & Cooperrider, 2015). Research on sign language points on the same direction, with the inclusion of the vertical axis to indicate months and weeks as they are represented in a calendar (Engberg-Pedersen,1999).

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

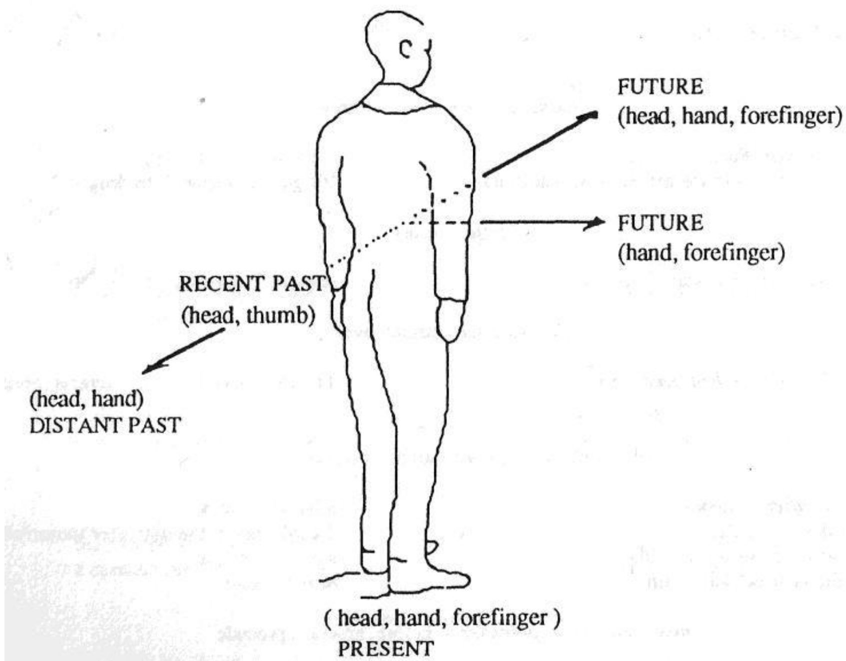

Figure 1. Axes localization in the gesture space (taken from Calbris, 1990)

The present study examines a great number of temporal expressions uttered by speakers in television shows. We analyze the gestural information that speakers deploy while speaking about basic elements of temporal deixis, namely, the beginning, end, and overall duration of a process or period. Gesture information has been shown to reveal conceptualization patterns (Emmorey, 2001; Goldin-Meadow, 2003; McNeill, 1992, 2005), and has proved essential for our understanding of the multimodal nature of oral communication. In the next section we briefly review the role of multimodality in language studies; we also describe the source from which the linguistic examples used in our analysis have been taken: the NewsScape Library of Television News.

Multimodality and the NewsScape Library of Television News

Multimodal data is becoming an increasingly important source of information for the elucidation of the conceptualization patterns used in language. Co-speech gesture has been acknowledged as a highly useful tool for uncovering issues of mental representation for some time now (e.g. Alibali et al., 1999) but the perception of its relevance has increased in recent decades (see Goldin-Meadow & Alibali, 2013 for a recent review). One of the reasons for this renewed interest in multimodality can be attributed to the paradigm change in the study of meaning construction, in which the "decoding" view of meaning has been losing weight in favor of an "inferential" approach to understanding. In this new perspective, there is not a one-to-one coding of meaning into linguistic forms; linguistic forms do not “contain” directly or even “activate” the meanings recovered by the hearer: language vastly underdetermines meaning (e.g. Fauconnier, 1997). Speakers provide words and linguistic expressions merely as cues in the hope that the hearer will integrate them with the rest of the information included in the communicative setting, and reach an understanding of the communicative intention of the speaker. This recovery of the speakers’ intentions is perceived as the basic task of a fellow interlocutor, a task that is carried out with the help of multimodal information, to the extent that “our analysis of others’ intentions can be rampantly incomplete when we lack multi-modal information about their behavior” (Baldwin, 2002, p. 288).

Thus, nowadays, a growing number of scholars agree that communication uses a "composite" (Clark, 1996, calls it "signal"), a complex unit that includes not only the words and linguistic units contained in the utterance, but also the intonation, the gestures used, and the speaker's facial expression and eye gaze. Hearers may use some or all of these cues for the construction of the final meaning. Multimodal information, however, has not been fully incorporated into linguistic research so far, which has chosen to rely almost exclusively on the morphosyntactic information of the signal, disregarding the rest (Jewitt, 2009). This could be due to the technical problems involved in an adequate, objective, and measurable treatment of these different types of information, as well as to the lack of a clear theoretical model than can integrate such disparate sources of information into the same packet, showing their (probably quite) complex interactions.

In the field of multimodality, gesture studies have stressed the need to go beyond explicit verbal utterances in order to uncover the complex, flexible and adaptative patterns that regulate situated, face-to-face communication (McNeill, 1992; Müller et al., 2013, 2014). More specifically, co-speech gesture has been often understood as a "window into the mind", insofar as we see through it the kind of "actions of the mind" – i.e. conceptualization patterns and mappings – that underlie the dynamic use of language (McNeill, 2013, p. 30). In this paper, we assume McNeill's (1992, 2013) definition of gesture as an expressive movement of hands, mainly, or other body parts – e.g. head, eyes – that is part of the process of speaking, usually enacting a certain degree of imagery, though we will focus on hand gestures.

The present study addresses some of these problems by using the NewsScape database of Television News, developed by the Distributed Little Red Hen Lab, an international consortium for research on multimodal communication (see: https://sites.google.com/site/distributedlittleredhen/). NewsScape, which is curated by the Library of the University of California at Los Angeles, contains over 250,000 hours of recorded TV news, with timestamped subtitles/close captioning, synchronized with their corresponding video files. The textual material comprises more than 3 billion words, which makes it bigger than most of the standard corpora used in linguistic research (e.g. the British National Corpus contains roughly 100 million words and the Corpus of Contemporary American English (COCA) about 450 million words). These timestamped subtitles can be searched like any other textual corpus, allowing for the easy location of a video clip with the moment in which the verbal pattern searched was uttered. In this way, the multimodal information associated to a linguistic expression can be accessed, listed, and analyzed with unprecedented speed and ease.

Gesturing the timeline: a multimodal study

Research questions and hypotheses

In the present study, we focus our attention in the following questions:

(1) In real-life oral communication, do people use a mental timeline, typically lateral, to organize temporal meanings when speaking about the start, end, and duration of a process?

(2) If this timeline exists, how fixed or flexible is it? What are the parameters of variation/adaptation of this mental timeline evidenced by gestural imagery?

