ARTICLE

Time as space vs. time as quantity in Spanish: a co-speech gesture study

Daniel Alcaraz Carrión[1],* 🆔 and Javier Valenzuela[2] 🆔

[1]University of Wisconsin-Madison; [2]Universidad de Murcia
*Corresponding author. Email: alcarazcarri@wisc.edu

Abstract

There is a distinction between languages that use the DURATION IS LENGTH metaphor, like English (e.g., *long time*), and languages like Spanish that conceptualise time using the DURATION IS QUANTITY metaphor (e.g., *much time*). The present study examines the use of both metaphors, exploring their multimodal behaviour in Spanish speakers. We analyse co-speech gesture patterns in the TV news setting, using data from the NewsScape Library, that co-occur with expressions that trigger the DURATION IS QUANTITY construal (e.g., *durante todo* 'during the whole') and the DURATION IS LENGTH construal in the *from X to Y* construction (e.g., *desde el principio hasta el final* 'from beginning to end'). Results show that both metaphors tend to co-occur with a semantic gesture, with a preference for the lateral axis, as reported in previous studies. However, our data also indicate that the direction of the gesture changes depending on the construal. The DURATION IS QUANTITY metaphor tends to be performed with gestures with an outwards direction, in contrast with the DURATION IS LENGTH construal, which employ a left-to-right directionality. These differences in gesture realisation point to the existence of different construals for the concept of temporal duration.

Keywords: metaphor; time; quantity; gesture; space

1. Introduction

The mental domain of time has attracted a great deal of attention throughout history. Different aspects of the conceptualisation of time have been investigated from disciplines such as philosophy, neuroscience, psychology, linguistics, anthropology, and history (e.g., Block, 1990, Munn, 1992, Damasio, 1994, Savitt, 1995, or Evans, 2004, inter alia). Out of this myriad of studies, it is the use of space to conceptualise time that emerges as the most agreed upon fact. Since time is an abstract domain, humans are forced to exploit other more concrete domains in order to structure it, and space seems to be the preferred domain for this temporal structuring across cultures.

CrossMark

How exactly space is used in this temporal conceptualisation has many variants. Most languages use an egocentric frame of reference (Radden, 2004), with temporal concepts such as *front*, *back*, *up*, or *down*, which locate the temporal event spatially with respect to the position of the conceptualiser (also known as D-Series or Deictic Time). This is not the only possibility: languages which use different frames of spatial reference (e.g., geocentric or field-based) can exploit their own specific spatial mechanisms for the conceptualisation of time, as Boroditsky and Gaby (2010) have documented for some Australian languages. Speakers are also able to arrange temporal events with respect to each other, which gives rise to another agreed-upon category of temporal reference: Sequential time or B-series (McTaggart, 1908; Clark, 1973). This conceptualisation mode is based on the arrangement of a number of events in relation to each other, establishing relations of anteriority and posteriority in time (Moore, 2014). Expressions such as *before that* or *after that* are perhaps the most common to express Sequential time in English.

Both Deictic time and Sequential time conceptualisations are thus used to locate temporal events with respect to the speaker or with respect to each other, placing temporal events in different spatial axes and directions. In most languages, like English, the most frequent is the sagittal axis: the future is normally located in front of the conceptualiser, and the past is located behind (e.g., *it will happen far ahead in the future* and *it happened way back in the past*; Casasanto & Jasmin, 2012). Other languages, such as Aymara (Moore 2011; Núñez & Sweetser, 2006) or Vietnamese (Sullivan & Buil, 2016), place the past in front of them, since it has been witnessed and it is known, and the future behind, since it is still unknown and cannot be seen. Within-language variations depending on task and context have also been reported (Callizo et al., 2020). As for the other two axes, languages such as Mandarin Chinese use the vertical axis, the past being located in the upper part of the axis and the future below (Boroditsky, 2001). The remaining axis, the lateral, does not figure explicitly in the linguistic repertoire of any oral language; however, many psycholinguistic experiments (e.g., Santiago et al., 2007) have proved that it does play an important role in structuring time. Additionally, a great number of studies have found that, in spontaneous speech, speakers conceptualise time using the lateral axis, as evidenced by their gesturing patterns (Casasanto & Jasmin, 2012; Pagán Cánovas et al., 2020; Valenzuela et al., 2020).

Finally, another aspect of time that can be expressed through language is temporal duration. While in Deictic and Sequential time the aim is to arrange sequences of temporal events, temporal duration refers to the magnitude of the temporal event itself (Núñez & Cooperrider, 2013). English tends to express the concept of temporal duration as the separation (or length) between two points in a timeline. This way of understanding temporal duration relies thus on *one* dimension of space: the *length* or distance between two points (Dolscheid et al., 2013):

(1) We are the extreme team, covering this thing **from beginning to end** and all over the place. (2013-02-08_1600_US_KABC_Good_Morning_America, NewsScape Library).

On the other hand, different languages, such as Greek and Spanish, tend to express the concept of temporal duration by employing a different type of metaphor: the *quantity* metaphor, related to the TIME IS SUBSTANCE metaphor (Lakoff & Johnson, 1980, 1999; Casasanto 2008; Baksteen, 2016; Bylund & Athanasopolous, 2017;

Valenzuela & Alcaraz Carrión, 2020). In this case, rather than conceptualising the duration of an event as one-dimensional length, temporal duration is understood as a unit located in a three-dimensional space, whose quantity is measured (Dolscheid et al., 2013):

> (2) Ni suman los votos hoy, ni los van a sumar en **mucho tiempo**. (2020-12-10_0300_ES_La-1_Telediario_2, NewsScape Library).
> 'The votes don't add up now, and they won't add up in a **long time**.' (lit. 'much time')

The preference for a one-dimensional conceptualisation of duration in English (DURATION IS LENGTH), and the preference for a three-dimensional conceptualisation of duration (DURATION IS QUANTITY) in languages such as Greek and Spanish has been supported by psycholinguistic studies (Casasanto 2008; Bylund and Athanosopolous, 2017) and more recently by corpus studies (Baksteen, 2016; Alcaraz Carrión & Valenzuela, 2021). This does not mean, of course, that the two duration metaphors cannot co-exist within the same language, but rather than each language tends to favour one construal over the other, perhaps related to the fact that English is a monochronic culture while Spanish (and other Latin cultures) are considered polychronic (Hall, 1976; Valenzuela & Alcaraz Carrión, 2020).

