Getting the ball rolling: the cross-linguistic conceptualization of caused motion*

GUILLERMO MONTERO-MELIS

Centre for Research on Bilingualism, Stockholm University

AND

EMANUEL BYLUND

Centre for Research on Bilingualism, Stockholm University; Department of Swedish, Linnaeus University; Department of General Linguistics, Stellenbosch University

ABSTRACT

Does the way we talk about events correspond to how we conceptualize them? Three experiments ($N = 135$) examined how Spanish and Swedish native speakers judge event similarity in the domain of caused motion ('He rolled the tyre into the barn'). Spanish and Swedish motion descriptions regularly encode path ('into'), but differ in how systematically they include manner information ('roll'). We designed a similarity arrangement task which allowed participants to give varying weights to different dimensions when gauging event similarity. The three experiments progressively reduced the likelihood that speakers were using language to solve the task. We found that, as long as the use of language was possible (Experiments 1 and 2), Swedish speakers were more likely than Spanish speakers to base their similarity arrangements on object manner (rolling/sliding). However, when recruitment of language was hindered through verbal interference, cross-linguistic differences disappeared (Experiment 3). A compound analysis of all experiments further showed that (i) cross-linguistic differences were played out against a backdrop of commonly

represented event components, and (ii) describing vs. not describing the events did not augment cross-linguistic differences, but instead had similar effects across languages. We interpret these findings as suggesting a dynamic role of language in event conceptualization.

KEYWORDS: language and thought, linguistic relativity, event cognition, caused motion, similarity arrangement, verbal interference, cross-linguistic differences.

1. Introduction

To what extent does our conception of reality bear a relation to the language we speak? This question, known as the principle of linguistic relativity, or the Whorfian hypothesis (Whorf, 1956), has been subject to extensive research and fierce debate within the cognitive sciences, notably during the past two decades (for recent overviews, see Gleitman & Papafragou, 2012; Regier & Kay, 2009; Wolff & Holmes, 2011). Classic test beds of the Whorfian hypothesis comprise the domains of colour, time, space, and objects. Though traditionally under-represented, the domain of motion has attracted increasing attention over the past decade, with a growing number of studies investigating cross-linguistic differences in motion event cognition (Athanasopoulos et al., 2015; cf. Athanasopoulos & Bylund, 2013; Bylund, Athanasopoulos, & Oostendorp, 2013; Flecken, Von Stutterheim, & Carroll, 2014; Gennari, Sloman, Malt, & Fitch, 2002; Loucks & Pederson, 2011; Malt, Ameel, Imai, Gennari, Saji, & Majid, 2014; Papafragou, Hulbert, & Trueswell, 2008; Papafragou, Massey, & Gleitman, 2002; Papafragou & Selimis, 2010; Trueswell & Papafragou, 2010; von Stutterheim, Andermann, Carroll, Flecken, & Schmiedtová, 2012). While previous work on cross-linguistic differences in the conceptualization of motion has yielded mixed evidence, all studies to date have focused on a single type of event: voluntary motion, in which an entity travels along a trajectory by its own force (e.g., 'the man walked into the barn').

In this paper, we extend the scope of previous research by investigating a type of motion that hitherto has remained unexplored in the study of linguistic relativity: CAUSED MOTION, that is, the motion of an entity as a result of an external force exerted by another entity (e.g., 'The man pushed the tyre into the barn').[1] We ask to what extent our native language mirrors the way we conceptualize caused motion, which is arguably a conceptually more complex

---

[1] Throughout the paper we limit our discussion of earlier studies on motion events to those falling within Talmy's (2000b) framework. Thus, we will not be concerned with the sub-stantial line of work pursued by Wolff and colleagues on direct and indirect causation and the individuation of causation (Wolff, 2003; Wolff, Jeon, & Li, 2009), or the study of grammatical aspect and goal-oriented motion (Athanasopoulos & Bylund, 2013; Flecken, Gerwien, Carroll, & Stutterheim, 2015).

447

phenomenon than voluntary motion. Specifically, by implementing a novel methodological approach that combines the experimental principles used in previous motion cognition research, we ask which aspects of event representation differ and which ones are shared between speakers of typologically different languages. We focus on Spanish and Swedish, because these languages exhibit maximal contrast in their way of encoding motion, and thereby constitute an ideal test bed for assessing the relationship between linguistic structure and conceptualization.

2. Cross-linguistic differences in the linguistic packaging of motion

Caused motion involves a simple motion event and some external force that causes the simple event to happen (Talmy, 2000a), as in example (1):

(1)   The man pushed the tyre into the barn.

The simple event in (1) involves the non-volitional motion of a figure (the tyre) with respect to a ground (the barn) following a certain PATH (inwards). In addition, the situation described in (1) includes a causing event, namely an agent (the man) causing the tyre to move by manipulating it in a certain manner (he pushes it). Compared to voluntary motion (e.g., 'the man walked into the barn'), caused motion contains a larger number of event components that can be linguistically encoded. In (1), the tyre could move in different manners: it could roll or slide along the ground. In the present study, we manipulate the manner in which the agent *causes* an object to move, henceforth MANNER OF CAUSE, and the manner of motion of the *object*, henceforth MANNER OF OBJECT. Both are illustrated in (2).

(2)   a. He *pushed* the tyre into the barn. (MANNER OF CAUSE)
        b. He *rolled* the tyre into the barn. (MANNER OF OBJECT)

Both examples involve an agent moving a tyre into a barn, but they profile information about different event components: (2a) highlights the manner of cause (he *pushes* the tyre, rather than pulling it), while (2b) highlights manner of object (the tyre *rolls*, rather than sliding).

As in voluntary motion, we can apply Talmy's (1985, 2000b) well-known typological distinction between satellite-framed languages (S-languages) and verb-framed languages (V-languages) to caused motion. S-languages like Swedish or English encode path outside of the main verb root (e.g., in verb particles like *into*), and the verb typically expresses manner information (e.g., *push/roll*). In contrast, V-languages like Spanish encode path in the verb root (*entró* '(he) entered'). This is illustrated in examples (3) and (4), for Swedish and Spanish, respectively:

448

(3) Han sköt/rullade   in hjul-et   i   lada-n.
     He   pushed/rolled   in tyre-DEF   to barn-DEF
     MANNER-CAUSE/MANNER-OBJECT PATH
     'He pushed/rolled the tyre into the barn.'
(4) Entr-ó   en el granero   con   una rueda
     Enter-3SG   in the barn   with   a tyre
     PATH
     (empujándo-la/rodándo-la).
     pushing-it/rolling-it
     (MANNER-CAUSE/MANNER-OBJECT)
     'He entered the barn with a tyre (pushing it/rolling it).'

This contrasting lexicalization pattern makes manner information less codable in V-languages like Spanish, because there is no obligatory syntactic slot that encodes this information. Manner information *can* be expressed in Spanish, for example in a gerund as in (4), but it is often omitted without the description being ungrammatical or sounding odd. Swedish, on the other hand, represents an even more extreme case of S-language than English (Ragnarsdóttir & Strömqvist, 2004; see Cadierno & Ruiz, 2006, for Danish). For instance, Swedish lacks most of the English Latinate verbs encoding path, such as *ascend*, *descend*, *enter*, or *exit*. Moreover, Swedish has no generic motion verb (like the English *go* or the Spanish *ir*), which leads to a high use of various manner verbs in everyday discourse.

Presently, empirical investigations on cross-linguistic differences in the linguistic packaging of caused motion are scarce. In a pioneering study, Choi and Bowerman (1991) demonstrated that Korean lexicalizes caused and voluntary motion differentially, such that the former is encoded through a conflation of motion and path in the main verb (on a par with V-languages), whereas the latter is expressed with separate constituents for motion, path, and (optionally) manner.