(3) If we can gather abundant, relevant gestural data about the psychological reality of the timeline and its use in authentic communicative situations, how can these inform

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

theories about mappings between space and time, embodiment, and multimodal meaning construction?

In order to answer these questions, we analyze a wealth of ecologically valid utterances about time that present relevant co-speech gestures. We look for evidence of an underlying conceptualization behind these utterances, that is, evidence that time is organized laterally along an imaginary path or linear object.

Materials and methodology

Data collection

The functionalities of the NewsScape Library include the capacity to perform linguistic searches in ways similar to the more standard textual corpora. This allowed us to select a number of lexical items that refer to temporal points, either the start or the end of a temporal stretch. These items could be reasonably expected to co-occur with

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

some type of related gesture in spontaneous speech. The following is the list of the terms we initially searched for:

Starting points: *beginning, creation, inception, origin, start, onset, outset, initiation, introduction, square one, genesis, inauguration, kickoff, starting point.*

Ending points: *completion, conclusion, termination, culmination, finale, outcome, end, finish, goal.*

These words were combined with prepositions to form prepositional phrases, using "*from, since, to, until*". Ex:

{*since/from*} {*0/its/the/their/its very/their very/*} {*inception/creation/*etc.}

This created a number of “middle level constructions” (Fillmore, Kay & O’Connor, 1988; Golberg, 1995, 2006), which contained fixed items (*from* and *to*) as well as open variables. The final result was a group of demarcative expressions indicating the start/end of durations/processes, or the durations/processes as a whole. Some of these expressions contained the prepositions *since* and *until*, which are not used to indicate spatial relations (except in non-prototypical metaphorical uses). Others contained *from* and *to*, which have a primarily spatial meaning. However, all the *from/to* phrases searched were standard, entrenched English expressions for temporal demarcation, and had no exact "literal" counterparts that would allow for the suppression of the spatial prepositions: e.g. "*from beginning to end*" vs. "**since beginning until end*". In the cases in which they did have a counterpart with *from* or *to*, the two phrases were interchangeable, with no clear semantic distinction between them, as in "*since/from the inception*." This criterion allowed us to rule out the possibility that some of the expressions were perceived by speakers as primarily spatial, in which case they might be gesturing mainly to depict spatial relations, as in *from Madrid to New York* or *from home*. These expressions were searched in all English-speaking television networks available in NewsScape during the period 3-Jan-2005 to 10-Feb-2013 (40 TV-stations). This search had as a result a total of 4,578 hits in the NewsScape database, each with an associated video clip providing the moment in which the phrase was uttered (see Figure 1).

Figure 2. Lateral gesture produced while uttering "from beginning to end". An .mp4 video with examples can be found at: http://blind_location

Filtering

The hits were initially screened by two independent coders. As an initial decision, they discarded all clips where a clearly relevant gesture could not be observed. The reasons for discarding a clip were:

1. There was a voice-over, that is, the speaker did not appear on screen.
2. The speaker was shown from a perspective that did not allow the identification of any hand-gesture. For example, only head and shoulders could be seen (what is technically known as a "medium close-up" or MCP take), with the hands staying out of the screen. Also excluded were cases in which a visual obstacle hindered the inspection of the hand movements.
3. The speaker did not produce any hand movement.
4. The speaker produced a gesture that was obviously not related to the content of the linguistic utterance. This may have been because the gesture was routinely repeated throughout the utterance and thus not linked specifically to the temporal expression, or because it was used to delimit and manage conversational units and not to represent the temporal meaning encoded in the lexical items. It is well known that gestures, along with prosody, are routinely used to structure discourse irrespectively of conceptual content (Kendon, 1995, 2004; Richter, 2014)

This set of criteria led to an initial list of 384 possible time-related gestures (8.4% of the initial hit list). As shown by an examination of the percentages of each of the previous cases for exclusion of a clip (Table 1), just the first two cases explain a full 75.34% of the reasons for discarding a clip: an absence of the speaker on the screen (a voiceover; 43%) or the poor visibility of the hands (31.96%), due to different reasons (position of the speaker, presence of obstacles in the line of vision, close-up shot of the person, etc). Only in 18.95% of the cases could the speaker be observed as uttering the expression while clearly not producing any type of gesture. An additional 5.71% corresponded to gestures that were unrelated to the content of the temporal expression (that is, the speaker was indeed gesturing but the function of the gesture was clearly related to some other communicative purpose, e.g. beat gestures, see Kendon. 2004) or to gestures that were not clearly in sync with the temporal expression and therefore could not be undoubtedly related to the linguistic utterance. Thus, a careful analysis of the filtering process indicates that both the gesture rate (irrespectively of the meaning and function of the gesture) and the percentage of gesture linked to deictic or

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

demarcative time expressions in speech, at least in American English, is in all probability much higher than the one we report here. At the very least, it should be higher for the type of temporal expressions analyzed: start/end of duration or process and demarcation of a process or duration in its entirety.

Voice-over	Hands Non Visible	No gesture	Out-of-sync/unrelated gesture
43.38%	31.96%	18.95%	5.71%

Table 1. Percentages of the different reasons for discarding a clip

Coding

Once we filtered out the utterances with conceptually-relevant gestures, we organized all the information in a database, mainly around two sections. The first one contained information that would allow us to quickly and accurately locate the clip featuring the relevant gesture.

Section 1) Gesture ID (what was searched, and where it was found)

- Precise phrase searched (e.g. *from the inception*)
- Immediate co-text of the phrase
- Program, Date and Time of the clip
- Link to the clip

The second section contained information about the gestures themselves. It should be reminded that in this study we approach co-verbal gesture as a tool to understand the spatial cognition of time; this is why our analysis of each gestural instance is in no way meant to be exhaustive. While we acknowledge the valuable advances and consensus reached in the thorough notation of the categories and features of bodily communication (see, for instance, Bressem, 2013), here we pay attention only to a part of the information provided by the gesture: basically, the axis, direction and general shape of the movement. Thus, instead of offering a fine-grained analysis of the formal features of each gesture, we examine the characteristics that may signal the preferential use of a particular type of timeline in spontaneous speech. This broad and rather schematic analysis of gesture facilitates the realization of a large-scale study aiming to examine the conceptualization patterns that come into play across a great number of instances.