Co-speech gesture studies in the domain of time have been a great source of evidence to study temporal cognition (Cienki, 2008). However, there are no gesture studies that have addressed the concept of temporal duration, nor these different metaphorical realisations in the domain of time. On top of that, most of the research on time gestures (semantically related, iconic/metaphorical gestures that co-occur with temporal expressions) has focused on English; in comparison, fewer studies have addressed other languages such as Chinese (Gu et al., 2019), Aymara (Nuñez & Sweetser, 2006), Yucatec Maya (Le Guen & Pool Balam, 2012; Le Guen, 2017) or languages that employ geocentric frames of reference, such as Yupno (Levinson, 2003; Boroditsky & Gaby, 2010). Spanish belongs to this under-researched group of languages in the area of the multimodality of temporal conceptualisation.

Our aim in this paper is to further expand the research that has so far been performed on time gestures by performing an observational study on the speech–gesture realisations of DURATION IS LENGTH and DURATION IS QUANTITY metaphors in Spanish. Spanish presents a particularly interesting case for the study of duration metaphors, since it reportedly favours the conceptualisation of temporal duration as a three-dimensional unit (Casasanto, 2008; Valenzuela & Alcaraz Carrión, 2020). We hypothesise that these different construals of temporal duration will have an influence on the patterns of co-speech gestures used by speakers. If true, the gesture evidence provided by this study would provide converging evidence for the existence of different construals of temporal duration in cognition and communication.

2. Methodology

2.1. Dataset

The data have been extracted from the Red Hen Lab dataset *NewsScape* (http://newsscape.library.ucla.edu/), a multimodal television news repository managed by the UCLA library and the Case Western Reserve University library, co-directed by Mark Turner and Francis Steen (https://www.redhenlab.org/). The original

component of the Red Hen Lab dataset is the UCLA NewsScape Archive, which was developed by the Department of Communication at UCLA. Worldwide researchers have supplemented that Archive substantially. The NewsScape dataset now contains around 500,000 hours of television news from 2004 until the present day, as well as over 5 billion words of television subtitles. The main source of audiovisual data are television news programmes in English, but there is also a wide range of television programmes from a great variety of languages available, including French, Italian, German, Arabic, and Hindi. For the purpose of this study, we will specifically focus on the NewsScape Spanish subset, which encompasses both Mexican Spanish and European Spanish; the total size of this Spanish subset is over 100 million words.

The type of co-speech gestures obtained from the NewsScape library presents a more spontaneous, ecologically valid set of data from co-speech gestures than could be obtained in laboratory settings. The gestures that are produced in a television setting are difficult to predict, and they are mostly performed unconsciously and spontaneously. These co-speech gestures are produced in a wide variety of situations: two-by-two interviews, TV anchors, stand-up comedians, and general conversations. This type of dataset is probably as close as one can get to the type of co-speech gestures produced by the general population in everyday settings.

2.2. Linguistic search criteria

Since the NewsScape repository contains audiovisual as well as textual data formed by the television subtitles, it is possible to search for a particular set of linguistic expressions and have access to the exact moment in which those structures were uttered on television. That is, we can use verbal language as an entry point to observe the multimodal signals that co-occur with concrete linguistic expressions.

The data collection process begins with the creation of two linguistic search packages that address the two temporal construals that are compared in this study: DURATION IS LENGTH and DURATION IS QUANTITY. The initial set of linguistic expressions was based on previous research that linguistically compared these construals (Valenzuela & Alcaraz Carrión, 2020).

The first search packet contained linguistic expressions that triggered a DURATION IS LENGTH construal, with temporal duration being conceptualised as the amount of space that is comprehended between two points in a line (Casasanto et al., 2004; Casasanto, 2010; Bylund & Athanasopoulos, 2017). The most frequent linguistic structure that evokes this construal is *de/desde X a/hasta Y* ('from X to Y'; see Valenzuela & Alcaraz Carrión, 2020). To ensure that the temporal meaning was always present in the linguistic searches, the *X* component of the construction contained a noun with temporal meaning (e.g., *principio, comienzo, origen, hoy* …; 'start, beginning, origin, today …'), while the Y component could contain any type of noun or clause with a temporal meaning (e.g., *el equipo alemán apretó **desde el principio hasta marcar el primer gol** al filo del descanso*, 'the German team pushed **from the beginning until scoring the first goal** just before the break'; NewsScape Library, 2013-10-03 ES La-1 Noticias 24 horas). Owing to the low frequency of these types of expressions in the Spanish section of the multimodal corpus (Valenzuela & Alcaraz Carrión, 2020), we expanded the search list with a total of 15 different linguistic expressions that resulted in a total of 629 hits in the NewsScape dataset (see Appendix 1).

The second search packet contained expressions that were related to the domain of quantity. The linguistic structures employed for this search were again extracted from Valenzuela and Alcaraz Carrión (2020): *durante todo/a el/la* and *durante todos/as los/las* ('during the whole/during all'). Since this type of construal is reported to be favoured in Spanish (see Casasanto et al., 2004; Bylund & Athanasopolous, 2017; Valenzuela & Alcaraz Carrión, 2020), these two linguistic structures presented a higher frequency than the ones employed in the DURATION IS LENGTH construal, with a total of 405 hits in the Spanish NewsScape repository (see Appendix 1).

2.3. Data processing

The linguistic searches performed in the DURATION IS LENGTH and DURATION IS QUANTITY search packages were classified according to the multimodal data that they provided: it was necessary to discriminate between clips that contained useful gestural information and those cases that did not. Thus, we reviewed each clip individually and filtered the data through a three-stage process.

The first filtering stage focused on removing all instances in which the multimodal information in the video clip could be considered to be noise data. This included cases in which the same clip was repeated several times (e.g., the same clip of a politician during a speech in different TV channels), cases in which the video or audio of the clip was not functioning correctly (e.g., misaligned audio), and cases in which the expression uttered by the speakers did not correspond to the expression recorded in the television subtitles.

