

ARTICLE

Language background impacts motion event categorization: evidence from English and Turkish speakers

Angelica Fulga¹  and Panos Athanasopoulos^{2,3}

¹Linguistics and the English Language, Lancaster University, UK; ²Centre for Languages and Literature, Lund University, Sweden and ³Department of General Linguistics, Stellenbosch University, South Africa

Corresponding author: Angelica Fulga; Email: angelicafulga@gmail.com

(Received 27 August 2025; Revised 09 January 2026; Accepted 13 March 2026)

Abstract

The languages individuals speak have been observed to influence how they perceive and interpret certain aspects of the world, but the conditions that give rise to these effects are not always clearly defined. One way to address the issue is to examine specific task contexts in which language is likely to modulate perception. Building on this, the aim of the present study was to investigate the role that native language (L1) plays in shaping motion processing in functional monolingual Turkish and English speakers (i.e., without intermediate or advanced proficiency in a second language) by using a supervised classification paradigm that involved sorting animated events based on either manner of motion or path. A total of 120 participants, including 60 native English speakers and 60 native Turkish speakers, were randomly assigned to one of the two conditions. The findings revealed a performance advantage for the English speakers in the manner-discrimination condition, in line with specific linguistic distinctions between the two languages. In contrast, overall performance was similar across the two groups when discrimination was based on path, a feature inherently expressed in both languages.

Keywords: cross-linguistic differences; language and thought; linguistic relativity; motion representation; working memory; the label-feedback hypothesis; predictive processing

1. Introduction

1.1. Mechanisms linking language and perception

Cross-linguistic differences were once believed to exert a pervasive effect on cognition, but current research has moved away from this absolutist perspective by adopting a more nuanced approach. This approach examines the mechanisms through which language may influence categorical perception in nonverbal tasks, that is, tasks in which language is not explicitly used (e.g., Athanasopoulos & Bylund,



2013; Athanasopoulos et al., 2015; Kersten et al., 2010). One of these mechanisms, the label-feedback hypothesis (Lupyan, 2012), suggests that during categorization, the human perceptual system automatically retrieves labels from long-term semantic memory to facilitate the abstraction of features from the environment. This process seems to occur even for linguistic content beyond word level, including grammatical structures and lexical constructions that encode motion events (Sato & Athanasopoulos, 2018). To gain insight into how labels/larger verbal structures modulate perception, the label-feedback hypothesis needs to be further explained in the contexts of the predictive processing framework (Lupyan & Clark, 2015) and the working memory model (Baddeley, 2000, 2003).

Predictive processing proposes that the brain uses prior knowledge to actively anticipate sensory input, and this knowledge is constantly updated through experience. In this scenario, perception is shaped by the ongoing interplay between prior knowledge, such as linguistic knowledge (e.g., knowledge of domain-specific labels) or assumptions about the world, and the continuous input provided by the senses. When these anticipations and the perceptual input match, language exerts an upregulating or augmenting effect on perception (e.g., Lupyan & Clark, 2015; Lupyan & Thompson-Schill, 2012; Lupyan et al., 2007). For example, hearing the word 'cat' activates a network of sensory and conceptual information related to cats which includes, among other features, their physical characteristics. This upregulation allows the brain to be more attuned to noticing these aspects in the environment, thus facilitating the faster and more accurate recognition of cats among other animals (Lupyan, 2012). Seen through the lens of predictive processing, language used in communication with others, along with self-directed speech, becomes a powerful tool that shapes thought processes and reasoning (Lupyan & Clark, 2015).

These observations are crucial within the scope of the working memory model (Baddeley, 2000, 2003). Within the latter approach, sensory information is briefly stored in a short-term store before being transferred to a phonological output system (a phonological output buffer), resulting in either verbal output or mental rehearsal (Baddeley, 2003). These processes are supported by the phonological loop, a component that maintains and manipulates verbal information through subvocal (mental) rehearsal. The phonological loop interacts with the central executive, a system responsible for the allocation of attentional resources and decision-making, and the visuospatial sketchpad, which is involved in processing nonverbal information. Finally, information from the phonological loop, visual sketchpad and long-term memory is integrated through the episodic buffer (Baddeley, 2000), forming a coherent representation of reality.