More recently, Hickmann and colleagues investigated how child and adult native speakers of English and French (Hendriks, Hickmann, & Demagny, 2008; Hickmann & Hendriks, 2010), and English and Chinese (Ji, Hendriks, & Hickmann, 2011), express path, manner of cause, and manner of object. Interestingly, the two former studies found no cross-linguistic differences in the frequency of encoding of these motion components in English and French. The lack of cross-linguistic differences in the information conveyed may be due to the fact that French is not straightforwardly classified as a V-language, but has been argued to constitute a "mixed case" in Talmy's typology (Stringer, 2005, p. 210). However, French speakers were found to exhibit greater variation in their structural patterns than English speakers, encoding path, manner of cause, and manner of object interchangeably in

449

main verbs, subordinated verbs, or adverbial constructions. Ji et al. (2011), in contrast, found developmental differences in how speakers of English and Chinese (an equipollent framing language) encoded caused motion events. These results highlight the importance of choosing optimal language pairs, so as to avoid null effects that stem from a lack of clear-cut cross-linguistic differences in event descriptions.

3. Cross-linguistic differences in motion event cognition

Using Talmy's typology as a starting point, several studies have set out to investigate the potential effects of the linguistic packaging of voluntary motion on the way we conceptualize motion events. A prominent notion is Slobin's (1996, 2003) THINKING-FOR-SPEAKING HYPOTHESIS, which predicts that cross-linguistic differences in grammaticized and lexicalized concepts will lead speakers of different languages to attend to different features when preparing and producing speech. In one very clear demonstration of this, Papafragou et al. (2008) used an eye-tracking paradigm to probe whether speakers of English and Greek allocated visual attention to different components of an unfolding motion event (e.g., a man skating to a snowman on an ice rink). Unlike English, Greek is a V-language and typically encodes the path in the main verb. In the linguistic condition, participants had to inspect the events while preparing to describe them. Consistent with thinking-for-speaking, gaze allocation was distinct in the two groups and followed the specific lexicalization patterns of each language. However, this difference disappeared in the condition where participants freely inspected the events without having to describe them (see also Trueswell & Papafragou, 2010).

Moving beyond speech planning, researchers have probed the extent to which cross-linguistic differences in motion event encoding influence cognitive processes such as recognition memory, similarity judgements, and category learning. According to the SALIENCE HYPOTHESIS (cf. Papafragou, 2008), the repeated use of linguistically mediated conceptual categories, in this case the distinct lexicalization patterns of manner and path, will lead to non-linguistic conceptual schemata that are largely in agreement with linguistic patterns. In a series of groundbreaking studies, Gennari et al. (2002) and Papafragou and associates (Papafragou et al., 2002; Papafragou & Selimis, 2010) tested this hypothesis using a triads-matching paradigm with native speakers of English, Greek, and Spanish. In this paradigm, each experimental stimulus consists of a reference event (e.g., a man walking up the stairs), a same-path alternate in which the manner is changed (a man *running* up the stairs), and a same-manner alternate in which the path is changed (a man walking *down* the stairs). Gennari et al. (2002)

450

found that Spanish speakers had a greater preference for same-path alternates compared to English speakers, but only if they had described the events before judging their similarity (a similar finding was also reported by Lai, Garrido Rodriguez, & Narasimhan, 2014). A possible flaw in Gennari et al.'s study is that the stimuli do not seem to have consistently discriminated between voluntary and caused motion (cf. Gennari et al., 2002, pp. 66–67). Papafragou and Selimis (2010) similarly found that when the task instructions contained a linguistic bias, English speakers were more likely to base their similarity judgements on manner than Greek speakers. In these studies cross-linguistic differences disappeared in the absence of biasing instructions (Papafragou & Selimis, 2010) or prior verbalization (Gennari et al., 2002), as well as in the presence of a verbal interference task (Gennari et al., 2002).

The studies just reviewed all used a task in which participants had to make a forced choice between path and manner. By its very design, this task confounds path and manner preferences, because proportions of path and manner choices always add up to one. Thus, higher path preference becomes equivalent to lower manner preference and vice versa. However, both S-languages and V-languages consistently encode path information, albeit mapped onto different linguistic elements. The crucial difference lies in the fact that manner is less codable in V-languages and thus gets more often omitted in discourse. Therefore, a design that teases out path bias from manner bias more adequately reflects the linguistic state of affairs, which after all is what is hypothesized to drive any potential conceptual differences between language groups.

Following this logic, Kersten, Meissner, Lechuga, Schwartz, Albrechtsen, and Iglesias (2010) adopted a methodological approach which did not oppose manner and path. They implemented a supervised classification paradigm in which English and Spanish speakers had to learn to correctly identify alien species depending on their motion patterns. The diagnostic criterion for classification differed between subjects; it was either path or manner. A series of experiments consistently showed that English speakers had an advantage over Spanish speakers when the relevant classification criterion was manner, but not when it was path, in which case both groups performed equally well. One limitation of this study is that it did not manipulate the involvement of language, thus leaving open the question of how much participants were relying on verbal mediation to solve the task.

The evidence reviewed above provides mixed support for the salience hypothesis. If it were the case that conceptual categories of motion were invariably aligned with linguistic categories, Whorfian effects should not fluctuate as a function of experimental condition and task. On the one hand,

the negative evidence lends support to the idea that motion cognition is mainly guided by our shared perceptual system with language exerting no influence on how we perceive motion events (Papafragou, 2008). On the other hand, the available positive evidence provides a qualified picture of language effects on motion event cognition. If language-specific labels are made salient in an experimental situation, they may modulate event categorization criteria. Linguistic labelling might be achieved by means of using instructions that encourage the use of language (Papafragou & Selimis, 2010), or by implementing a challenging task that promotes the use of labels (Kersten et al., 2010). These effects are consistent with what Wolff and Holmes (2011) have termed THINKING-WITH-LANGUAGE effects, whereby linguistic and non-linguistic processes might be activated in tandem, with the consequence that linguistic categories mediate similarity judgements and facilitate category learning (for similar findings on grammatical aspect and goal-oriented motion, see Athanasopoulos & Bylund, 2013; Bylund & Athanasopoulos, 2014; Flecken et al., 2014).

4. The present study

The overall aim of the present paper is to extend the cross-linguistic study of motion event cognition to the domain of caused motion. To this end, we implemented a free arrangement task (Goldstone, 1994), in which participants had to arrange different scenes of caused motion on the basis of their perceived similarity. Unlike the commonly used forced triads-matching task, this paradigm does not confound path and manner preference, but instead captures the possibility that the relative importance of event components be weighted differently. The arrangement task was carried out under three different encoding conditions that manipulated the participants' likelihood to use language to solve the task. Our method thus represents an improvement compared to previous studies, which either have manipulated linguistic task mediation but confounded path/manner preference (Gennari et al., 2002; Papafragou et al., 2002; Papafragou & Selimis, 2010), or vice versa (Kersten et al., 2010). The experimental approach, combined with the choice of prototypical S- and V-languages (Swedish and Spanish, respectively), thus maximizes the likelihood of detecting potential cross-linguistic differences.

Our first and narrower research question concerns whether speakers of Spanish and Swedish differ in their focus on manner of cause and manner of object when judging event similarity, reflecting how these events are described in their respective languages. Our second and broader question concerns the commonalities and differences in how speakers of different languages represent motion, and whether event representation varies as

452

a function of verbal mediation. Experiments 1 through 3 speak to the first question; we address the second question in a subsequent compound analysis of all three experiments.[2]

5. Experiment 1: similarity assessment under linguistic encoding

Under the assumption that cross-linguistic variability in describing caused motion implies corresponding differences in conceptualizing these events, such a pattern should be strongest when language-specific representations have been explicitly activated. For this reason, Experiment 1 tested how speakers of Spanish and Swedish judged event similarity after first having described the events in their respective languages.