The second filtering phase focused on choosing clips which presented a speaker on screen whose hands were visible. Many of the video files extracted from the NewsScape repository contain instances in which the speaker is not present on screen, and instead we find an image or video with a voice-over. Additionally, there are also cases in which the shot of the camera does not include the hands of the speaker, or part of the gesture is performed out of frame.

The third and last stage further divided the remaining video clips, which showed clearly the speakers' hands while uttering the expression, into three sub-categories: clips in which the speakers did not perform any type of co-speech gesture; clips in which the speakers performed a gesture that was not related to the temporal domain (e.g., a beat gesture, or a self-adaptor; see McNeill, 1992, for more precise definitions); and clips in which speakers performed a co-speech gesture that was related to the temporal domain (see Appendix 1 for a full breakdown of the filtering process).

This last category, with video clips that contained a co-speech gesture with a temporal meaning, was then further analysed in terms of gesture axis and gesture direction, also noting the hands that were used during the gesture and whether they were free or busy (holding a microphone or some papers, for example).

Thus, when performing a co-speech gesture, speakers may deploy it along the lateral, sagittal, or vertical axis. Depending on the axis chosen by the speaker, the gesture can have different directions. Lateral gestures can be performed leftwards or rightwards, but they also can be performed inwards or outwards when speakers use both hands (that is, both hands moving simultaneously towards each other or in the opposite direction). Sagittal gestures can be performed away from the speaker or towards the speaker. Vertical gestures can have an upward or a downward directionality. Additionally, there are instances in which the gesture does not present axial movement; the speaker signals a

point in space without a clear axial movement (but this gesture is not a beat gesture because it is not repeated through discourse; we classified these as punctual gestures). We also identified the hand(s) that the speakers used when performing the co-speech gesture: right, left, or both hands. In the case of gestures that were performed with both hands, we also coded for handshape as well as palm orientation (cf. Bressem, 2013). Lastly, we identified whether the hands of the speakers were free or were busy holding an item, and, in the case of being busy, which hand was holding the item (see Appendix 2 for a full breakdown of the data analysis). We adopted a conservative approach and did not include bodily movements that are often considered extensions of gestures, such as head movement or gaze, since several of the features relevant for our analysis (e.g., axis, direction of gesticulation, movement towards or away from the body) can only be consistently extracted from hand gestures.

3. Results

The linguistic searches performed in the DURATION IS LENGTH and DURATION IS QUANTITY search packages resulted in 1034 hits in the NewsScape repository (629 and 405 hits, respectively). The first filtering phase removed a total of 303 (30%) video clips from the dataset, which left 731 hits. A total of 502 (69% of the remaining data) video clips were excluded in this second filtering phase, resulting in 229 video clips in which the speakers (and their hands) were clearly visible. During the last filtering phase, a total of 66 hits (28.82%) were classified as being an instance of *no gesture*, 21 hits (9.17%) as cases with a non-temporal gesture, and finally 142 hits (62%) that contained a co-speech temporal gesture (see Appendix 1).

For each of the two categories (and for their individual expressions), we measured the gesture frequency ratio. In the DURATION IS LENGTH search packet, speakers performed 61 temporal gestures (64.21% of the times in which their hands were visible), while they performed 10 non-temporal gesture (e.g., beat gesture) (10.52% of the occasions), and did not perform any gesture in 24 of the clips of the instances (25.26%). Similarly, speakers in the DURATION IS QUANTITY search packet performed 81 temporal gestures (60% of the occasions in which the hands were visible); only 8 clips (9.17%) contained a non-temporal gesture, and 36 cases (28.82%) did not present a gesture realisation. These results are consistent with previous research on the frequency of speech–gesture co-occurrence (Pagán Cánovas et al., 2020; Woodin et al., 2020). No statistically significant differences were observed between our two search packages, nor across the individual expressions of each category regarding gesture frequency (Figure 1)

This last category of clips including a temporal co-speech gesture was further classified in terms of gesture axis, direction, and gesturing hand, in both the DURATION IS LENGTH and DURATION IS QUANTITY search packages, as follows.

For the DURATION IS LENGTH search packet, there was a total of 61 co-speech gestures (42.95% of the total gesture dataset). The axis employed by these gestures was distributed as follows: 51 gestures (83.6%) in the lateral axis; 4 gestures (6.55%) in the sagittal axis; 1 gesture (1.6%) in the vertical axis; 4 gestures (6.55%) with no axis (punctual); and 1 gesture (1.6%) whose axis was unclear. For each axis, the direction employed by the gesture was analysed as follows: in the lateral axis (Figure 2), 33 gestures were performed with a rightward direction (64.7%), 15 gestures with a leftward direction (29.4%), and 2 gestures (3.9%) with an outward motion (there were

Gesture frequency across construals
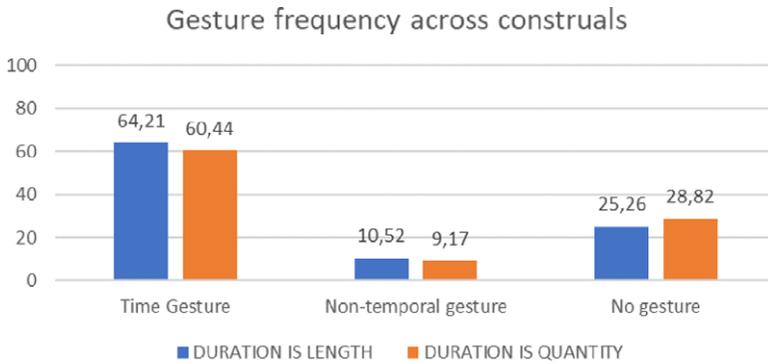
Fig. 1. Gesture frequency percentage in DURATION IS LENGTH and DURATION IS QUANTITY construals.
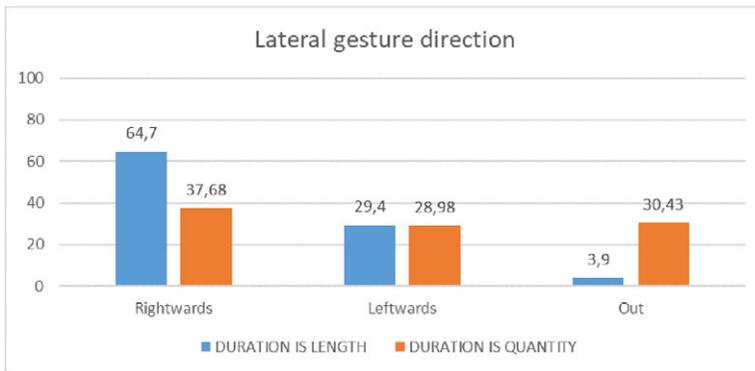
Fig. 2. Lateral gesture direction percentage in DURATION IS LENGTH and DURATION IS QUANTITY construals.