These insights suggest that knowledge of labels may establish a neural and biological environment in which speakers may be immersed, even during nonverbal tasks, via, for instance, self-directed speech. For this reason, an analysis of the effects of language on thought would not be complete without including these mental processes (Bylund & Athanasopoulos, 2013).

1.2. Nonverbal paradigms targeting early perceptual processing

Nonverbal paradigms are methods that do not require overt use of language, namely linguistic input and output. Studies using nonverbal methods that tap into early, automatic perceptual processing, such as breaking continuous flash suppression

(b-CFS), have shown cases where language appears to modulate low-level visual sensitivity. Low-level perception represents the incipient stages of sensory processing, which involve the detection and encoding of basic visual features, such as shape, color, or manner and path of motion. In the domain of motion, b-CFS was used by Vanek and Fu (2022) in a feasibility study to assess how quickly functionally monolingual speakers of English and Mandarin Chinese become aware of visual properties of motion events, including manner salience, path salience and neutral events. These two languages were chosen because of differences in how they encode manner and path when expressing motion events, which may be reflected in the timing of participants' detection of motion features. More specifically, Mandarin Chinese plays equal emphasis on manner and path, which are frequently expressed through serial-verb constructions (Talmy, 2000). In contrast, English tends to express manner in the verb and path in satellites such as prepositional phrases (see *Section 2.1*).

The b-CFS technique measures visual stimulus processing, as it emerges into awareness, by initially making the motion stimulus invisible using a high-contrast, rapidly flashing mask presented to the dominant eye. A low-contrast motion event is then introduced to the nondominant eye, with its contrast gradually increased. Observers indicate as quickly as possible when and where on the screen they detect the motion event, and the detection time serves as a measure of how speakers' language may influence perception during event suppression (Vanek & Fu, 2022). Compared with the Mandarin speakers, the English speakers exhibited faster reaction times when manner was the salient feature.

While these findings need to be considered carefully, given the pilot nature of the study, they are consistent with the engagement of predictive mechanisms in the visual system. Because the stimulus emerged gradually from suppression, the paradigm provided a temporal window in which prior knowledge and expectations could influence the timing of visual awareness in a top-down fashion, making language-specific effects detectable (Lupyan & Clark, 2015). Contrary to this finding, a parallel body of work indicates that linguistic differences do not always penetrate low-level perceptual processing. Studies in motion cognition repeatedly found no cross-linguistic differences when stimuli were processed online, without memory demands or explicit linguistic activation (e.g., Gennari et al., 2002; Papafragou & Selimis, 2010). In these cases, the perceptual system may have processed the visual stimuli directly, without the need to predict missing information, which resulted in similar sensitivity to the raw visual features of motion across languages. Language-specific effects become detectable again, however, when the tasks require higher level categorization involving working memory (Bylund & Athanasopoulos, 2013; Bylund et al., 2013; Montero-Melis & Bylund, 2016), as detailed below.

1.3. *Nonverbal paradigms targeting higher level categorization*

Higher level processes represent complex cognitive functions that involve analyzing information, rather than just perceiving it, and include tasks such as classifying objects and events. Among these, triad-matching tasks revealed an effect of language when the latter was implicitly activated via task instructions (e.g., during classification of motion events: Papafragou & Selimis, 2010), or when the task required holding the events in working memory before making a decision, as the target and

the alternative clips were played sequentially rather than simultaneously (e.g., studies examining how grammatical aspect guides motion-event classification: Athanasopoulos & Bylund, 2013; Bylund & Athanasopoulos, 2013; Bylund et al., 2013). Unlike triad-matching paradigms, which allow participants to select whichever alternative seems most similar to the target, another type of task, supervised classification, requires mapping events onto predetermined categories, such as manner-based or path-based groups. By providing participants with feedback on their categorization choices (e.g., a green checkmark indicates correct category assignment, whereas a red X indicates an incorrect answer), the task guides attention to the relevant dimension (e.g., manner or path). If the relevant feature (e.g., manner) is habitually encoded in the participants' language (e.g., English), predictive processing may make the pattern easy to detect and learn. By the same token, if the critical feature is less reliably encoded in the language, it might compete with features that are regularly expressed, making detection of the task-relevant feature more difficult.