5.1. METHOD

5.1.1. *Participants*

Twenty-four native Spanish speakers (mean age = 23.4 years, SD = 3.0) and twenty-two native Swedish speakers (mean age = 25.7 years, SD = 4.3) participated. Spanish speakers were students at the Complutense University of Madrid, and Swedish speakers were students at Stockholm University. Participants used their native language routinely since they lived in a monolingual context. Some had familiarity with other languages, but none of them had expert knowledge. Crucially, the Spanish speakers had no knowledge of Swedish, and vice versa. Participation was remunerated.

5.1.2. *Materials*

The stimuli were thirty-two animated cartoons (each approximately 7 seconds long) depicting caused motion events in which the same human-like agent displaced an object along a certain path. The animations were originally developed by Hickmann and associates (e.g., Hickmann & Hendriks, 2010). Three motion components were systematically crossed in the stimuli: path (four levels: up, down, across, into); manner of cause (two levels: push, pull); and manner of object (two levels: roll, slide). See Fig. 1 for examples. The sagittal direction of motion (left-to-right, right-to-left) was counterbalanced, so that each of the sixteen possible combinations of motion components

---

[2] All data and analyses reported in this paper are publicly accessible via the Harvard Dataverse at: <http://dx.doi.org/10.7910/DVN/JZZIYU>.

453

Fig. 1. Three example events to illustrate the design. The first and second scenes share the manner of object (both objects *roll*); the first and third scenes share the path (motion *downwards*); the second and third scenes share the manner of cause (the agent is *pulling* the object). Our analysis measures how much participants relied on each of these components when carrying out similarity arrangements. Note that the real items were videos showing actual motion.

corresponded to two target items, one for each direction. In addition, stimuli varied with respect to the ground in which the event took place (8 different grounds) and the object that was moved (16 different objects). This variation was necessary: for the path to vary (e.g., 'into' versus 'down') the ground necessarily has to vary as well; similarly, an object can either roll or slide, but typically cannot do both. A full description of the target items is given in the Supplementary Materials (available at <http://dx.doi.org/10.1017/langcog.2016.22>.

5.1.3. *Norming study*

We took the measure of establishing the degree to which speakers of Spanish and Swedish actually differed in their descriptions of caused motion as depicted in the videos. Eighteen native Spanish and nineteen native Swedish speakers, who did not take part in the main experiment but belonged to the same two participant populations, provided descriptions of the thirty-two stimulus videos and seven additional filler items presented in pseudo-randomized order. The instructions were: "Describe what happens in each scene after having watched it in its entirety." Participants were not given any limitation regarding length. Descriptions were coded for whether they mentioned path (e.g., 'into', 'across'), manner of cause (e.g., 'push', 'pull'), and manner of object (e.g., 'roll', 'slide'), and whether this information was expressed in a main verb root or outside of it (see TABLE 1). As expected, descriptions in both languages were very likely to express path (Spanish: 95% of all descriptions, Swedish: 99%), but Swedish more often than Spanish descriptions mentioned manner of cause (Spanish: 55%, Swedish: 78%) and manner of object (Spanish: 10%, Swedish: 29%). Lexicalization patterns followed each language's typological status, with Spanish verbs mainly expressing path and Swedish verbs manner (TABLE 1). Comparison

454

TABLE 1. *Percentage of descriptions in norming study that contained each of the three semantic components and whether they were lexicalized in the main verb root (V) or outside of it (Outside V). Percentages are computed out of 576 Spanish descriptions and 608 Swedish descriptions*

Component	Spanish (%)			Swedish (%)		
	V	Outside V	Total	V	Outside V	Total
Path	68.2	32.3	94.6	0.5	98.5	99.0
Manner of cause	26.2	29.3	54.5	77.1	9.9	77.6
Manner of object	3.8	5.9	9.7	28.9	0.0	28.9

NOTES: The first two columns by language do not always add up to the total proportions (Total) because the same description can have redundant information in the main verb root and outside of it (e.g., *sube para arriba* 'he ascends up'). The V columns need not add up to 100 because verbs can express neither of the components (e.g., Spanish *va* 'he goes') and because there can be two main verbs in a description (e.g., Swedish *skjuter och rullar* 'pushes and rolls').

of these figures to previous studies confirms the suitability of the current language pair.[3]

We fitted three separate logistic mixed models (Jaeger, 2008) to assess whether Spanish and Swedish reliably differed in their likelihood to express each of the three components in TABLE 1. In each model the binary dependent variable was whether the component was expressed in a description or not. The sole predictor was language, which was dummy coded (Spanish = 0, Swedish = 1), so that the intercept in each model expresses the log-likelihood that Spanish speakers encoded a given component, while the language coefficient expresses the difference in log-likelihood between languages. Here we only report the latter, which provide the critical comparison (see Supplementary Materials for full model details including random effects structure). Results confirmed that Spanish and Swedish speakers were equally likely to express path ($\hat{\beta}_{\text{Swedish-vs.-Spanish}}$ = 3.03, p = .113), but that Swedish speakers were more likely than Spanish speakers to encode manner of cause ($\hat{\beta}_{\text{Swedish-vs.-Spanish}}$ = 5.01, p < .001) and manner of object ($\hat{\beta}_{\text{Swedish-vs.-Spanish}}$ = 2.55, p = .018). The norming study thus confirms the validity of our stimuli for the two tested language populations.

[3] The figures for Swedish confirm that it is a more prototypical example of a S-language than English. Collapsing manner of cause and manner of object into a single manner component for comparability with previous research on voluntary motion, Swedish speakers encoded manner in 99% of their descriptions. In comparison, Gennari et al. (2002, pp. 65–67) report 86% of English descriptions mentioning manner; Papafragou et al. (2008, p. 167, Table 1) and Papafragou and Selimis (2010, p. 238, Table 3) both report 90% of English descriptions mentioning manner. As for Spanish descriptions, 63% encoded manner in the present study, which is less than the 71% reported by Gennari et al. (2002); hence, there is no support that our stimuli elicited more manner mentions overall.

5.1.4. *Procedure*

Participants were tested individually by a native speaker of the relevant language. They first described the stimuli following the same procedure as in the norming study, and then moved on to the arrangement task. The exact instructions, translated into English, were: "Your task is to arrange the scenes on the screen depending on how similar they are. Video-clips showing similar actions should be placed near each other; if the action is different, the scenes should be placed far away from each other." Participants were informed they would carry out three series of arrangements and that they would not be able to move a video-clip once it had been placed. The test phase, which started after a brief training phase, consisted of three arrangement blocks of twenty-two video-clips each. The progression per block was as follows: first, a video-clip was played on the screen in its entirety. The participant then moved to a screen where they had to place the scene by clicking with the mouse; upon clicking, a still of the video-clip appeared on the screen. This procedure was repeated until the end of the block. Participants could arrange items freely on the screen (they were not constrained to form piles or clusters). They moved forward by clicking on a centred message box to prevent spatial bias in the arrangements.

The specific items appearing in each block were randomized subject to the following constraints: no item appeared more than once in the same block; all 496 possible pairs of video-clips appeared at least once in some block; the number of videos appearing three times across blocks was minimized. We preferred this algorithm to the one used in Goldstone (1994), that is, choosing the items completely at random from the full set, since fully randomizing the choice of items would entail missing observations and lead to an unbalanced dataset. This resulted in a total of 693 pairwise similarity values per participant (231 for each block). The arrangement task was programmed in E-Basic and run in E-Prime (Schneider, Eschman, & Zuccolotto, 2002).