Section 2) Gesture information

- Hand: left, right, bimanual
- Axis: sagittal, vertical, lateral
- Direction: toward (the body), away (from body), downwards, upwards, leftward, rightward, in (both hands meeting in the center), out (separating both hands) (McNeill, 1992)
- Shape: shape of the hand motion (linear or curved)

- Hands interaction: how the hands moved with respect to each other (e.g. one hand staying in a place signalling a landmark and the other moving towards an end point).
- Deictic content: incorporation in the gesture of an object or event present in the joint attentional frame

Though not all these features were used in the final analysis (which focused mainly on the hand used, the axis, and its direction), we decided to keep them for their use in future studies. This second coding phase involved the same two initial coders plus a third one, this time a member of the research team. Disagreements among coders were found in 87 cases (22.65%), which were thus discarded and left us with a final sample of gestures to be studied which amounted to 297 cases (6.49% of the initial hit list rendered by the linguistic search).

Results

Construal

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

We initially looked for the beginning and end points of demarcative temporal expressions, separately: that is, expressions that indicate only one point of the temporal stretch (e.g. *from the start* or *till the end*). However, the results showed a marked preference for their use in combination, therefore referring to the totality of a process (e.g. *from start to finish*). The gesture rate for whole-process expressions was thus much higher. As a result, most of the expressions analyzed made reference to the whole process (almost 75%) and not only the start or the end of the temporal process, as can be seen in Table 2.

Start	16.84% (N=50)
End	8.42% = (N=25)
Whole process	74.75% (N=222)
Total	297 (100%)

Table 2. Percentages and hits for expressions including start, end or whole process

Demarcative-temporal gestures were found much more often when speakers made reference to the whole process than in the other construals (Wald χ^2=72,75, *df*=1, $p<0.001$). Thus, the constructional patterns included expressions with “from/since” and “to/until”, instantiating the middle-level constructions such as [FROM X TO Y] or [SINCE X UNTIL Y]. This basic structure could sometimes be enriched with additional elements, typically referring to the “path”, so that the expression includes a “source”, a “path” and a “goal” (e.g. *from the beginning* ***all the way*** *to the end*). More interestingly, we found that the choices for the basic open variables X and Y constrained each other; that is, they tended to come “in pairs”. For example, when X = *beginning*, then Y = *end*; when X = *start*, Y = *finish*. Just these two middle-level constructions accounted for 80% of the whole process expressions found. Another entrenched pair was *from inception to completion*. Examples of other expressions found were:

FROM beginning of pregnancy *onwards*

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

FROM the beginning *TO the very last frame*

FROM the introduction in the elevator *TO today*

FROM the start of the (marathon/picture/race, etc.) *TO the end*

FROM genesis *TO revelation*

FROM inauguration day *TO where we are*

FROM problem definition *TO implementation*

FROM shower *TO out the door*

This points strongly in the direction of a multimodal view of grammatical patterns; the presence of specific, relatively low-level constructional patterns may include, at least as an optional part of its formal realization, its association with a given gesture. Such a multimodal view of linguistic constructions is currently being explored (Steen & Turner, 2013; Zima, 2014, Blind Reference 2016).

Axis

Out of the three axes, the lateral axis appeared as the most frequent orientation by far (83.16%) as compared with the rest (Wald χ^2=130,67; df=1; p<0.001); the sagittal axis came a very distant second (3.37%), with only two more instances than the vertical axis (3.03%), as shown in Table 3.

Lateral	83.16% (N=247)
Sagittal	3.37% (N=10)
Vertical	3.03% (N=9)
Punctual	10.44% (N=31)
Total	100% (N=297)

Table 3. Percentages of gestures for each axis

Additionally, a number of gestures not clearly connected to any axis were found (10.44%); they were almost always co-occurring with isolated expressions for start or end. We decided to include them in an additional category we termed “punctual” (see Table 3). We discuss each of these types in turn.

As can be seen in Table 3, 31 gestures could not be classified in any of the three axes, because there was no clear motion involved and the gesture was located right in front of the speaker. These gestures were usually tied to a great emphasis of a single point; in this sense, they are close to *beat* gestures, which lack semantic content (Leonard & Cummins, 2011; Wang & Chu, 2013). However, they were not repeated across discourse, and seemed to point instead a specific point in space. Out of the 31 punctual gestures, 16 were connected to the starting point, with phrases including words strongly connected to a single starting point in the temporal sequence, such as *inception* (the most abundant in this type of gestures), *creation, onset* or *starting point.* These words contrast with other ways of making references to the starting point which are more usually coupled with an end point, such as *start* or *beginning*, which appear most frequently with their closing pairs (*finish* or *end*, respectively). Only in one of these punctual gesture expressions the word *beginning* was used, but its punctualness was emphasized with the adverb *very* (from the *very* beginning).

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

The distribution of the axes with respect to the portion of the temporal demarcative stretch highlighted (what we have called “construal”) was, however, not the same. Demarcative expressions that signaled the whole process were more strongly linked to a lateral gesture (90.09%) than in those that signaled only the start (62%) or the end (64%)(Wald χ^2=23.41, *df*=2, p<0.001). Another noticeable difference was found in the percentages of punctual gestures, which were less frequent when the whole process was highlighted (4.5%) than when only one limit of the demarcative expression was highlighted (start only, 32%; end only, 20%).

	Lateral	Sagittal	Vertical	Punctual	Total
Whole Process	90.09% (N=200)	2.70% (N=6)	2.70% (N=6)	4.50% (N=10)	100% (N=222)
Start	62% (N=31)	0	6% (N=3)	32% (N=16)	100% (N=50)
End	64% (N=16)	16% (N=4)	0	20% (N=5)	100% (N=25)

Table 4. Percentage of axis depending on which part of the demarcative expression is highlighted

Directionality

Within the more numerous group, lateral gestures, the great majority (performed with either one or two hands) followed the expected directionality, left-to-right (72.87%); this however means that a sizeable proportion of them, 23.48%, were executed with an inverse directionality (right-to-left), which was a significant difference (Wald χ^2=62,53, *df*=1, p<0.001). Only a small proportion of gestures were carried out by moving both hands away from each other, starting at a central point (3.64%), as shown in Table 5.

Rightward	180 (72.87%)
Leftward	58 (23.48%)
Out	9 (3.64%)
Total	247 (100%)

Table 5. Directionality of lateral gestures

As we just mentioned, 23.48% of the lateral gestures did not follow the typical, rightward directionality, congruent with the direction of reading/writing. There are many possible reasons for this, as shall be mentioned in the discussion. One of the most obvious and also easiest to quantify is hand use; we thus looked at the hand with which these leftward and rightward gestures were carried out. The results are shown in Table 6:

	left-hand	right-hand	both
Leftward	50% (N=29)	32,76% (N=19)	17.24% (N=10)
Rightward	18.33% (N=33)	45.55% (N=82)	36.11% (N=65)

Table 6. Percentages of hand used in leftward and rightward lateral gestures

This means that, though half the leftward lateral gestures (50%) can be accounted for by the primacy of the left hand in the execution, 32.76% of them were carried out exclusively with the right hand (and 17.24% with both hands). A similar pattern was found with rightward gestures: most of them are carried out with the right hand (45.55%) or both hands (36.11%); however, almost one out of five (18.33%) are carried out with the left hand.