Supervised classification was successfully used by Kersten et al. (2010) to investigate manner-based versus path-based categorization in speakers of English and Spanish. These two languages were selected because they differ in how manner and path are encoded in motion events, with Spanish generally exhibiting a lower tendency than English for encoding manner in the main verb (see Section 2.1). In their task, participants viewed animated 'bug-like' creatures that moved along one of four possible paths while exhibiting one of four possible manners of motion. On each trial, only one dimension (either manner or path) was relevant for determining the correct category. Kersten et al. found that English speakers outperformed Spanish speakers in the manner condition, a result consistent with typological differences between these two languages.

Taken together, research employing nonverbal paradigms suggests that linguistic experience can modulate processing at multiple levels, but not in a uniform or obligatory manner. Paradigms targeting early perceptual processing indicate that linguistic categories can shape predictive expectations and influence the timing or sensitivity of visual responses under conditions that engage top-down effects. By contrast, higher level nonverbal categorization tasks consistently show that cross-linguistic differences emerge primarily when tasks require abstraction, selective attention and the retention of task-relevant dimensions over time. Crucially, these effects are contingent on task demands and cognitive load rather than reflecting a pervasive perceptual influence. Together, this body of work supports a view of linguistic relativity as context-dependent, motivating the present study's focus on how typological differences in motion encoding interact with learning and attentional mechanisms during supervised nonverbal categorization.

2. The current study

Kersten et al.'s (2010) study has become a seminal reference point for eliciting linguistic effects in motion categorization without explicit use of language, especially since this field is marked by mixed findings (Casasanto, 2008; Ünal & Papafragou, 2016). To date, however, no independent replication has been conducted. The present study addresses this gap by re-employing Kersten et al.'s supervised classification task with native speakers of English and extending the methodology to native speakers of Turkish. While previous studies have examined Turkish speakers' language-driven performance

in motion tasks (e.g., Özçalışkan, 2005; Özçalışkan & Slobin, 2003; Özyürek et al., 2005), the perception of manner and path of motion has not been systematically explored in nonverbal tasks. This kind of approach is necessary to provide a nuanced examination of how Turkish influences motion cognition beyond linguistic expression.

2.1. Cross-linguistic encoding of motion events

When analyzing cross-linguistic differences in motion cognition, the categorical perception of path and manner of motion is particularly intriguing due to the variety in the encoding of these two dimensions across different languages (Cardini, 2010; Czechowska & Ewert, 2011; Finkbeiner et al., 2002; Feinmann, 2020; Filipović, 2011; Gennari et al., 2002; Hohenstein, 2005; Kersten et al., 2010; Papafragou & Selimis, 2010; Papafragou et al., 2002, 2008). This distinction inspired the classification of linguistic systems into equipollent (Slobin, 2004; Zlatev & Yangklang, 2004), satellite-framed, and verb-framed languages (Talmy, 2000). Equipollent languages express both path and manner in equivalent verbs, as observed in, for example, Thai and Mandarin Chinese (Slobin, 2004; Zlatev & Yangklang, 2004). Satellite-framed languages, including English and German, typically express manner within the verb, while path is encoded in a grammatical element associated with a verb, such as a particle, prefix, or prepositional phrase. For instance, a motion event can be rendered in English as *The dog bounced down the stairs*. Here, the verb *bounced* expresses how the dog moved, whereas *down the stairs* denotes trajectory or path. Within the latter structure, *down* is the preposition, with *the stairs* representing the ground, or landmark. Notably, this contrast is not absolute, as English also employs path verbs, such as *enter* and *descend*.

Unlike satellite-framed languages, verb-framed languages, which include Turkish and Spanish, typically express path in the verb, while manner is either expressed in additional structures or ignored if implicit (Talmy, 2000). The equivalent of the aforementioned English sentence would be in Turkish, *Çocuk merdivenlerden zıplayarak indi* ‘the child descended the stairs while bouncing’. Here, the verb *indi* (descended) encodes the path, while manner is encoded in the additional element, *zıplayarak* (while bouncing).