no cases of lateral gestures with an inward direction). For the sagittal axis, all gestures were performed with an away from the body direction. The only vertical gesture that was found had a downward motion. Gestures that presented no axial movement or presented a circular motion were not included in the final analysis. Finally, the hands employed to perform these co-speech gestures were distributed as follows: 31.03% of the co-speech gestures employed the right hand; 34.48% the left hand; and 34.48% both hands. No patterns of interest were observed across the different individual expressions regarding axis, direction, or gesturing hands. The only expression that presented some slight differences was *desde hoy hasta …* 'from today until …', which was the only one that employed sagittal gestures (but with an overall preference for the lateral axis; see Appendix 2). Concerning the shape and the orientation of the gestures performed with both hands, most of the cases (95.23% of all the both-hands gestures) presented open palms facing each other (similar to the one presented in Figure 3 with both hands), with the remaining 4.77% of the cases containing instances of both palms being together, facing each other. The data were analysed by two different coders. An inter-coder reliability measure was calculated for each of the three gesture features, with coders presenting perfect agreement on the analysis of the gesture axis ($\kappa=1$), and a strong agreement in both gesture direction ($\kappa=0.807682$) and hand selection ($\kappa=0.879518$).

For the DURATION IS QUANTITY search packet, there were 81 temporal co-speech gestures (57.04% of the total dataset). The distribution of the axis employed was as follows: 69 (85.18%) lateral, 2 (2.46%) sagittal, 2 (2.46%) vertical, 2 (2.46%) with no axis, 3 (3.7%) with circular motion, and finally 3 (3.7%) of gestures whose axis was unclear. Lateral gestures were performed with a rightward (26 gestures, 37.68%), leftward (20 gestures, 28.98%), and outward (21 gestures, 30.43%) motion; again, there were no cases of gestures with an inward motion (Figure 2). There were only two instances of sagittal gestures, one with an away from the body motion and another with a towards the body motion. Similarly, the two instances of vertical gestures presented an upward and a downward direction, respectively. The most frequent hand employed to perform these gestures was a combination of both hands (35.52%), followed by the right (34.21%) and the left hand (30.26%). Lastly, we also coded the co-speech gestures performed with both hands in terms of hand shape and orientation. All the cases included were performed with an open palm: 85.18% of the instances showed gestures with both palms facing each other; in 11.11% of the instances the palms had a downward orientation and in 3.7% an upward orientation.

This analysis was also performed by two coders, which presented a strong agreement for axis (κ=0.84106) as well as direction (κ=0.887588), and a perfect agreement for hand analysis (κ=1). Concerning the different individual expressions, there were no significant differences among them in terms of gesture axis, direction, or shape.

4. Discussion

The first fact that is immediately apparent from the data is that there is no statistical difference in the frequency of temporal co-speech gesture realisation between those linguistic expressions associated with a DURATION IS LENGTH construal and those linked to a DURATION IS QUANTITY construal. In Spanish, 64.21% of the DURATION IS LENGTH linguistic expressions co-occur with a temporal gesture; and a very similar ratio (60.44%) is found for the expressions associated to a DURATION IS QUANTITY construal. Thus, the data obtained from Spanish seem to mirror the gesture frequency ratio reported in previous quantitative studies. For example, Valenzuela et al. (2020) found that that 72.33% of the co-speech gestures that were performed with TIME IS SPACE temporal expressions in English were strongly related to the temporal meaning present in speech (though that study included different types of temporal expressions, not only durational). Such a high frequency of co-speech gesture production has also been found in another similar study, focused on the domain of number (Woodin et al., 2020). The authors found that expressions such as *tiny number*, *small number*, or *huge number* were accompanied by a co-speech gesture 78.4% of the time. It thus seems that in both English and Spanish, speakers tend to produce a co-speech gesture when they employ language that makes reference to some sort of abstract magnitude: the measurement of time as the length between two points in space, through the DURATION IS LENGTH metaphor in English and Spanish, the QUANTITY IS SIZE metaphor represented through different hand configurations in English, and finally the DURATION IS QUANTITY metaphor in Spanish, in which the temporal period is conceptualised as units located in a three-dimensional space (Casasanto 2008; Valenzuela & Alcaraz Carrión, 2020). The embodiment of abstract thinking in co-speech gestures is not limited to time: other domains, such as space (see Alibali, 2005, for an overview) have also been reported to be represented through co-speech

gestures. Another domain also included in this list is the domain of number. which frequently employs co-speech gesture as well as other graphical representations such as number lines and graphs (Gunderson et al., 2015; Alibali et al., 2019; Pier et al., 2019; Woodin et al., 2020).

Concerning the axial location of the temporal co-speech gestures, speakers showed an overwhelming preference for the lateral axis. Similarly to other studies based on English (Cooperrider & Núñez, 2009; Casasanto & Jasmin, 2012; Valenzuela et al., 2020), our data show that Spanish speakers also show a preference for the lateral axis when gesticulating about temporal concepts. In fact, part of this data can be compared with the equivalent linguistic structures associated with DURATION IS LENGTH temporal gestures in Pagán Cánovas et al. (2020). In their study, they found that, in television discourse, 80% of demarcative expressions (e.g., *from start to finish*, *from beginning to end*) in English were accompanied by a gesture along the lateral axis. Spanish shows almost the exact same tendency, with demarcative expressions such as *de principio a fin* 'from beginning to end' or *desde ahora hasta* 'from now until …' using the lateral axis 83.6% of the time. It is unknown, however, whether the same equivalent ratios can also be observed when looking at co-speech gesture realisations of the DURATION IS QUANTITY metaphor in English, since no data are available on that topic.