5.1.5. *Design and analysis*

To assess whether speakers of Spanish and Swedish relied on different event components in their arrangements, we analyzed to what extent two events sharing a component led to increased similarity scores (Fig. 1). For example, if two events shared the manner of object (e.g., in both events the object was rolling), did this on average lead to an increase in similarity ratings with respect to events that did not share the manner of object? Crucially, was this increase different for Spanish and Swedish speakers?

The dependent variable was similarity between event pairs. Similarity was a continuous measure bounded between 0 (minimal similarity) and 1 (maximal similarity). It was computed from the similarity arrangements in two steps.

456

First, we normalized the distance in pixels between all pairs of scenes for each participant–block combination to yield a measure between 0 and 1, as follows:

$$Sim_{ijbs} = 1 - \frac{Dist_{ijbs}}{MaxDist_{bs}}, \qquad (1)$$

where Sim_{ijbs} denotes the normalized similarity score between scenes i and j in block b for subject s. $Dist_{ijbs}$ denotes the distance in pixels between the coordinates of the centres of scenes i and j in arrangement b for subject s, and $MaxDist_{bs}$ denotes the maximal distance between two scenes in block b for subject s. Second, for each participant we averaged the normalized similarity between pairs of scenes across blocks, so as to yield one single measure per participant and event pair.[4] The final number of observations per participant was 496, one for each pairwise combination of the stimuli.

We examined the effect of the critical event components (path, manner of cause, and manner of object) on similarity scores and their interactions with language using linear mixed effects regression models (Baayen, Davidson, & Bates, 2008). The analyses controlled for the effect of other non-target variables (ground, object, and left/right direction). Predictors were first centred so that the reported coefficients represent the estimated difference in similarity ratings between the two levels of each predictor. We included the maximal by-subject random effects structure and by-item random intercepts (Barr, Levy, Scheepers, & Tily, 2013). All analyses were run in in R (R Development Core Team, 2013) using the *lmer* function of the *lme4* library (Bates, Maechler, Bolker, & Walker, 2014). See Supplementary Materials for full specification of the linear mixed effects models.

5.2. RESULTS AND DISCUSSION

5.2.1. *Reliance on event components*

Results by language are plotted in Fig. 2A. There were main effects of path (Path$_{same-vs.-different}$: $\hat{\beta} = 0.18$, $t = 8.43$, $p < .001$), manner of cause (MannerCause$_{same-vs.-different}$: $\hat{\beta} = 0.08$, $t = 4.74$, $p < .001$) and manner of object (MannerObject$_{same-vs.-different}$: $\hat{\beta} = 0.04$, $t = 3.88$, $p < .001$), indicating that all of these components were used to judge event similarity. We were especially interested in interactions of these components with language to assess cross-linguistic differences in similarity judgements (Fig. 2A). The analysis revealed both cross-linguistic similarities and differences. There was no difference in how much speakers of each language relied on path

---

[4] No other similarity measures were explored, to avoid inflation of Type I errors.
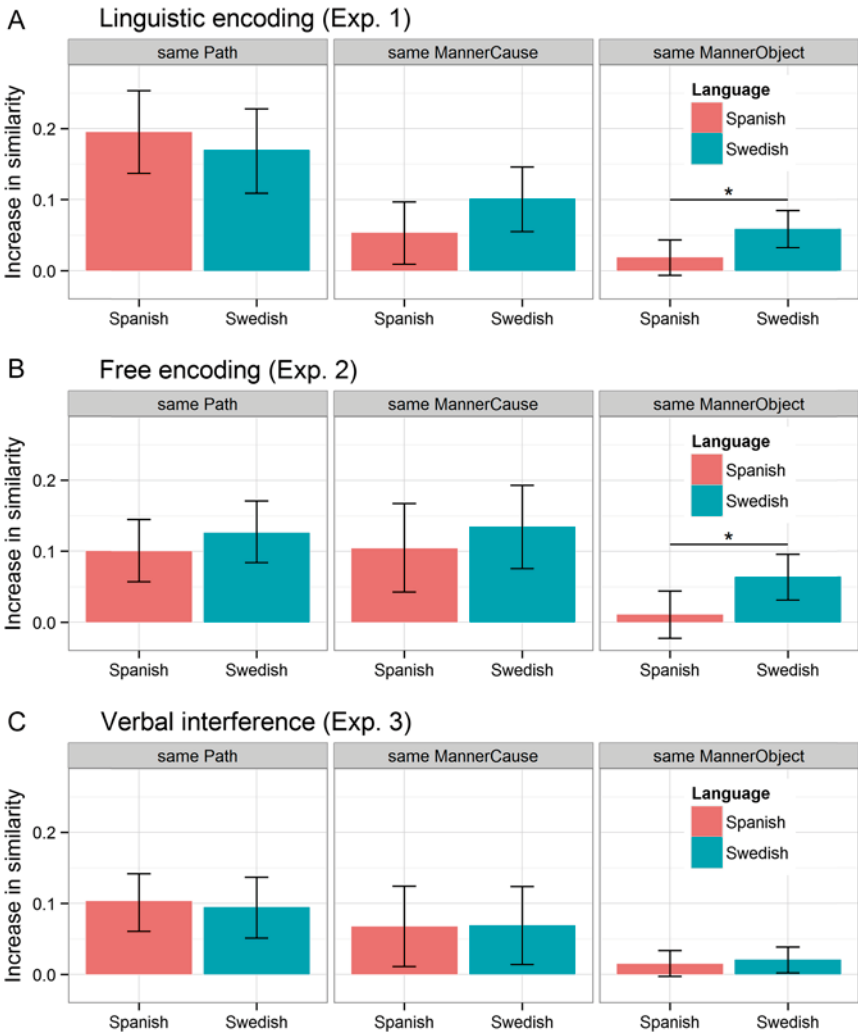
457

Fig. 2. Results from Experiments 1 through 3. Each figure shows the increase in pairwise event similarity due to shared event components (path in left panel; manner of cause in middle panel; manner of object in right panel), as a function of language group. The *y*-axis indicates how much more similar an event was perceived to be when it shared a component (e.g., same path) as opposed to when it did not share it (different path). Bars show by-language means estimated from our analysis. Error bars indicate 95% confidence intervals obtained from posterior simulations via *sim* function in R package arm (Gelman & Su, 2014). Significant differences are marked with asterisks.

458

(Path$_{\text{same-vs.-different}}$ × Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = -0.02$, $t = -0.58$, $p > .10$). Swedish speakers did, however, rely significantly more on manner of object than Spanish speakers (MannerObject$_{\text{same-vs.-different}}$ × Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.04$, $t = 2.28$, $p < .05$). Finally, with respect to manner of cause, while Swedish speakers did numerically rely more on this component than Spanish speakers, the difference was not significant (MannerCause$_{\text{same-vs.-different}}$ × Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.05$, $t = 1.54$, $p > .10$).

These results provide several insights. First, they show that participants in the current task and with our specific instructions were not just adopting the very simple strategy of labelling actions with verbs to gauge event similarity. In that case we should have observed that Spanish speakers relied almost exclusively on path (based on Spanish path verbs like *subir* 'ascend', *entrar* 'enter', etc.) and Swedish speakers on manner (based on Swedish manner verbs like *skjuta* 'push', *rulla* 'roll', etc.). Instead, participants seemed to engage in a complex decision-making procedure in which several components were used to assess event similarity.[5] Second, the lack of a difference in reliance on path further confirms that this component is equally salient for speakers of both languages (recall that path was systematically encoded in both languages in the norming study; see also Kersten et al., 2010). Finally, one of the manner manipulations yielded cross-linguistic differences: Swedish speakers relied more than Spanish speakers on the manner of object motion (rolling or sliding); but no significant difference was found with respect to manner of cause (pushing or pulling). This suggests that speakers are doing more than just basing event representations on the typical descriptions in their language.