Although this classification provides a useful basis for differentiating motion expression patterns, it is important to note that, as mentioned above with regard to path expression in English, both satellite-framed and verb-framed languages employ various strategies for encoding path and manner. This reflects a typological tendency rather than an absolute distinction (Beavers et al., 2010).

2.2. Motion-event expression in Turkish

In contrast to Spanish, which is an Indo-European language in the Romance language family, Turkish is a non-Indo-European language belonging to the Turkic language family. Similar to Spanish, however, Turkish is a verb-framed language, in which path is encoded in the verb stem, while manner is expressed only optionally in adjuncts (Talmy, 1991). For example, the English sentence *The ball bounced down the hill* can be rendered as *Top zıplayarak tepeden aşağı indi* in Turkish, which translates as ‘The ball descended the hill while bouncing’ (Özyürek et al., 2005). Here, the verb *indi* encodes path, and manner is relegated to a subordinate converbial form.

Production studies corroborate this typological contrast between Turkish and satellite-framed languages such as English, further detailing where these differences lie. Although Turkish has various means to express manner, speakers encode it significantly less often than English speakers, and generally only when it is causally central to how an event unfolds (Özçalışkan, 2005; Özçalışkan & Slobin, 2003). Under time pressure, they frequently omit manner altogether, while consistently expressing path (Özyürek et al., 2005). This reduced expression frequency is partly attributed to the structures required for manner expression in Turkish, which include subordinate constructions (1) or converb forms (2):

- (1) Sürüne sürüne çıktı.
crawl-ADV crawl-ADV exit-PST
'She/he exited crawling, crawling'.
- (2) Koşarak içeri girdi.
run-CONV inside enter-PST
'She/he entered running'.

This analysis should not be taken to imply that Turkish lacks manner verbs. Özçalışkan and Slobin (2003) found that one of the differences between English and Turkish motion descriptions across the novels and oral narratives they analyzed lies not in the absence of a manner lexicon in Turkish but in the relative frequency and variety of the manner verbs employed in English. Similarly, Ünal et al. (2022) also found that descriptions of motion events (e.g., one character moving relative to a landmark) contained both path and manner information, with the path expressed in the verb; however, gestures mainly encoded path alone.

2.3. Theoretical implications of motion typology for predictive processing and working memory

Within the predictive-processing framework, these observations suggest that Turkish speakers receive consistent linguistic reinforcement for path in everyday language use, whereas manner, being encoded less reliably, receives weaker predictive support. On the other hand, for English speakers, frequent and obligatory manner encoding leads to more stable reinforcement for manner features. From a working memory perspective, habitual encoding of patterns may influence how efficiently Turkish and English speakers maintain motion information during categorization. Rather than reflecting differential use of the working memory per se, cross-linguistic variation may reflect the strength and accessibility of motion representations, with consequences for which dimension is prioritized when categorization requires retaining task-relevant information.

Taken together, typological differences between English and Turkish, cross-linguistic production patterns and experience-based label-feedback mechanisms converge on a consistent prediction: manner information will be prioritized during motion categorization for English speakers, whereas Turkish speakers' attention will be biased toward path. This integrated theoretical account provides the foundation for the hypothesis tested in the present study.

Hypothesis. Given the lack of consistency in manner encoding in their language, Turkish speakers are expected to show lower categorization accuracy than English speakers in the manner-discrimination condition.

3. Method

3.1. Participants

A total of 60 English-speaking and 60 Turkish-speaking participants took part in this study. This sample size was chosen a priori to match the design of Kersten et al. (2010), Experiment 2, whose supervised classification paradigm we replicate and extend. A priori power analysis was not conducted because the main statistical models in this study were mixed-effects models. Unlike simple tests such as t-tests or ANOVAs, mixed-effects models do not generally allow analytical power calculations due to their multilevel structure and multiple sources of variance. To estimate power, one would need to run large simulations, and these simulations require knowing the variability between participants and items ahead of time. Because Kersten et al. (2010) did not report these values, and they cannot be meaningfully guessed before collecting data, any simulated power estimate would have been unreliable. For this reason, the sample size followed the original study, which used 30 participants per group. Table 1 presents background information for 120 participants, including 60 functional monolingual Turkish speakers and 60 functional monolingual English speakers.