Regarding the other two less frequent axes, out of the ten sagittal gestures, nine of them followed an orientation coherent with the more classic front-later time, back-earlier time; the starting point of the temporal sequence was located at deictic origo – in front of the speaker – while the end of the sequence was located in front of and away from the speaker. Only in one case (discussed in the next section) was this pattern reversed. As for the nine vertical gestures, all of them had a downwards orientation (up=earlier, down=later).

Discussion

The lateral timeline: new evidence from spontaneous speech

We have presented what constitutes, to the best of our knowledge, the first large-scale study of temporal co-speech gestures using a multimodal database of authentic/spontaneous discourse. In this corpus study, we examined all the several-thousand occurrences of a specific set of basic demarcative expressions throughout eight years of television news shows from 40 US stations, and filtered out a dataset of almost 300 valid gestures. These utterances and gestures were not the result of elicitation tasks in the lab or of anthropological fieldwork, but arose from speech acts carried out with a great variety of goals and within many different contexts and communicative situations, in a medium that is still the most familiar one for the dissemination of multimodal public discourse around the world.

The clearest result found, at least initially, is the overwhelming prevalence of the lateral axis over the other two in the demarcation of temporal stretches (83.16%). This confirms previous experimental findings with data from spontaneous discourse, and even points more emphatically at the psychological reality of a lateral timeline. The reduced presence of a vertical axis does not strike as too surprising; generally speaking, English does not use vertical spatial terms to organize time, beyond some isolated and somewhat fragmented cases. For example, the word "up" can sometimes be used in combination with motion verbs such as "come" to indicate future (cf. *departure times are coming up*); on the other hand, the word "down" can also be used for future, as exemplified in the expressions *down the line* or *way down the future*. However, English does indicate time in a consistent way using the sagittal axis: *looking **forward** to seeing you*, *a bright future **ahead** of you*, *look **back** in anger*. This is in fact the main means of spatial organization of time in language, since, as mentioned in the initial section, there is no linguistic evidence for the use of the lateral axis in the organization of time (i.e. no "left" or "right" month, to indicate "previous" or "next" or any other temporal-related meaning). Despite this fact, the vast majority of the gestures in the present study use the lateral axis. Gestures which could clearly be ascribed to a sagittal axis amounted to a meagre 3.69%.

This marginal use of the sagittal axis goes against the expectations produced by much psycholinguistic and linguistic research on time-space mappings. The frequency is much lower than the percentages reported so far in the scarce literature on temporal co-speech gestures, even when these studies already report a prevalence of the lateral axis under conditions in which subjects are likely to be unaware of their gesturing (Cooperrider & Núñez, 2009, Walker & Cooperrider, 2015). Casasanto & Jasmin (2012), who compared an experiment requesting deliberate gestures with one in which they sought to elicit spontaneous gestures, found that lateral gestures were three times more frequent than the sagittal ones in the spontaneous condition. Our study indicates that the lateral axis is used 24 times more frequently than the sagittal axis for the type of linguistic expressions searched. In fact, the sagittal axis happens to be as marginal as the vertical one. Although the specific statistical results from those other studies are not comparable with ours due to the differences in the materials and methods, what we found in spontaneous speech does not only seem to confirm, but also to intensify the tendencies shown in the laboratory.

This increase in the prevalence of the lateral axis, up to making the other two marginal for the demarcative expressions studied, is likely to be due, in the first place, to the different verbal cues used. Casasanto & Jasmin (2012), for example, included in their study time expressions that were spatial and, moreover, that clearly referred to the sagittal axis and to an ego-moving or time-moving perspective. The results of Walker & Cooperrider (2015) are even harder to compare with ours, since they used isolated words as cues for their spontaneous gesture study. These metaphorical time expressions are indeed usual and representative, and on most occasions, it is quite unlikely that they are perceived as figurative by speakers. But even though a sagittal verbal expression and a lateral gesture can very well co-occur (as also shown by both Casasanto & Jasmin, 2012 and Walker & Cooperrider 2015), the use of specific spatial language may have had an overall influence on the choice of axis in the gestural modality. This is an interesting issue for further research.

Our study, on the other hand, sought to discard any gesture that could be motivated by explicit spatial vocabulary, even if this meant being extremely restrictive, and only examining a small subset of expressions referring to durations or processes. Therefore, we searched the subtitles in the TV repository only for expressions that do not prompt for an overt mapping between spatial and temporal knowledge. Another reason for our findings of an increased use of the lateral axis may be due to the methodology used. Our results stem from the direct observation of speech in real contexts and in the absence of any elicitation tasks, as opposed to a laboratory setting, which is always susceptible to experimenter’s effects.

Consequently, as a more specific result, beyond the general prevalence of the lateral axis, this gesture study with the NewsScape TV Library suggests that, when the linguistic expression does not encode spatial meaning or motion along any particular axis, the mental disposition of temporal sequences and durations is lateral by default. It was our ability to select the precise expressions to be included in the study, thanks to the computational tools of the NewsScape database, that has allowed us to reach this type of result. When elements explicitly prompting for a sagittal disposition are included (mainly verbal cues that use the default front-back vocabulary for time), this overwhelming prevalence of the lateral axis may decrease (that is an empirical question that still requires much more evidence), although lateral gestures would quite probably

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

still be preferred, as shown by Casasanto & Jasmin (2012) and Walker & Cooperrider (2015) in the category of “spontaneous gestures” obtained in laboratory studies. In fact, our own study, just like Casasanto and Jasmin’s, has also located some examples of lateral gestures produced with verbs indicating motion or with prepositions suggesting a use of the sagittal axis; for example, we have observed speakers uttering expressions such as “move *forward* to its conclusion” that, nonetheless, produce a lateral rightward gesture while saying “forward”.

The reasons for this prevalence of the lateral axis in communicative acts has been a source of speculation. There are a number of physical features, such as postural constraints, the shape of the human body, avoiding intrusions in the space of the interlocutor, or facilitating the perception of two different points in the gesture space, that make it advantageous to present temporal relations along a lateral axis. However, sign languages do not show this preference so clearly (Engberg-Pedersen, 1993; 1999), so there must be some further motivations.