There are several reasons that could explain the preference for Spanish speakers (as well as English speakers) to employ the lateral axis when gesturing about time. First, as reported by several gesture studies scholars (Calbris, 2008; Casasanto & Jasmin, 2012; Burns et al., 2019; Pagán Cánovas et al., 2020), the lateral axis offers the largest amount of gesture space for the speaker to perform gesticulations, as well as being the most anatomically comfortable axis for performing hand movements. The other alternatives (sagittal or vertical gestures) are more challenging to use effectively, especially in the television news context, since they are often more difficult to discern. This does not mean that these gesticulations are not present on television, but that their frequency is significantly lower. Finally, there are a wide number of cultural practices and artefacts for time conceptualisation in Spanish that, similarly to English, tend to employ the lateral axis. To name a few: writing direction (Casasanto & Bottini, 2010), timelines (Valenzuela et al., 2020), and time-related diagrams or calendars are often presented with a lateral orientation in which the past is located on the left and the future is located on the right.

In the results reported so far, both construals have presented an almost identical distribution of gesture realisation, both in terms of frequency and in terms of axis. However, our data indicate that there are some clear differences between the DURATION IS LENGTH and the DURATION IS QUANTITY construals concerning the direction of the lateral co-speech gestures. Gestures that co-occurred with a DURATION IS SPACE expression were performed most of the time with a rightward direction: 64.7% of them followed a left-to-right motion. Consider the following example:

(3) Nos vemos este domingo a las diez de la mañana, a las nueve si vive en el centro de Estados Unidos. – **Desde el principio hasta el final**. (2010-11-13, US KEMX Noticiero Univisión, NewsScape Library; clip available here <https://gallo.case.edu/go/dac/0001/2010-11-13_0230_US_KMEX_Noticiero_Univision_1770-1777.mp4> or scan QR code.)
'See you this Sunday at 10 in the morning, nine if you live in the centre of the United States – **From beginning to end**.'

Fig. 3. Lateral rightward gesture with a ᴅᴜʀᴀᴛɪᴏɴ ɪs ʟᴇɴɢᴛʜ construal.

As can be observed in Figure 3, the speaker (right) performs a lateral gesture with a rightwards (left-to-right) motion by performing two chopping-like gestures in synchrony with the beginning (*desde el principio*) and the end (*hasta el final*) of the event that is indicated in speech. This two-stroke chopping-like gesture is very similar to the type of co-speech gestures that are performed with demarcative expressions in English (see Pagán Cánovas et al., 2020), in which speakers locate the start of the event on their left and the end of the event on their right with two different strokes.

Speakers, however, can also reverse the flow of time, performing these gestures with a leftward motion and locating the first point in the sequence of events on their right and the second one on their left, albeit less frequently (29.4%). The number of incongruent instances of time gestures in Spanish seems to be slightly higher than those reported in previous studies (Casasanto & Jasmin, 2012, reported that 26% of the gestures performed in the lateral and sagittal axes were incongruent; Pagán Cánovas et al., 2020, reported 27.13% incongruent gestures in the lateral axis). The motivation(s) behind the production of incongruent gestures are not still clear, although some of the reasons could be linked with handedness, viewpoint, or the spatial arrangement of the speakers (Valenzuela et al., 2020).

Co-speech gestures performed with the ᴅᴜʀᴀᴛɪᴏɴ ɪs ǫᴜᴀɴᴛɪᴛʏ construal show a very clear difference in gesture direction when compared to the ᴅᴜʀᴀᴛɪᴏɴ ɪs ʟᴇɴɢᴛʜ construal: while 37.68% of the gestures presented a rightward direction, and 28.98% showed a leftward direction, the case of 'out' gestures, in which speakers moved both hands away from each other across the lateral axis, accounted for 30.43% of the co-speech gestures. When comparing the gesture direction between the ᴅᴜʀᴀᴛɪᴏɴ ɪs ǫᴜᴀɴᴛɪᴛʏ and the ᴅᴜʀᴀᴛɪᴏɴ ɪs ʟᴇɴɢᴛʜ construals, we can observe some statistically significant differences, namely the difference between the rightward gesture direction in length (64.7%) vs. quantity (37.68%) construals (χ^2=8.49, p<.01) as well as the difference in the 'out' direction in the length (3.9%) and quantity (30.43%) construals (χ^2=13.27, p<.001).

Consider the following example:

(4) Las lluvias se quedan con nosotros durante **todo el fin de semana**, y probablemente también durante el inicio de la próxima semana. (2018-09-08, ES 24h Noticias 24h, NewScape Library; clip available here <https://gallo.case.edu/go/dac/0002/2018-09-08_1800_ES_24h_Noticias_24h_1785-1794.mp4> or scan QR code.)
'The rain will stay with us during **the whole weekend**, and probably also during the start of the next week.'

In Figure 4, the speaker begins the clip with both hands together close to his waist. Then, immediately after saying *todo* 'whole, all', he performs an outward gesture with both hands while saying *fin de semana* 'weekend', moving them shoulder-width and opening both palms as if holding an item between them.

We believe that these differences in gesture directionality found for the DURATION IS LENGTH vs. DURATION IS QUANTITY expressions (the former preferring left-to-right gestures and the later presenting a high number of out gestures) can be explained by the fact that speakers are representing through their co-speech gestures two different conceptualisation patterns of time. The DURATION IS LENGTH metaphor tends to evoke a timeline, mapping the temporal sequence onto a straight line. Each of the temporal events is placed on this timeline, and the temporal duration is measured as the distance between the start and the end point signalled in the timeline. As has been mentioned before, numerous studies have shown that this type of construal favours a left-to-right co-speech gesture: the gesticulation signals the beginning and the end of the timeline following the canonical direction of time in Western languages (past-on-left and future-on-right). This preference for the lateral axis has also been argued to be related to cultural artefacts such as writing direction (Casasanto & Bottini, 2014), and even pragmatic motivations, since the use of sagittal gestures could invade the interlocutor's personal space (Cienki, 1998). Another factor mentioned is the greater information value offered by lateral gestures in comparison to sagittal gestures (for a full review on these issues, see Casasanto & Jasmin, 2012). On the other hand, the DURATION IS QUANTITY construal would make speakers more likely to conceptualise

Fig. 4. Lateral out gesture with a DURATION IS QUANTITY construal.

temporal duration as a unit that can be held between their two hands. Temporal duration would then no longer be conceptualised in terms of the physical distance between point A and point B, but rather as a unit that is deployed in a three-dimensional space (Casasanto, 2008). When looking at the hand shape, especially in the cases when gestures were performed with both hands, we found a very similar configuration in both construals: an open-palm gesture with the two palms facing each other. In the case of the DURATION IS LENGTH construal, this hand configuration could be used to mirror the boundaries of the timeline that is being represented, with each of the hands presenting sequentially the *beginning* and the *end* of the event, with a left-to-right motion. In the case of the DURATION IS QUANTITY construal, the same hand configuration could be representing the 'quantity' of time, which the speaker is holding with both hands; in this last case, both hands move and arrive at their positions at the same time.