What the between-group analysis leaves open, however, is whether there is a tight link between how individual speakers describe the events and then judge their similarity. A strong link would be indicative that participants were using their descriptions to perform the task. In other words, it would support a thinking-for-speaking effect (as in Gennari et al., 2002). In contrast, the absence of such a link would suggest that the source of the influence was not coming from the linguistic descriptions per se, but was possibly an effect of habitual linguistic encoding on attention, as predicted by the salience hypothesis (cf. Kersten et al., 2010). We explore this in the next analysis.

5.2.2. *Are descriptions predictive of similarity judgements?*
To test the relation between descriptions and similarity judgements, we performed by-participant correlation analyses between the frequency with which an event component was mentioned in the descriptions during

---

[5] We are grateful to an anonymous reviewer for pointing this out.

459

the linguistic encoding phase and the degree to which that component was used in the subsequent similarity arrangement task.[6] If during the arrangement task speakers are mainly retrieving previously activated linguistic representations, the extent to which they mentioned an event component should offer a good index of how much they relied on that component for judging similarity.

Fig. 3 shows scatterplots with regression lines and 95% confidence intervals for each of the three semantic components in Spanish and Swedish. The large confidence intervals for the regression lines suggest that overall the link was not very strong. First, path showed very limited spread over the *x*-axis because speakers tended to systematically include this component in their descriptions, which makes inferences about the relation between descriptions and similarity judgements difficult (Fig. 3, left panel). Manner of cause and manner of object (middle and right panels) show a greater spread over the *x*-axis, but here again the confidence intervals are broad. Note that in the case of manner of cause in Swedish the correlation numerically even runs counter to expectations, with *more* mentions of manner of cause predicting *less* reliance on this component during the similarity task (dashed line in middle panel).

Results from the six separate by-speaker correlation analyses for each component and language are shown in TABLE 2. Only one of the correlations reached significance at the .05 level, namely manner of object in Spanish. These results suggest at best a very weak link between participant descriptions and performance on the similarity task. The low correlations between subject descriptions and their similarity arrangements are suggestive of a thinking-with-language effect and thus open up for the possibility that there may be cross-linguistic differences even in the absence of prior linguistic encoding. We tested this in the next experiment.

6. Experiment 2: similarity assessment under free encoding

Do people pay attention to aspects of the events typically encoded in their language, even when linguistic representations are not explicitly evoked as in Experiment 1? To test this, we let speakers of Spanish and Swedish carry out the same similarity arrangement task without providing prior event descriptions (free encoding condition). If it was the descriptions that drove the patterns in Experiment 1, we should find no differences between language groups in the current experiment. A lack of differences under free encoding would support a thinking-for-speaking type of effect and would replicate the findings in Gennari et al. (2002). If, on the other hand, the same

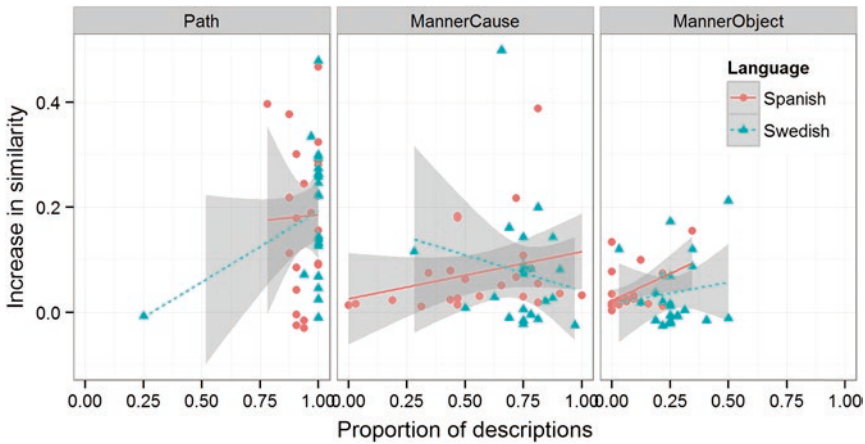[6] We are grateful to an anonymous reviewer for suggesting this analysis.

Fig. 3. By-subject scatterplots showing the relation between verbally encoding a semantic component and using it for similarity judgements. Shapes represent participants (Spanish: dots; Swedish: triangles). The *x*-axis in each panel indicates the proportion of descriptions that contained a given component (path, manner of cause, manner of object). The *y*-axis indicates the by-subject effect size of that component on similarity judgements, as extracted from the mixed-effects model. Each participant contributes one point per panel. Linear regression lines by language are shown with 95% confidence intervals.

TABLE 2. *Pearson's* r *and associated* p-*value between by-speaker proportion of descriptions containing a given semantic component and that speaker's reliance on the component in the similarity task (as estimated from the mixed model). The only significant correlation is marked in* **boldface**.

Component	Spanish		Swedish	
	r	*p*	*r*	*p*
Path	0.02	.93	0.35	.11
Manner of cause	0.27	.20	−0.17	.45
Manner of object	**0.45**	**.03**	0.13	.56

cross-linguistic patterns persist (i.e., Swedish speakers still rely more on manner-related components than Spanish speakers), this would suggest a deeper relation between the language we speak and the concepts we form, one that does not depend on explicit verbal mediation.

6.1. METHOD

6.1.1. *Participants*

Twenty-four native Spanish speakers (mean age = 22.3 years, SD = 2.8) and twenty-five native Swedish speakers (mean age = 25.8 years, SD = 4.0)

461

participated. Participants in each group were drawn from the same student populations as in Experiment 1. Participation was remunerated.

6.1.2. *Procedure*

The procedure was identical to that in Experiment 1 except for the familiarization phase preceding the similarity task: instead of describing the events, participants went through a familiarization phase where they silently watched the target stimuli presented in random order.

6.1.3. *Materials, design, and analysis*

The materials, design, and analysis were all the same as in Experiment 1.

6.2. RESULTS AND DISCUSSION

The results are plotted in Fig. 2B. As in Experiment 1, we found main effects of path (Path$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.11$, $t = 6.66$, $p < .001$), manner of cause (MannerCause$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.12$, $t = 5.45$, $p < .001$), and manner of object (MannerObject$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.04$, $t = 3.11$, $p < .01$), again indicating that all of these components were used to judge event similarity. Critically, the interactions of these components with language were qualitatively identical to Experiment 1. Once again, Swedish speakers relied significantly more on manner of object than Spanish speakers (MannerObject$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.05$, $t = 2.27$, $p < .05$). As expected, there was no difference in how much speakers of each language relied on path (Path$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.03$, $t = 0.87$, $p > .10$). Finally, while the numerical difference with respect to manner of cause persisted (greater mean reliance by Swedish speakers), it was far from significant (MannerCause$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.03$, $t = 0.70$, $p > .10$).

In Experiment 2 participants did not describe the events prior to carrying out the similarity arrangement task. Yet precisely the same qualitative patterns obtained as under linguistic encoding (Experiment 1). This outcome, together with the relative lack of correlations between speaker descriptions and their similarity judgements in Experiment 1, is suggestive of a general tendency for Swedish speakers to pay more attention to the manner of object than Spanish speakers, even when linguistic representations are not explicitly invoked, congruent with thinking-with-language and the saliency hypothesis. However, the results leave open the possibility that speakers were covertly using language during the task. That is, to help them carry out the arrangement task, participants might have described the events subvocally. If so, the fact that

462

the same cross-linguistic difference was found in Experiments 1 and 2 would have a simple explanation: participants might have been using language *to the same extent* in both experiments, only overtly in the first and covertly in the second. We tease out these different possibilities in two steps. First, we present the results of a third experiment in which the ability to subvocally describe the events was reduced (Experiment 3). Second, we run a compound analysis of the experiments to find out whether there is evidence that participants are solving the similarity task in equivalent ways irrespective of encoding condition.