Also, the fact that speakers shift perspectives between language (using a sagittal axis) and gesture (use of lateral axis), or even use both in combination, still needs to be accounted for. This apparent divergence of perspectives in temporal representation is a puzzle for cognitive theories postulating a direct-transfer model for mappings between sensorimotor experience (such as spatial relations) and abstract concepts (such as time). In conceptual metaphor theory (Lakoff & Johnson, 1980, 1999), the TIME IS SPACE set of projections (Lakoff, 1993) cannot account for a combination of axes or perspectives; the fact that the future may be both ahead and on the right is difficult to explain from a direct mapping perspective.

The phenomenon is much less puzzling if analyzed within a network model like the one proposed by blending theory (Fauconnier & Turner, 1998, 2002; Turner, 1996, 2014). In this theory, information does not flow directly from a source domain to a target domain; instead, there is an intermediate space, a *blend*, which, in our case, is what would allow us to accommodate both the sagittal ego-moving perspective (the one used in language), and the lateral perspective obtained from the interaction with an external object acting as a “material anchor” for time. Casasanto and Jasmin (2012) review this model, including the cognitive training or habits provided by the interaction with artifacts that anchor temporal relations, in their list of possible explanations for the data, although they qualify this hypothesis as post hoc and lacking sufficient data for its empirical verification. They also suggest, as a further problem, that it violates the invariance principle of conceptual metaphor theory (Lakoff, 1990), which, in the early stages of the theory, was used to argue that conceptual projections preserve the structure of the source, or (as argued later), seek to avoid a clash between the topologies of source and target (Turner, 1996, pp. 53–54, 108–109). Invariance has been discussed extensively (starting with Turner, 1996, pp. 30–31, 108–9; Stockwell, 1999), and it was precisely the appearance of the blending model that did away with the rigid need to preserve the structure from the source on the target of a conceptual projection. By providing a middle space between source and target (or between any other number of inputs, regardless of whether they are behaving as source or targets), blending theory proposes to explain how the organizing frames fuse and interact within a blend, without insisting on shared generic structures that need to remain fixed. This is a basic difference between a direct transfer model (metaphor) and a network model (blending).

If we decide to adopt the blending hypothesis as a model that can better encompass the data, then the notion of “material anchor” (Hutchins, 2005) becomes crucial. Anchoring a conceptual blend means that a material or perceptual structure constitutes one of the inputs to the network. This allows for the emergent conceptual relations in the blend to become perceptual or material relations at the same time. An example of how material anchors work is found in the cultural practice of queuing. The spatial configuration formed by one person standing behind another based on the order of arrival is a direct way of encoding the temporal sequence. Your spatial location, that is, whether you stand at the beginning, middle or end of the line, allows a direct, perceptually-based way of performing the mental estimation of how long you will have to wait compared to the rest of the people in the line. Queues are just one of the different anchors we use for time or sequentiality: other material artifacts serving this function (that is, making conceptual relations directly accessible as perceptual relations) are calendars, clocks, sundials and, of course, timelines. It is not exactly that the interaction with artifacts influences the blending process, as Casasanto and Jasmin (2012) seem to suggest; rather, in order to use artifacts and their meanings as anchors, a process of blending has to take place: we queue in order to make a particular sequential meaning emerge from the practice; the habit of queueing can then be transported to many other situations in which similar meanings and projections are being established. This process is different from constructing an independent representation (e.g. lexical items such as *tomorrow* or *mañana*) in order to prompt for a temporal meaning. Instead, the material or perceptual anchor seeks to ease the cognitive load by providing a shared ground of joint attention in which complex conceptual relations may be *directly perceived*. This is what a timeline, a clock or a sundial allow us to do: “seeing” time.

Except perhaps for Blind reference 2015 (which is based on a deep psycholinguistic analysis of a single linguistic example), blending has not been explicitly tested against competing models in any study of time representation that we are aware of. But, if we do use the blending model as a working hypothesis, then the idea of the

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

material anchoring of the blend is required to explain the preference for the lateral axis. In this view, the co-speech gesture in the present study is overwhelmingly lateral because it activates a timeline, which is an anchor rather than merely a sign. In material anchors such as timelines, form is related to meaning through a direct perceptual connection. On the other hand, signs, such as words, prompt for meaning using communicative conventions, with a substantial degree of arbitrariness. For example, the words *tempus*, *time*, and *χρόνος* can be different formal cues for approximately the same range of meanings, across different cultural settings. Although there are exceptions, words generally do not use their own perceptual properties, such as length, to cue for differences in magnitude: e.g. Spanish *año* (year) and *siglo* (century) are shorter than *semana* (week), and of course there is no linear mapping associating the days at the beginning of the week with the initial syllable of *semana*, or the weekend with the final one. In anchoring, on the other hand, the relevant perceptual features chosen for the material representation of the blend cannot be handled with the least degree of arbitrariness. Shorter segments on a timeline generally do not mean the same as longer ones. Modifying relative positions of aligned points/events is not likely to be irrelevant either. This is why the differences between timelines, sundials, or clocks across cultures cannot affect the essentials of these representations. Their variation will never be comparable with that of lexical items across languages.

The mental timeline would be providing the perceptual basis for a blend that, as a middle mental space resulting from a network rather than from direct A-to-B projections, would not require the preservation of the source structure (invariance) and would allow for the integration of sagittal and lateral viewpoints, and therefore for the axis disparity between modalities.

Within the blending model, these are, then, the reasons why both perspectives are found in the case of the communication of spatialized time: on the one hand, we have an “internal” point of view, in which the conceptualizer assumes an ego-moving perspective, with its associated sagittal axis. Since forward motion is the most natural and familiar case, meanings such as “*the summer is now behind us*” or “*winter is getting closer*” are understood with utmost ease. On the other hand, and especially when we are talking about the processes themselves or about sequences of events independent of our own perspective (e.g. *the game was followed by a big celebration*), it is also very useful to assume an external viewpoint. In these cases, anchoring the sequence on a lateral timeline makes the access to the temporal relations much easier, rendering them as perceptual relations that the speaker can easily manipulate within the gesture space, while the listener can also perceive them at a glance. Our ability to shift and combine viewpoints is again not unusual or specific to the case at hand: it actually lies at the core of many fields, from multiperspectivity in narratology to perspectivization in linguistics; from point of view in art to perspective in mental imagery, and it has been studied with considerable detail (e.g. Dancygier, 2012; Dancygier & Sweetser, 2012; Turner, 1996). All in all, using the notion of material anchor suggests a gestural basis that can provide evidence for the generalized cognitive habit of downloading cognitive effort on material or perceptual structure, which is a basic proposal of the distributed cognition paradigm (Zhang & Patel, 2006; Clark, 2008; Barrett, 2015).