The participants' native languages and their knowledge of additional languages were assessed using the self-reporting Language Experience and Proficiency Questionnaire (LEAP-Q; Marian et al., 2007). Functional monolingualism was defined as a lack of proficiency in a second language, based on self-reports and QPT scores for the Turkish speakers. The Turkish group, selected via referrals, comprised 34 males and 26 females aged 18–42 ($M = 20.2$, $SD = 4$), who were university students or professionals residing in Türkiye. They all mandated Turkish to be their native, primary language, with no proficiency in an additional language other than English. They reported acquiring English early in their schooling (M age of onset = 10.63 years, $SD = 2.60$) and self-rated their proficiency in English as poor to basic. This was supported by their scores on the 60-item English Oxford Quick Placement Test (QPT; University of Cambridge Local Examinations Syndicate, 2001), with a mean score of 21.5 ($SD = 5.4$), corresponding to beginner to elementary levels on the CEFR scale (A1–A2). Turkish participants in the manner-discrimination ($M = 22.73$, $SD = 3.97$) and path-discrimination conditions ($M = 21.63$, $SD = 4.16$) showed comparable English proficiency based on the QPT scores.

The English-speaking group comprised 44 females and 16 males ($M = 31$, $SD = 12$), recruited through advertisements from among graduate students in a UK university and English teachers residing in Türkiye. The English speakers mandated

Table 1. Participant background information and task assignment

Language group	Age	Gender		Formal education (years)	Categorization task	
	<i>M/SD</i>	Male	Female	<i>M/SD</i>	Manner	Path
English speakers	30.88/11.97	16	44	16.07/1.78	30	30
Turkish speakers	20.21/4.03	34	26	13.13/.74	30	30

English as their native, primary language. The native English speakers living in Türkiye reported using exclusively English in their daily interactions, both at work and with friends. We excluded data from the English-speaking participants who reported a proficiency score above 5 out of 10 in Turkish or any other language (e.g., German, Spanish, French) or who reported using Turkish in their daily interactions. All participants received remuneration for their participation in this study.

A **Welch's two-sample *t*-test** was conducted using R version 4.3.1 (R Core Team, 2023) to compare background differences between the groups. An alpha level of 0.05 was adopted for all these analyses. The results indicated that the **English speakers** had a significantly higher mean age ($M = 30.88$, $SD = 11.97$) compared to the **Turkish speakers** ($M = 20.21$, $SD = 4.03$), $t(439.12) = 16.02$, $p < .001$, 95% CI [9.36, 11.98]. The English speakers also had significantly higher levels of **formal education** ($M = 16.07$, $SD = 1.77$) compared to the **Turkish speakers** ($M = 13.13$, $SD = 0.74$), $t(481.2) = 28.98$, $p < .001$, 95% CI [2.73, 3.13]. These differences arise because all Turkish students who do not demonstrate English proficiency at the B2 (high-intermediate) level are required in Türkiye to complete a one-year English preparatory program before beginning their undergraduate education. Therefore, in order to recruit functional monolingual Turkish speakers, participant recruitment should take place among students enrolled in these preparatory programs at the beginning of the academic year, when students are typically between 18 and 20 years old. On the other hand, the English-speaking participants were recruited from students who had already begun or were nearing completion of their undergraduate or graduate programs, as well as from professional groups, including English teachers.

3.2. Materials

The stimuli consisted of bug-like creatures (Ada, 2018) animated using HTML, CSS and Javascript (Eminoğlu, 2019). Screen recordings of the manner and path categorization tasks, each containing four clips illustrating the four categories, can be found at the following link: https://osf.io/8b4c6/overview?view_only=16ecd0c689b4b62889cb6d0996caba6.

These materials are similar to those used in Kersten et al. (2010) Experiment 2, with the following difference: instead of using the labels Species 1–4, the categories in this study were labeled with the numbers 1–4. As in Kersten et al. (Experiment 2), no verbal labels (novel nouns or verbs) were presented.