7. Experiment 3: similarity assessment under verbal interference

Do the results in Experiments 1 and 2 reflect a deep cognitive bias that leads Swedish speakers to pay more attention than Spanish speakers to manner-related event components, even when the use of language is blocked? We tested this by letting speakers of both languages carry out the similarity arrangement task under verbal interference throughout both the encoding and the test phase. If the same cross-linguistic difference persists, this will constitute evidence that the effect is not mediated by the on-line recruitment of language. If the effect disappears, it will suggest that the difference was due to the on-line activation of linguistically mediated event representations.

7.1. METHOD

7.1.1. *Participants*

Twenty native Spanish speakers (mean age = 20.7 years, SD = 1.4) and twenty native Swedish speakers (mean age = 24.3 years, SD = 4.1) participated. In each group, participants were drawn from the same student populations as in Experiments 1 and 2. Participation was remunerated.

7.1.2. *Procedure*

The procedure was identical to that in Experiment 2, except that participants repeated out loud random series of three two-digit numbers throughout the experiment (cf. Trueswell & Papafragou, 2010). A new series was presented aurally before each trial during both the familiarization and the arrangement phase.

7.1.3. *Materials, design, and analysis*

The materials, design, and analysis were all the same as in Experiments 1 and 2.

463

7.2. RESULTS AND DISCUSSION

Results for the interference condition are plotted in Fig. **2**C. There were main effects of path (Path$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.10$, $t = 6.33$, $p < .001$), manner of cause (MannerCause$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.07$, $t = 3.46$, $p < .001$), and manner of object (MannerObject$_{\text{same-vs.-different}}$: $\hat{\beta} = 0.02$, $t = 2.46$, $p < .05$), so all of these components were used to some extent to judge event similarity. Critically, none of these components interacted significantly with language: path (Path$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = -0.01$, $t = -0.32$, $p > .10$), manner of cause (MannerCause$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.00$, $t = -0.01$, $p > .10$), manner of object (MannerObject$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.00$, $t = 0.41$, $p > .10$).[7]

When participants' ability to engage in linguistic mediation was reduced, the higher reliance on manner of object by Swedish speakers disappeared. The vanishing of cross-linguistic differences under verbal interference lends no support to the strong version of the salience hypothesis that language can change our underlying perceptual machinery. The result also replicates previous findings under verbal interference (Gennari et al., 2002; Trueswell & Papafragou, 2010). Regarding the effects of Experiments 1 and 2, however, they are still consistent with two accounts. It could be the case either that Experiments 1 and 2 yielded the same results because linguistic mediation was equally active in both experiments, or that the difference persisted although language was not being used to the same extent or in the same fashion. The final compound analysis speaks to this question.

8. Compound analysis of Experiments 1–3

Experiments 1 to 3 had their main focus on cross-linguistic differences. This final section adopts a broader perspective. On the one hand, we aim to explore the weights given to the different event components, including the control variables for which we did not have explicit cross-linguistic predictions (direction, ground, and object); this will help understand how participants solved the kind of complex similarity task we gave them. On the other hand, a compound analysis lets us directly assess whether the weights assigned to

---

[7] In the model for Experiment 3, there was one higher-order interaction with language which was significant at the .05 level, namely the interaction of path, manner of cause, and language (Path$_{\text{same-vs.-different}} \times$ MannerCause$_{\text{same-vs.-different}} \times$ Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta} = 0.05$, $t = 2.51$, $p < .05$). We merely report this result here but refrain from interpreting it as a meaningful cross-linguistic difference, for two reasons. First since this higher-order interaction was not predicted, and since the model included many parameters, there is an increased risk of Type I error (Gelman & Loken, n.d.). Second, the interpretation of this effect is not straightforward. Speculatively, it might reflect a higher propensity by Swedish speakers to cluster same-path and same-manner events.

464

each component varied as a function of encoding condition; this will be informative as to the nature of event representations that participants create depending on whether language is engaged or not.

8.1. ANALYSIS

For this analysis we fitted a very similar model to those in Experiments 1 to 3, except that now we used the data from all three experiments and added the factor ENCODING CONDITION, with three levels: linguistic, free, and interference (corresponding to Experiments 1 to 3, respectively). Encoding condition was forward coded, so that we can compare linguistic encoding against free encoding (Experiments 1 vs. 2), and free encoding against verbal interference (Experiments 2 vs. 3). As in experiments 1–3, all other predictors were centred with a difference of 1, so that the reported coefficients represent the estimated difference in similarity ratings between the two levels of each predictor. See Supplementary Materials for full specification and output of the model.

8.2. RESULTS AND DISCUSSION

Results by encoding condition are shown in Fig. 4. When considering all three experiments together, the analysis again found main effects of path (Path$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.13, t = 11.01, p < .001), manner of cause (MannerCause$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.09, t = 7.75, p < .001), and manner of object (MannerObject$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.03, t = 4.71, p < .001). There were also main effects of the control variables: left/right direction (Direction$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.10, t = 7.40, p < .001), ground (Ground$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.03, t = 2.77, p < .01), and object (Object$_{\text{same-vs.-different}}$: $\hat{\beta}$ = 0.05, t = 2.44, p < .05), indicating that participants also used these components in their similarity judgements. Crucially, we found two significant interactions of event components with encoding condition: path was used *more* under linguistic encoding (Experiment 1) than free encoding (Experiment 2) (Path$_{\text{same-vs.-different}}$ × Encoding$_{\text{linguistic-vs.-free}}$: $\hat{\beta}$ = 0.07, t = 2.97, p < .01), while the opposite was true for left/right direction (Direction$_{\text{same-vs.-different}}$ × Encoding$_{\text{linguistic-vs.-free}}$: $\hat{\beta}$ = −0.10, t = −3.42, p < .001). No other event component interacted with encoding condition. As for interactions of event components with language, the analysis again yielded a greater reliance by Swedish speakers on manner of object (MannerObject$_{\text{same-vs.-different}}$ × Language$_{\text{Swedish-vs.-Spanish}}$: $\hat{\beta}$ = 0.03, t = 3.02, p < .01), an effect we know was driven by Experiments 1 and 2. Additionally, however, we found a reliable difference with respect to ground, indicating that Spanish speakers paid more attention than Swedish speakers to the particular landscapes shown in the stimuli (Ground$_{\text{same-vs.-different}}$ × Language$_{\text{Swedish-vs.-Spanish}}$:
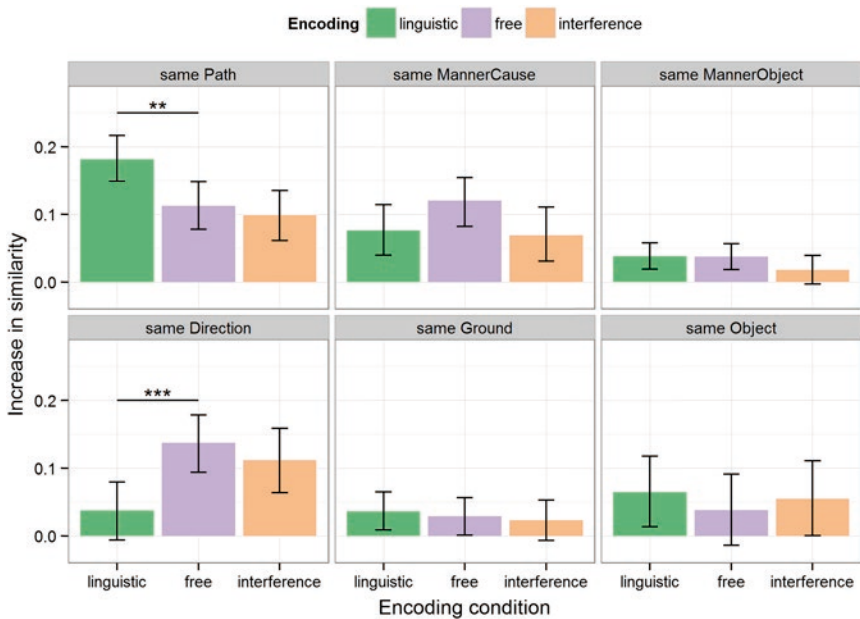
465

Fig. 4. Results from compound analysis of Experiments 1–3. Each panel shows the increase in pairwise event similarity due to sharing each of the critical event components (path, manner of cause, and manner of object) or control components (direction, ground, and object). Results are shown as a function of encoding condition (linguistic, free, verbal interference). The y-axis indicates how much more similar an event was perceived to be when it shared a component (e.g., same path) as opposed to when it did not share it (different path). Bars show mean effects as estimated from our analysis. Error bars indicate 95% confidence intervals obtained from posterior simulations via *sim* function in R package arm (Gelman & Su, 2014). Significant differences are marked with asterisks.