Flexibility

Yet another fact that needs to be accounted for is the solid presence of gestures with non-standard directionality within the lateral axis. Almost one out of four lateral gestures (23.48%) were based on a timeline in which time flowed from right to left. This is slightly larger than the results reported in both Casasanto & Jasmin (2012) and Walker & Cooperrider (2015), and is in need of some explanation beyond the differences in methodology. As shown in Table 6, the hand used in the realization of the gesture is one of the factors influencing its directionality; of all the incongruent directionality gestures (right-to-left), 50% were done with the left hand. Well-known left-handers such as Jay Leno or Larry King have appeared in our clips using the lateral axis with this "inverse" directionality, using their left hand. Handedness could even influence people while they are using their non-dominant hand: in one of our cases, Bill Gates, another famous left-hander, can be seen gesturing with his right hand in a leftward direction. However, there are other factors that could force the use of a left hand; for example, as found in some cases, the speaker may be holding something in his/her right hand (e.g. a book or a microphone), with the result that the only hand available for gesturing is the left hand. Nonetheless, handedness (be it natural or “forced”) can hardly be the only factor, since 32.76% of those "inverse" gestures were carried out with the right hand, and in another 17.24% of cases, that directionality was indicated by using both hands. A qualitative analysis of our list of gestures provides possible alternative factors, such as the relative position of the interlocutors or even the position of the camera; a more exhaustive study should be done in order to further investigate these possibilities.

It should also be remarked that in order to explain some of the behaviours that people display in these cases, it is sometimes necessary to go beyond the linguistic utterance and its associated co-speech gesture and look at the whole scene. Systematic consideration of the influence of such contextual factors is one of the advantages offered by observational studies with a large audiovisual dataset such as NewsScape. Factors such as the physical position of the interlocutors, the topic being discussed or the presence of certain landmarks in the environment, can influence how people shape their gestures. For example, we found a sagittal gesture in which the

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

directionality was reversed: the beginning of the temporal stretch was located far from the speaker, while the end was at deictic origo. This gesture was produced while the speaker, the coach in a dancing contest show who is being interviewed, utters the phrase "*from beginning to end*". In this case, the speaker is explaining the evolution of his partner and talks about her great progress from the beginning of the show (and he points at the stage where dancing coaching takes place) till the end (pointing at the place where they are standing at the moment of utterance). In this way, the timeline used includes highly deictic start and end points, which cannot be understood outside this context; the physical point where the temporal period being described began is joined by a gesture with the end point of the temporal stretch, which corresponds to the present moment/place. These are indeed extremely interesting results, though our current dataset presents so far only a handful of cases; we hope that future stages of the research will provide enough data to establish more robust hypotheses for the role of contextual elements in timeline gestures

The psycholinguistic evidence (see Santiago, Román & Ouellet, 2011 for a review) seems to point in this direction: we constantly perform mappings between mental structures and store many of those sets of mappings and integrations for further reuse. But what we store are not fixed, ontological correspondences between domains, but rather flexible instructions for partially re-creating conceptual networks that have offered useful functionalities in the past. In the case of the gestures studied, they suggest that a full explanation of how timelines work must necessarily go beyond a simple connection between the broad domains of space and time. It is true that the spatial characteristics of timelines are quite concrete and specific, involving unidimensional vectors in one of the three possible spatial axes and including very clear biases regarding their orientation and their shape (see Blind Reference, 2015). But at the same time, timelines are not fully-stored entities that can be recovered from long-term memory and directly applied "out of the can". Instead, they are dynamic entities that can be flexibly adapted to the discourse situation (Blind Reference, 2013b). In real-world linguistic interactions, speakers adapt dynamically the orientation of timelines, whose gestural construction is contingent on a number of heterogeneous environmental constraints such as the relative position of the speaker and hearer (or of any other third party involved in the communicative event), the handedness of the speaker, or the possible unavailability of one of the hands for the gesture (due for example to the presence of an object in one of the hands). Speakers can also opportunistically include the presence of different spatial landmarks in their timelines, creating orientations that only make sense in a given specific situation. All in all, the model that seeks to explain these effects must go beyond the direct and fixed mappings from a spatial to a temporal domain and must make room for flexible, opportunistic, and goal-directed adaptation.

Therefore, timelines are more than a mere straight line running in the direction of reading or a path along which the speaker or the time units move. In this sense, timelines, as other patterns of conceptual mappings, can be seen as adaptable recipes that can be activated by minimal formal prompts, with prototypical outcomes that can be reached in different ways, and which undergo a satisfaction-constraint process (Blind reference*,* 2015). It should also be noticed that signing for a timeline or expressing temporal relations are never the *only* purpose of the speaker. In this sense, these meanings cannot be isolated from the communicative situation in which they are being built and the type of interaction that each particular setting promotes. As a result, the end of a canonical left-to-right gesture can be used to simultaneously point at a panel or a book on a table, or the prototypical realization of the gesture may be completely altered to adjust to a particular layout, or the performance of the gesture may add stylistic features such as humor or elegance, or it may be only hinted at with minimal motion of hand or head because the situation poses limitations on gesturing.

Conclusion

Gesture accompanying demarcative time utterances observed in real communicative settings privilege the lateral axis up to leaving the other two axes in a marginal position. The study was carried out with gestures co-occurring with expressions that were selected for their minimal space-time metaphorical import (e.g. *since the inception, from beginning to end*, etc.). More explicit spatial, metaphorical language may reduce the prevalence of the lateral axis considerably. Moreover, a considerable flexibility in the use of the lateral timeline was evidenced, especially regarding orientation and the particular realization of the gesture. In combination, all of this evidence suggests a default lateral (but adaptable) timeline used for anchoring diachrony or sequence beyond the semantics of temporal phrases*.* A useful strategy to explain the behaviour of speakers can be to regard this mental timeline as a material anchor for a conceptual blend This anchor would assist both speaker and hearer in the process of meaning construction and facilitate the processing of temporal relations. Understanding timelines in this way allows us to explain features such as the clear predominance of the use of a lateral axis, which follows naturally from its function: ideally, the beginning and end points should be clearly differentiated and easily visible points in space and should merge well with current cultural practices, hence the preference for a left-to-right straight line. It also allows us to explain its flexibility, given the widely-established opportunistic nature of the blending process and its regular accommodation of further inputs or adaptations to local purposes and circumstances.