It should also be pointed out that, even though the *out*-gesture realisations in the DURATION IS QUANTITY construal already comprise a third of the gestures, a closer look at the dataset suggests that their frequency could be even higher. Since all the data that have been analysed belong to television news, there are many occasions (almost half of the annotated gestures) that show a speaker holding an item (typically a microphone or some papers) while performing a co-speech gesture with their free hand (see Appendix 2). We believe that the presence of an item on the hands of the speakers should be taken into account, since speakers will not be able to perform an *out* gesture with both hands if they are holding a microphone (often close to their face) with one of their hands. Hence, we have analysed more closely the instances in which speakers performed a co-speech gesture while having one of their hands busy.

If our initial hypothesis is correct, we would find that speakers still try to perform an *out* gesture using only their free hand. Thus, the busy hand would remain stationary in a location, while the free hand would perform a stroke in the opposite direction. If this is the case, speakers that employ a DURATION IS QUANTITY construal would then be more likely to perform a *leftward* gesture when their right hand is busy, aiming to create a container between the free hand and the busy hand, while speakers that employ a DURATION IS LENGTH construal would try to mirror the past-left/future-right direction of time, and accordingly perform a *rightward* gesture with their left hand when the right hand is busy. When the left hand is busy and the gesture is performed with the right hand, both construals would favour a rightward gesture, since this direction is congruent both with the flow of time in the case of the DURATION IS LENGTH construal, and with the creation of a container in the case of the DURATION IS QUANTITY construal.

Indeed, this is what our data show. When looking at the DURATION IS LENGTH co-speech gestures performed with the left hand because the right hand was busy, 31.57% of the cases presented a rightward gesture (congruent of flow of time). In contrast, no such instances were found in the DURATION IS QUANTITY co-speech gestures (0%), and all the cases favoured a left hand-leftward gesture.

Thus, a variety of evidence points to the existence of a favoured DURATION IS QUANTITY construal by Spanish speakers when discussing temporal duration. First, Spanish speakers tend to overestimate amounts of time when presented with quantity-related stimuli, as shown by Casasanto (2010) and Bylund and Athanasopolous (2017). Second, corpus research has shown that Spanish speakers tend to express temporal duration mainly by using quantity metaphors, while English seems to favour length metaphors to express this notion (Figure 5). And third, the use of a LENGTH vs. QUANTITY metaphor when talking about temporal duration in Spanish also
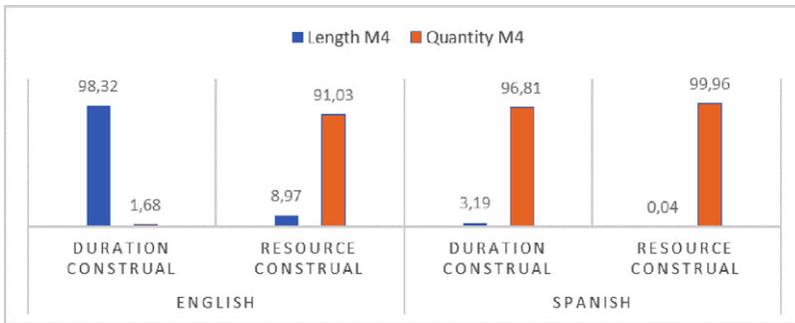
Fig. 5. Frequency of length and quantity metaphors in the duration and resource construal in English and Spanish (extracted from Alcaraz Carrión & Valenzuela, 2021).

involves changes in the type of gesture realisations that speakers make. This evidence points in the direction that two very different metaphors are employed when conceptualising time, or rather that different metaphors are employed to refer to *different* aspects of time, at least in Spanish. While English mostly relies on the TIME IS SPACE metaphor to refer to most temporal meanings, conceptualising time in one-dimensional space, Spanish speakers tend to change from a one-dimensional conceptualisation of time when talking about temporal sequences, to a three-dimensional conceptualisation of time when talking about temporal duration.

The high frequency of gesture in this domain could be seen as supporting the need for a more concrete referent (e.g., timelines, shapes, containers) as a way to more easily anchor abstract concepts. Alternatively, and more in consonance with enactive views of cognition, these gestures could be seen as 'material anchors' which allow the offloading of conceptual information onto the world (Hutchins, 2005; see also Goodwin, 2000). In this sense, the gestures are not external expressions of some internal state of affairs, but the way in which the cognizer *enacts* the meanings: they do not merely express a given conceptualisation, but create it as they are realised (Alcaraz Carrión & Valenzuela, 2021).

Further research on the possible Whorfian consequences of these differences in conceptualisation would also be extremely interesting. This path was started by Casasanto et al. (2004) and their line-growing experiments, but given the apparent psychological soundness of the distinction it would be worth further possible exploration of the Whorfian effects along the lines of Filipovic (2011). It should also be mentioned that work on the DURATION IS QUANTITY construal has been limited to only a handful of languages (English, Spanish, Greek, and Indonesian), and, in this sense, future work with a wider number of languages (and time construals) would be needed in order to clarify a number of issues, such as whether the high gesture production ratio is universal across languages regardless of the temporal construal they are using, or possible commonalities in the multimodal profiles or construal details found in different languages. Concerning research in Spanish, widening the empirical base of this incipient work with more expressions would allow us to fill in the details with more precision, such as the different variations of hand use, or the investigation of other possible factors that could affect the realisation of the gesture: type of concrete expression, position of speaker in the scene, role of type of genre/discourse, etc. Hopefully, as our Spanish multimodal database keeps growing, such studies will become possible.