$\hat{\beta} = -0.04, t = 2.71, p < .01$). Finally, there were no significant three-way interactions between either of the event components, language, and encoding condition, suggesting that encoding condition had the same effect across languages.[8]

The compound analysis showed that numerically the most predictive component of event similarity across experiments was path. It was closely

---

[8] As in Experiment 3, there were higher-order interactions between two event components and encoding condition and/or with language (e.g., Path × MannerCause × Language). We follow the same principle as in Experiment 3 and refrain from interpreting these interactions here (see previous note). There is an additional reason to be cautious about these higher-order interactions in the compound analysis: the fitted mixed model did not include by-subject random effects for interactions between event components, because such a model failed to converge (see Supplementary Materials). Thus, the model estimates of the coefficients of these higher-order interactions are likely to be anticonservative with misleadingly low *p*-values.

466

followed by manner of cause and direction, with only slightly lower mean effects. Manner of object, ground, and object had an overall lower, but significant, effect on similarity judgements. However, event components were not equally important across conditions: the effect of path was stronger if events had been previously described, whereas the effect of left/right direction was smaller under this condition. This implies that participants were not simply adopting the same strategy in the linguistic and the free encoding conditions: participants in free encoding do not seem to have silently (i.e., subvocally) formulated full-blown linguistic descriptions of the events. The observed decrease in the importance of path and the concomitant increase in the importance of direction suggest that verbal mediation was stronger in the linguistic condition, boosting the importance of a component that is typically verbalized (path) and supressing the effect of a visually salient but linguistically irrelevant category (left/right direction). In the free condition, linguistically relevant and non-relevant categories seemed to be activated in tandem.

Last, the finding that Spanish speakers relied more on the ground (i.e., landscape) than Swedish speakers was not predicted on the basis of the linguistic descriptions in the norming study (ground information was systematically and equally often encoded in both languages, see Supplementary Materials). Slobin (1996) has shown that the rhetorical style in Spanish is particularly prone to encoding static locative descriptions (i.e., grounds), but adducing this as an explanation remains post-hoc. Given that the difference only arose in the compound analysis, but not in the individual experiments, we remain cautious about this unexpected finding.

9. General discussion

This study has extended research on motion cognition to the domain of caused motion. We used a paradigm that let participants choose among several event components when judging the similarity of dynamic scenes, and we explored the role of language in those choices. Our first and narrow question was about cross-linguistic differences: Do the components used by Spanish and Swedish speakers mirror the lexicalization patterns of their languages? We found the answer to be a partial 'yes'. Since manner information is expressed more often in Swedish than Spanish (see norming study), Whorfian accounts predict a greater focus by Swedish speakers on this component when mentally representing events. We found this to be the case for only one of the two manner manipulations in our stimuli: Swedish speakers tended to rely more than Spanish speakers on manner of object (whether the object rolled or slid), but both groups equally relied on manner of cause (whether the agent pushed or pulled the object). The difference in manner of object was found when participants judged event similarity after having described the events

(Experiment 1: linguistic encoding) or after having watched them silently (Experiment 2: free encoding). However, the difference disappeared under verbal interference (Experiment 3), lending no support to the view that language changes our underlying perceptual machinery. This seems to correspond to thinking-with-language, the hallmark of which is precisely that "it can be eliminated by having people engage in a verbal interference task" (Wolff & Holmes, 2011, p. 256).

Why not simply call it thinking-for-speaking? Recall that in the free encoding condition (Experiment 2) participants were not asked to produce verbal descriptions at any point, which is the typical scenario for thinking-for-speaking, as formulated by Slobin (2003). They did have the possibility of using language, and it is even quite likely that they did so, as verbal encoding could offer a way of keeping the events in memory during a rather long and complex task.[9] However, we know from the compound analysis that the linguistic and free encoding conditions did differ in ways suggestive that there was *more* verbal mediation in the former than in the latter. This, coupled with the weak correlations between descriptions and similarity judgements in Experiment 1, suggests that the effect is one of thinking-with-language, which in turn is very similar to what has been referred to as language-as-strategy (Gennari et al., 2002). Overall, then, our results from the domain of caused motion replicate previous cross-linguistic differences found in voluntary motion (Gennari et al., 2002; Kersten et al., 2010; Lai et al., 2014; Papafragou & Selimis, 2010). We conclude that – independent of the degree of complexity of the events, which is greater for caused motion than for voluntary motion – language-specific labels may modulate event conceptualization when made salient in a given task.

A question that remains is the lack of cross-linguistic differences for manner of cause in the arrangements, even though this component was encoded significantly more often in Swedish than in Spanish event descriptions. This result poses a challenge to any simple Whorfian account and calls for a more complex model of the correspondence between linguistic and conceptual patterns. One possibility is that manner of cause has higher cognitive salience than manner of object, which would override any language-induced difference. This may be because pushing or pulling an object pertains to the agent and is thus linked to a highly prominent role in the event (i.e., the protagonist). Alternatively, pushing and pulling can be reduced to the relative position of the object with respect to the agent (in front or behind), a property that was arguably more visually salient in our stimuli than the way the object moved. Analyzing linguistic data, Malt et al. (2014) recently provided evidence that

---

[9] We are grateful to an anonymous reviewer for pointing this out.

Spanish speakers make less fine-grained manner distinctions than speakers of S-languages, but that speakers of both language types converge for more general manner categories (see also Slobin, Ibarretxe-Antuñano, Kopecka, & Majid, 2014, for a similar argument). Pushing or pulling objects might correspond to one of these rather general manner categories. In any case, the question remains open for future research of why the conceptual representation of some but not other manner categories reflects cross-linguistic differences.

From a broader perspective, our findings also contribute new evidence to the more general question of how people weigh different components when judging event similarity. Uncontroversially, we found that cross-linguistic differences in event representation were played out against a backdrop of commonalities. First, the results confirmed our expectation that speakers of S- and V-languages should not differ in their focus on path because this component is routinely encoded in both language types (see also Kersten et al., 2010). More interestingly, the compound analysis found that the effect of path, a category that is often held to be basic and universal (Talmy, 2000a), diminished when participants did not describe the events (Experiment 2 vs. 1). This change in the importance of path across encoding conditions held for both languages and was accompanied by an increase in reliance on left/right direction, a component that is hardly ever expressed in motion descriptions. Together this suggests that speakers changed how they represented the events depending on the degree of verbal mediation.