The present study has been based on the direct observation of the communicative behaviour of speakers, facilitated by the NewsScape Library of Television News tool. Its data presents thus a high degree of ecological validity, since it has not been obtained from a laboratory setting or from anthropological interviews, but from a relatively unconstrained environment, namely, that of speakers communicating freely in TV programs, and generally unaware of their own gesture. Human communication is a complex, multimodal process, involving nested sets of interactional constraints, and only observational studies like the present one will permit the analysis of the different features of communication in a real-life, contextually-rich environment. These studies are thus able to tap into a wealth of information that hints at the underlying complex patterns of interaction. This will surely improve our understanding of both multimodal communication and of conceptualization processes in the near future.

References

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

Alibali, M. W., Bassok, M., Solomon, K. D., Syc, S. E., & Goldin- Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science*, *10*, 327-333.

Baldwin, D. A. (2002). The rise of intentional understanding: Analogies to the ontogenesis of language. In T. Givon & B. Malle (Eds.). *The evolution of language out of prelanguage* (pp. 285-305). Amsterdam: John Benjamins.

Barrett, L. (2015). *Beyond the brain: how body and environment shape animal and human minds*. Princeton (NJ): Princeton University Press.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645.

Blind Reference (2013a).

Blind Reference (2013b).

Blind Reference (2015).

Blind Reference (2016).

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

Boroditsky, L., & Gaby, A. (2010). Remembrances of times east: Absolute spatial representations of time in an aboriginal Australian community. *Psychological Science*, *21*, 1635–1639. doi:10.1177/0956797610386621

Boroditsky, L., Fuhrman, O., & McCormick, K. (2011). Do English and Mandarin speakers think about time differently? *Cognition,* 118(1), 126–132.

Bressem, J. (2013). A linguistic perspective on the notation of form features in gestures. In C. Müller *et al.* (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (vol. I, pp. 1079-1098). Berlin: Mouton de Gruyter.

Brewer, M. (2000). Research Design and Issues of Validity. In Reis, H. and Judd, C. (eds) *Handbook of Research Methods in Social and Personality Psychology* (pp. 11-26). Cambridge: Cambridge University Press.

Calbris, G. (1990). *The Semiotics of French Gesture*. Bloomington and Indianapolis: Indiana University Press.

Calbris, G. (2011). *Elements of Meaning in Gesture*. Amsterdam / Philadelphia: John Benjamins.

Casasanto, D. (2009). When is a linguistic metaphor a conceptual metaphor? In V. Evans & Pourcel, S. (Eds.), *New Directions in Cognitive Linguistics* (pp. 127-145). Amsterdam: John Benjamins.

Casasanto, D. & Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition, 106*, 579–93.

Casasanto, D. & Bottini, R. (2014). Mirror-reading can reverse the flow of time. *Journal of Experimental Psychology: General.* 143(2), 473-9.

Casasanto, D. & Jasmin, K. (2012). The Hands of Time: Temporal gestures in English speakers. *Cognitive Linguistics,* 23(4), 643 – 674.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. In Koenig, J.-P. (Ed.), *Discourse and cognition: bridging the gap* (pp. 189-204). Stanford (CA): CSLI Publications.

Cienki, A. & Müller, C. (2008). Metaphor, Gesture and Thought. In Gibbs, R. W. (ed). *Cambridge Handbook of Metaphor and Thought* (pp. 483-501). Cambridge: Cambridge University Press.

Clark, A. (2008). *Supersizing the mind: Embodiment, Action and Cognitive Extension*. Oxford: Oxford University Press.

Clark, H. H. (1973). Space, time, semantics, and the child. In Moore, T. (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 27-63). New York: Academic Press.

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

Cooperrider, K. and Nuñez, R. (2009) Across time, across the body: transversal temporal gestures. *Gesture* 9, 181–206

Dancygier, B. (2012). *The Language of Stories*. Cambridge: Cambridge University Press.

Dancygier, B., & Sweetser, E. (Eds.). (2012). *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press.

Dirven, R. & Parings, R. (eds.). (2002). *Metaphor and Metonymy in Comparison and Contrast*. Berlin: Mouton de Gruyter.

Engberg-Pedersen, E. (1993). *Space in Danish Sign Language: The semantics and morphosyntax of the use of space in a visual language*. Hamburg: SIGNUM-Verlag.

Engberg-Pedersen, E. (1999). Space and time. In J. Allwood & P. Gardenfors (Eds.), *Cognitive semantics: Meaning and cognition* (pp. 131-152). Amsterdam: John Benjamins Publishing Company.

Emmorey, K. (2001). Space on hand: The exploitation of signing space to illustrate abstract thought. In Gattis, M. (Ed.), *Spatial Schemas and Abstract Thought* (pp. 147–174). Cambridge: The MIT Press.

Fauconnier, G. (1997). *Mappings in Thought and Language*. Cambridge: Cambridge University Press.

Fauconnier, G. & Turner, M. (1998). Conceptual integration networks. *Cognitive Science,* 22(2), 133-187.

Fauconnier, G. & Turner, M. (2002). *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Fauconnier, G. & Turner, M. (2008). Rethinking metaphor. In Gibbs, Jr., R.W. (ed.), *Cambridge Handbook of Metaphor and Thought* (pp. 53-66). New York: Cambridge University Press.

Fillmore, C. J. (1971). *The Santa Cruz lectures on deixis*. Bloomington: Indiana University Linguistic Club

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

Fillmore, C.J., Kay, P. & O’Connor, M. (1988). Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language, 64*, 3, 501–538

Fuhrman, O. & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit nonlinguistic task. *Cognitive Science*, 34, 1430–1451. doi:10.1111/j.1551-6709.2010.01105.x

Goldberg, A. (1995). *Constructions: A Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, A. (2006). *Constructions at Work*. Oxford: Oxford University Press.

Goldin-Meadow, S. & Alibali, M.W. (2013). Gesture's role in speaking, learning, and creating language. *Annual Review of Psychology*, 2013, *123*, 448-453.

Goldin-Meadow, S. (2003). Gesture. *Encyclopedia of Cognitive Science*, volume 2 (pp. 289-293). London: Macmillan Reference Ltd..

Guyau, J.-M. (1988 [1890]). The origin of the idea of time. Reprinted in Michon, J. A., Pouthas, V. & Jackson, J. L. (Eds.), *Guyau and the idea of time* (pp.37-90). Amsterdam: North-Holland.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics, 37*: 1555–1577.

Ishihara, M., Keller, P. E., Rossetti, Y. & Prinz, W. (2008). Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex*, 44(4), 454–461.