All in all, the data in this study show clear evidence of a different scheme in Spanish for the conceptualisation of time, which adds to previous research that has shown differences between English and Spanish (Casasanto et al., 2004, Valenzuela & Alcaraz Carrión, 2020). The present study is thus another example of how gestural data can work together with linguistic frequency data to uncover conceptualisation patterns with a high degree of precision.

As a final note, it should be mentioned that the database that has been introduced in this study is, to the best of our knowledge, the first database of time gestures in Spanish (including both European Spanish and Mexican Spanish), employing contextualised instances of speech–gesture interaction on television to study the speech–gesture relation in Spanish. The interest in the use of large datasets with contextualised communicative situations has increased in recent years (Valenzuela & Alcaraz Carrión, 2020; Woodin et al., 2020) and their usefulness has become evident. Even though limited in quantity and scope, the present dataset establishes the bases for the quantitative study of Spanish co-speech gestures. We hope it will encourage other researchers to perform studies in a similar line in order to deepen our understanding of multimodal communication in Spanish.

Conflicts of Interest. We have no conflicts of interest to disclose.

References

Alcaraz Carrion, D. & Valenzuela, J. (2021). Duration as length vs amount in English and Spanish: a corpus study. *Metaphor and Symbol* 36(2), 74–84.

Alibali, M. W. (2005). Gesture in spatial cognition: expressing, communicating, and thinking about spatial information. *Spatial Cognition and Computation* 5(4), 307–331.

Alibali, M. W., Nathan, M. J., Boncoddo, R. & Pier, E. (2019). Managing common ground in the classroom: teachers use gestures to support students' contributions to classroom discourse. *ZDM Mathematics Education* 51(2), 347–360.

Baksteen, M. (2016). *Do longer lines take longer? Reconsidering the cognitive reflections of spatial duration metaphors: Evidence from Dutch.* Unpublished master's thesis, Leiden University.

Block, R. A. (1990). Models of psychological time. In R. A. Block (ed.), *Cognitive models of psychological time* (pp. 1–35). Hillsdale, NJ: Lawrence Erlbaum Associates.

Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology* 43, 1–22.

Boroditsky, L. & Gaby, A. (2010). Remembrances of times east: absolute spatial representations of time in an Australian aboriginal community. *Psychological Science* 21(11), 1635–1639.

Bressem, J. (2013). A linguistic perspective on the notation of form features in gestures. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill and S. Teßendorf (eds), *Body – language – communication / Körper – Sprache – Kommunikation. Vol. 1. Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of linguistics and communication science* (HSK 38.1) (pp. 1079–1098). Berlin & New York: Mouton de Gruyter.

Burns, P., McCormack, T., Jaroslawska, A. J., O'Connor, P. A. & Caruso, E. M. (2019). Time points: a gestural study of the development of space–time mappings. *Cognitive Science* 43(12), e12801.

Bylund, E. & Athanasopoulos, P. (2017). The Whorfian time warp: representing duration through the language hourglass. *Journal of Experimental Psychology: General* 146(7), 911–916.

Callizo-Romero, C., Tutnjevic, S., Pandza, M., Ouellet, M., Kranjec, A., Ilic, S. *et al.* (2020). Temporal focus and time spatialization across cultures. *Psychonomic Bulletin & Review* 27(6), 1247–1258.

Calbris, G. (2008). From left to right: coverbal gestures and their symbolic use of space. In A. Cienki & C. Müller (eds), *Metaphor and gesture* (pp. 27–53). Amsterdam: John Benjamins.

Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning* 58, 63–79.

Casasanto, D. (2010). Space for thinking. In V. Evans & P. Chilton (eds), *Language, cognition, and space: state of the art and new directions* (pp. 453–478). London: Equinox.

Casasanto, D., Boroditsky, L., Phillips, W., Greene, J., Goswami, S., Bocanegra-Thiel, S. *et al.* (2004). How deep are effects of language on thought? Time estimation in speakers of English, Indonesian, Greek, and Spanish. *Proceedings of the Annual Meeting of the Cognitive Science* 26.

Casasanto, D. & Bottini, R. (2010). Can mirror-reading reverse the flow of time? In C. Hölscher *et al.* (eds), *Spatial cognition* (pp. 1342–1347). Berlin & Heidelberg: Springer.

Casasanto, D. & Bottini, R. (2014). Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General* 143(2), 473–479.

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 counterparts. In J. P. Koenig (ed.), *Discourse and cognition: bridging the gap* (pp. 189–205). Stanford, CA: CSLI.

Cienki, A. (2008). Why study metaphor and gesture? In A. Cienki & C. Müller (eds), *Metaphor and gesture* (pp. 2–25). Amsterdam: John Benjamins.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (ed.), *Cognitive development and the acquisition of language* (pp. 27–63). New York: Academic Press.

Cooperrider, K. & Núñez, R. (2009). Across time, across the body: transversal temporal gestures. *Gesture* 9(2), 181–206.

Damasio, A. (1994). *Descartes' error: emotion, reason, and the human brain*. New York: Avon Books.

Dolscheid, S., Shayan, S., Majid, A. & Casasanto, D. (2013). The thickness of musical pitch: psychophysical evidence for linguistic relativity. *Psychological Science* 24(5), 613–621.

Evans, V. (2004). *The structure of time: language, meaning and temporal cognition*. Amsterdam: John Benjamins.

Filipović, L. (2011). Speaking and remembering in one or two languages: bilingual vs. monolingual lexicalization and memory for motion events. *International Journal of Bilingualism* 15(4), 466–485.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics* 32 (10), 1489–1522.

Gu, Y., Zheng, Y. & Swerts, M. (2019). Which is in front of Chinese people, past or future? The effect of language and culture on temporal gestures and spatial conceptions of time. *Cognitive Science* 43(12), e12804.

Gunderson, E., Spaepen, E., Gibson, D., Goldin-Meadow, S. & Levine, S. (2015). Gesture as a window onto children's number knowledge. *Cognition* 144, 14–28.

Hall, E. T. (1976). *Beyond culture*. New York: Doubleday.

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics* 37(10), 1555–1577.

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

Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: the embodied mind and its challenge to Western thought*. New York: Basic Books.