Recall that, at the other extreme of the Whorfian hypothesis, there is the idea that spatial cognition is guided by universal conceptual categories such that motion cognition is independent of language (Papafragou, 2008). In light of the results of the compound analysis, we submit that a theoretical framing that sets as its primary goal the identification of language-independent 'core' conceptual categories might miss out on certain crucial insights as to how humans conceptualize events. Concretely, should a category such as path – which differentiates between motion upwards, inwards, etc. – be part of this core conceptual repertoire? At the very least, the compound analysis suggests that certain aspects of how we construe path are susceptible to linguistic mediation, which resonates with the view that conceptual categories are formed dynamically as a function of the task (Barsalou, 1983; Casasanto & Lupyan, 2015), rather than being fixed categories. Recent theoretical accounts consider strict divisions between verbal and non-verbal cognition as moot, arguing instead for a tight link between high-level cognitive representations and low-level conceptual categories (Lupyan, 2012; Lupyan & Clark, 2015). These accounts see cognitive processing as essentially flexible and task-dependent. Under this view, instead of dealing with language effects as superficial 'intrusions' into pure cognition (cf. Gleitman & Papafragou, 2012), we might conceive of language as one of the main ingredients that allows us to flexibly form the type

469

of conceptual representations needed to understand events. In that process, it seems that the mental categories we form will at least partially be coloured by the particular language we happen to speak.

Supplementary Materials

For supplementary material for this paper, please visit http://dx.doi.org/10.1017/langcog.2016.22.

REFERENCES

Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, **37**(2), 286–309.

Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., … Thierry, G. (2015). Two languages, two minds: flexible cognitive processing driven by language of operation. *Psychological Science*, **26**(4), 518–526.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**(4), 390–412.

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

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, **11**(3), 211–227.

Bates, D. M., Maechler, D., Bolker, B. M., & Walker, S. (2014). lme4: linear mixed-effects models using Eigen and S4 (Version R package version 1.1-7). Online: <http://CRAN.R-project.org/package=lme4>.

Bylund, E., & Athanasopoulos, P. (2014). Language and thought in a multilingual context: the case of isiXhosa. *Bilingualism: Language and Cognition*, **17**(2), 431–441.

Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion event cognition and grammatical aspect: evidence from Afrikaans. *Linguistics*, **51**, 929–955.

Cadierno, T., & Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics*, **4**(1), 183–216.

Casasanto, D., & Lupyan, G. (2015). All concepts are ad hoc concepts. In E. Margolis & S. Laurence (Eds.), *The conceptual mind: new directions in the study of concepts* (pp. 543–566). Cambridge, MA: MIT Press.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition*, **41**(1/3), 83–121.

Flecken, M., Gerwien, J., Carroll, M., & von Stutterheim, C. (2015). Analyzing gaze allocation during language planning: a cross-linguistic study on dynamic events. *Language and Cognition*, **7**(1), 138–166.

Flecken, M., von Stutterheim, C., & Carroll, M. (2014). Grammatical aspect influences motion event perception: findings from a cross- linguistic non-verbal recognition task. *Language and Cognition*, **6**(1), 45–78.

Gelman, A., & Loken, E. (n.d.). *The garden of forking paths: why multiple comparisons can be a problem, even when there is no 'fishing expedition' or 'p-hacking' and the research hypothesis was posited ahead of time*. Online: <http://www.stat.columbia.edu/~gelman/research/unpublished/>.

Gelman, A., & Su, Y.-S. (2014). arm: data analysis using regression and multilevel/hierarchical models (Version R package version 1.7-07). Online: <http://CRAN.R-project.org/package=arm>.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, **83**(1), 49–79.

Gleitman, L., & Papafragou, A. (2012). New perspectives on language and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 543–568). Oxford: Oxford University Press.

Goldstone, R. L. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods*, **26**(4), 381–386.

Hendriks, H., Hickmann, M., & Demagny, A.-C. (2008). How adult English learners of French express caused motion: a comparison with English and French natives. *Acquisition et Interaction en Langue Étrangère*, **27**, 15–41.

Hickmann, M., & Hendriks, H. (2010). Typological constraints on the acquisition of spatial language in French and English. *Cognitive Linguistics*, **21**(2), 189–215.

Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, **59**(4), 434–446.

Ji, Y., Hendriks, H., & Hickmann, M. (2011). How children express caused motion events in Chinese and English: universal and language-specific influences. *Lingua*, **121**(12), 1796–1819.

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, **139**(4), 638–653.

Lai, V. T., Garrido Rodriguez, G., & Narasimhan, B. (2014). Thinking-for-speaking in early and late bilinguals. *Bilingualism: Language and Cognition*, **17**(1), 139–152.

Loucks, J., & Pederson, E. (2011). Linguistic and non-linguistic categorization of complex motion events. In J. Bohnemeyer & E. Pederson (Eds.), *Event representation in language and cognition* (Vol. 11, pp. 108–133). Cambridge: Cambridge University Press.

Lupyan, G. (2012). Linguistically modulated perception and cognition: the label-feedback hypothesis. *Frontiers in Psychology*, **3**, 54. Online: <http://doi.org/10.3389/fpsyg.2012.00054>.

Lupyan, G., & Clark, A. (2015). Words and the World Predictive Coding and the Language-Perception-Cognition Interface. *Current Directions in Psychological Science*, **24**(4), 279–284.

Malt, B. C., Ameel, E., Imai, M., Gennari, S. P., Saji, N., & Majid, A. (2014). Human locomotion in languages: constraints on moving and meaning. *Journal of Memory and Language*, **74**, 107–123.

Papafragou, A. (2008). Space and the language-cognition interface. In P. Carruthers, S. Laurence, & S. P. Stich (Eds.), *The innate mind. Volume 3: foundations and the future* (pp. 272–289). Oxford: Oxford University Press.

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, **108**(1), 155–184.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: the representation of motion in language and cognition. *Cognition*, **84**(2), 189–219.

Papafragou, A., & Selimis, S. (2010). Event categorisation and language: a cross-linguistic study of motion. *Language and Cognitive Processes*, **25**(2), 224–260.

R Development Core Team. (2013). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Online: <http://www.R-project.org/>.

Ragnarsdóttir, H., & Strömqvist, S. (2004). Time, space, and manner in Swedish and Icelandic: narrative construction in two closely related languages. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative. Vol. 2: Typological and contextual perspectives* (pp. 113–141). Mahwah, NJ: Lawrence Erlbaum.

Regier, T., & Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends in Cognitive Sciences*, **13**(10), 439–446.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Slobin, D. I. (1996). From 'thought and language' to 'thinking for speaking'. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.

Slobin, D. I. (2003). Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.

Slobin, D. I., Ibarretxe-Antuñano, I., Kopecka, A., & Majid, A. (2014). Manners of human gait: a crosslinguistic event-naming study. *Cognitive Linguistics*, **25**(4), 701–741.

Stringer, D. (2005). *Paths in first language acquisition: motion through space in English, French and Japanese*. Unpublished doctoral dissertation, University of Durham.

471

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description. Vol. III: Grammatical categories and the lexicon* (pp. 57–149). Cambridge: Cambridge University Press.

Talmy, L. (2000a). *Toward a cognitive semantics: concept structuring systems*. Cambridge, MA: MIT Press.

Talmy, L. (2000b). *Toward a cognitive semantics: typology and process in concept structuring*. Cambridge, MA: MIT Press.

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: evidence from dual-task paradigms. *Journal of Memory and Language*, **63**(1), 64–82.

von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., & Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics*, **50**(4), 833–867.

Whorf, B. L. (1956). *Language, thought, and reality: selected writings*, ed. J. B. Carroll. Cambridge, MA: MIT Press.

Wolff, P. (2003). Direct causation in the linguistic coding and individuation of causal events. *Cognition*, **88**(1), 1–48.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, **2**(3), 253–265.

Wolff, P., Jeon, Ga-hyun, & Li, Yu (2009). Causers in English, Korean, and Chinese and the individuation of events. *Language and Cognition*, **1**(2), 167–196.