Jewitt, C. (ed.) (2009). *The Routledge Handbook of Multimodal Analysis*, London: Routledge.

Kendon, A. (1995). Gestures as illocutionary and discourse structure markers in Southern Italian conversation. *Journal of Pragmatics 23*(3): 247-279.

Kendon, A. (2004). *Gesture: Visible Action as Utterance.* Cambridge: Cambridge University Press.

Kita, S. (2009) Cross-cultural variation of speech-accompanying gesture: a review. *Language and Cognitive Processes*, Volume 24 (Number 2), 145-167.

Kövecses, Z. (2005). *Metaphor in Culture. Universality and Variation*. Cambridge: Cambridge University Press.

Lakoff, G. (1993). The contemporary theory of metaphor. In A. Ortony (ed.), *Metaphor and Thought*, (pp. 202–251). Cambridge: Cambridge University Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh*. New York: Basic Books.

Le Guen O. & Pool Balam L.I. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. Frontiers *in cultural Psychology* (3)271. doi:10.3389/fpsyg.2012.00271

Leonard, T. and Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, 26(10):1457–1471.

Levinson, S. C. (1996). Frames of reference and Molyneux's question: cross-linguistic evidence. In Bloom, P., Peterson, M., Nadel, L. & Garrett, M. (Eds.), *Language and space* (pp. 109-169). Cambridge, MA: MIT Press

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

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

McNeill, D. (2013). Gesture as a window onto mind and brain, and the relationship to linguistic creativity and ontogenesis. In Müller, C. et al. (eds.), *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (vol. I, pp. 28-54). Berlin: Mouton de Gruyter.

Müller, C., Cienki, A, Fricke, E, Ladewig, S. H., McNeill, D. & Bressem, J. (eds.). (2013). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction. Volume I.* Berlin: Mouton de Gruyter.

Müller, C., Cienki, A, Fricke, E, Ladewig, S. H., McNeill, D. & Bressem, J. (eds.). (2014). *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction. Volume II.* Berlin: Mouton de Gruyter.

Nuñez, R. & Cooperrider, K. (2013).The tangle of space and time in human cognition. *Trends in Cognitive Sciences,* May 2013, Vol. 17, No. 5.

Núñez, R. & Sweetser, E. (2006). With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30, 401–450.

Ouellet, M., Santiago, J., Israeli, Z. & Gabay, S. (2010). Is the future the right time? *Experimental Psychology*, *57*, 308–314. doi:10.1027/1618-3169/a000036

Radden, G. (2004). The metaphor TIME AS SPACE across languages. In Baumgarten, N. , Böttger, C., Motz, M. & Probst, J. (Eds.), *Übersetzen, interkulturelle Kommunikation, Spracherwerb und Sprachvermittlung – Das Leben mit Mehreren Sprachen: Festschrift für Juliane House zum 60 Geburtstag* (pp. 225-238). Bochum, Germany: Aks-Verlag.

Richter, N. (2014). Structuring discourse: Observations on prosody and gesture in Russian TV-discourse. In C. Müller et al. (eds): *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (vol. II, pp. 1392-1400). Berlin: Mouton de Gruyter.

Santiago, J., Lupiáñez, J., Perez, E., & Funes, M. J. (2007).Time (also) flies from left to right. *Psychonomic Bulletin & Review*, *14*, 512–516.

Santiago, J., Román, A. & Ouellet, M. (2011). Flexible foundations of abstract thought: A review and a theory. *Spatial Dimensions of Social Thought*, 41–110.

Scott, A. (1989). The Vertical Dimension and Time in Mandarin. *Australian Journal of Linguistics* 9, 295–331.

Sell, A. J. & Kaschak, M. P. (2011). Processing time shifts affects the execution of motor responses. *Brain & Language, 117*, 39–44.

Shapiro, L. (2011). *Embodied Cognition*. New York: Routledge.

Sinha, C., Silva Sinha, V. da, Zinken, J. & Sampaio, W. (2011). When Time is not Space: The social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Language and Cognition, 3*(1): 137-169.

Stockwell, P. 1999. “The Inflexibility of Invariance.” *Language and Literature* 8 (2): 125-142.

Torralbo, A., Santiago, J. & Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cognitive Science* 30(4), 745–757

Traugott, E. C. (1978). On the expression of spatio-temporal relations in language. In Greenberg, J. H. (Ed.), *Universals of human language*, *vol.3: Word structure* (pp.369-400). Stanford: Stanford University Press.

Turner, M. (1996). *The literary mind*. New York: Oxford University Press US.

Turner, M. (2014). *The origin of ideas: blending, creativity, and the human spark*. New York: Oxford University Press US.

Steen, F. & Turner, M. (2013). Multimodal Construction Grammar. In Borkent, M., Dancygier, B. & Hinnell, J. (Eds.), *Language and the Creative Mind* (pp. 255-274). Stanford: CSLI Publications.

Ulrich, R., Eikmeier, V., Vega, I., Ruiz Fernández, S., Alex-Ruf, S., & Maienborn, C. (2012). With the past behind and the future ahead: Back-to-front representation of past and future sentences. *Memory & Cognition,* 40(3), 483–495.

Ulrich, R., & Maienborn, C. (2010). Left–right coding of past and future in language: The mental timeline during sentence processing. *Cognition, 117*, 126–138.

Post-print de:
VALENZUELA, Javier, Cristóbal PAGÁN, 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 of Cognitive Linguistics*, 18(2), 289-315. https://doi.org/10.1075/rcl.00061.val

Walker, E. & Cooperrider, K. (2015). The continuity of metaphor: evidence from temporal gestures. *Cognitive Science* 40 (2), 481-495.

Weger, Ulrich W. & Jay Pratt. 2008. Time flies like an arrow: Space–time compatibility effects suggest the use of a mental timeline. *Psychonomic Bulletin and Review* 15(2), 426–430.

Wang, L. and Chu, M. (2013). The role of beat gesture and pitch accent in semantic processing: an ERP study. *Neuropsychologia,* 51(13):2847–55.

Whorf, B. L. (1956). Language, mind, and reality. In Carroll, J.B. (Ed.), *Language, thought and reality. Selected writings of Benjamin Lee Whorf* (pp. 246-270). Cambridge (Mass): MIT Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9: 625–636.

Zhang, J. & Patel, V.L. (2006). Distributed cognition, representation, and affordance. *Pragmatics & Cognition* 14 (2), 333-341.

Zima, E. (2014). Gibt es multimodale Konstruktionen? Eine Studie zu [V(motion) in circles] und [all the way from X PREP Y]. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* (15), 1-48.