Le Guen, O. (2017). Una concepción del tiempo no-lineal en dos lenguas: el maya yucateco colonial y actual y la lengua de señas maya yucateca. *Journal de la Société des Américanistes* e15327.

Le Guen, O. & Pool Balam, L. I. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. *Frontiers in Psychology* 3, e00271.

Levinson, S. C. (2003). *Space in language and cognition: explorations in cognitive diversity*. Cambridge: Cambridge University Press.

McNeill, D. (1992). *Hand and mind*. Chicago, IL: University of Chicago Press.

McTaggart, J. M. E. (1908). The unreality of time. *Mind* 17, 456–473.

Moore, K. (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.

Moore, K. (2014). *The spatial language of time metaphor, metonymy and frames of reference*. Amsterdam: John Benjamins.

Munn, N. D. (1992). The cultural anthropology of time: a critical essay. *Annual Review of Anthropology* 21(1), 93–123.

Núñez, R. & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in Cognitive Sciences* 17(5), 220–229.

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(3), 401–449.

Pagán Cánovas, C., Valenzuela, J., Alcaraz Carrión, D., Olza, I. & Ramscar, M. (2020). Quantifying the speech–gesture relation with massive multimodal datasets: informativity in time expressions. *PLOS ONE* 15(6), e0233892.

Pier, E. L., Walkington, C., Clinton, V., Boncoddo, R., Williams-Pierce, C., Alibali, M. W. & Nathan, M. J. (2019). Embodied truths: how dynamic gestures and speech contribute to mathematical proof practices. *Contemporary Educational Psychology* 58, 44–57.

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

Santiago, J., Lupiáñez, J., Pérez, E. & Funes, M. (2007). Time (also) flies from left to right. *Psychonomic Bulletin & Review* 14(3), 512–516.

Savitt, S. (ed.) (1995) *Time's arrows today, recent physical and philosophical work on the direction of time.* Cambridge: Cambridge University Press.

Sullivan, K. & Bui, L. T. (2016). With the future coming up behind them: evidence that time approaches from behind in Vietnamese. *Cognitive Linguistics* 27(2), 1–29.

Valenzuela, J. & Alcaraz Carrión, D. (2020). Temporal expressions in English and Spanish: influence of typology and metaphorical construal. *Frontiers in Psychology* 11, e543933.

Valenzuela, J., Pagán Cánovas, C., Olza, I. & Alcaraz Carrión, D. (2020). Gesturing in the wild: evidence for a flexible mental timeline. *Review of Cognitive Linguistics* 18(2), 289–315.

Woodin G., Winter B., Perlman M., Littlemore J. & Matlock T. (2020). 'Tiny numbers' are actually tiny: evidence from gestures in the TV News Archive. *PLOS ONE* 15(11), e0242142.

Appendix 1

	LX expression	Filtering Stage 1				Filtering stage 2			Filtering stage 3			Time gestures
		Hits	Repeated	False	Broken	Real hits	Voice-over	Non-visible hands	Visible hits	No gesture	Non-Spatial	Gestures
DURATION IS LENGTH	desde el principio hasta	45	5	0	1	39	12	18	9	1	1	7
	de principio a fin	238	40	1	14	183	135	12	36	10	5	21
	desde el comienzo hasta	6	1	0	0	5	2	3	0	0	0	0
	desde el inicio hasta	1	0	0	0	1	0	1	0	0	0	0
	desde el origen	2	0	0	0	2	1	0	1	0	0	1
	desde entonces hasta	41	11	2	1	27	19	3	5	1	1	3
	desde hoy hasta	110	19	0	6	85	48	15	22	7	0	15
	desde su comienzo hasta	1	0	0	0	1	1	0	0	0	0	0
	desde su origen hasta	1	0	0	0	1	0	1	0	0	0	0
	de comienzo a fin	0	0	0	0	0	0	0	0	0	0	0
	de inicio a fin	1	0	0	0	1	0	1	0	0	0	0
	desde el principio al	14	2	2	1	9	1	4	4	3	0	1
	desde mañana hasta	20	3	0	4	13	8	1	4	1	1	2
	desde ayer hasta	10	0	1	0	9	4	2	3	1	1	1
	desde ahora hasta	139	111	2	4	22	10	1	11	0	1	10
DURATION IS QUANTITY	durante todo/a el/la	301	11	25	2	263	119	34	110	36	8	66
	durante todos/as los/las	104	29	3	2	70	36	10	24	6	3	15
	Total	1034	232	36	35	731	396	106	229	66	21	142

Appendix 2

	LX expression	Total Gestures	AXIS						DIRECTION									HANDS			BUSY HAND	
			L	S	V	P	U	C	R	L	IN	OUT	A	T	UP	D	O	L	R	B	YES	NO
DURATION IS SPACE	desde el principio hasta	7	7	0	0	0	0	0	7	0	0	0	0	0	0	0	0	3	1	3	3	4
	de principio a fin	21	19	0	1	1	0	0	13	5	0	1	0	0	0	1	0	7	8	4	12	9
	desde el comienzo hasta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	desde el inicio hasta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	desde el origen	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	1
	desde entonces hasta	3	3	0	0	0	0	0	1	2	0	0	0	0	0	0	0	3	5	5	1	2
	desde hoy hasta	15	9	4	0	1	1	0	7	2	0	0	4	0	0	0	0	0	0	0	4	10
	desde su comienzo hasta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	desde su origen hasta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	de comienzo a fin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	de inicio a fin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	desde el principio al	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	desde mañana hasta	2	2	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	1	1	1
	desde ayer hasta	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0
	desde ahora hasta	10	9	0	0	1	0	0	3	5	0	1	0	0	0	0	0	4	3	4	3	7
DURATION IS QUANTITY	durante todo/a el/la	66	57	2	2	2	1	2	21	16	0	19	1	1	1	1	4	20	21	22	29	37
	durante todos/as las/los	15	12	0	0	0	2	1	5	4	0	2	0	0	0	0	1	3	5	5	3	12
	Total	142	120	6	3	6	4	3	59	35	0	23	5	1	1	2	5	43	44	46	57	84

Axis L: lateral; S: sagittal; V: vertical; P: Punctual; U: Unclear; C: Circular *Hand* R: right; L: left; B: both *Direction* R: rightwards; L: leftwards; A: away from body; T: towards body; D: down; O: other.