Each clip featured a static bug (target) and a moving bug, with the latter exhibiting four different manners of motion and following four paths with respect to the target (see Figure 1). More precisely, the moving bug approached the static bug or moved away from it either along a straight trajectory or along a nonlinear (indirect) path, with a single 90° turn halfway through the event. These paths can be expressed in English as *directly/indirectly toward* and *directly/indirectly away* (Kersten et al., 2010). Conversely, in Turkish, 'indirectly toward/away' corresponds to *dolaylı yaklaşmak/uzaklaşmak*, whereas 'directly toward/away' corresponds to *doğrudan yaklaşmak/uzaklaşmak*.

The task consisted of 6 blocks, each containing 16 events (4 paths $\times$ 4 manners), resulting in a total of 96 clips. Creature appearance, including body shape, color and head shape, as well as background scene variations featuring simplified swamp vegetation, had four possible instantiations per event, which were randomly assigned. In the manner-discrimination condition, correct classification depended on the manner of motion, while for the path-discrimination task,

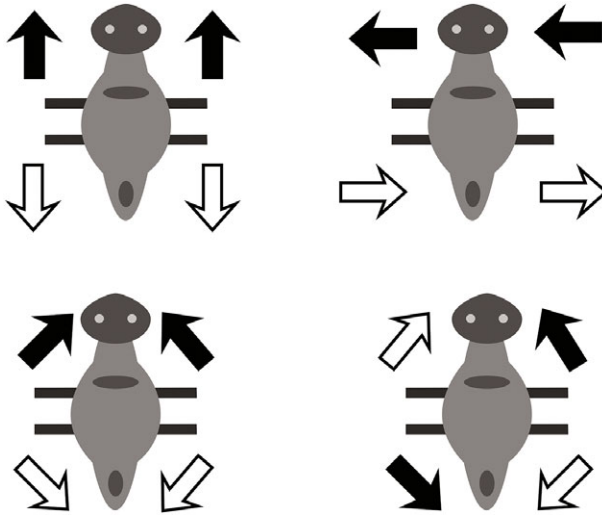


Figure 1. The four different manners of motion used in the experiment.

Note: Movement started from the depicted position, went in the direction indicated by the dark arrows, returned to initial position, and then moved in the direction shown by the clear arrow. *Source:* Adapted from Kersten et al. (2010).

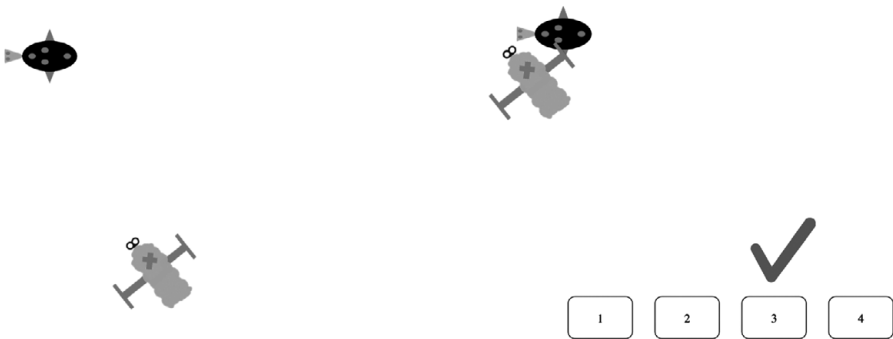


Figure 2. Two frames from an example event.

Note: Unlike Kersten et al. (2010), in which each category number is accompanied by the word *species*, only digits were used in this study.

correct category assignment was dictated by the path the moving entity traveled on. Because the two dimensions of motion differed orthogonally (Kersten et al., 2010), when manner of motion predicted the correct category, path was irrelevant, whereas when path of motion determined category assignment, manner was irrelevant. Bug location on the screen varied, and each motion event, lasting 2–3 seconds, was centered on the computer screen. Four buttons, representing categories 1, 2, 3, and 4, appeared postevent, allowing participants to categorize by clicking on the corresponding button. Feedback was provided in the form of a green checkmark for each correct answer and a red X for each incorrect classification (Figure 2).

3.3. Design

This study employed a mixed factorial design with one within-subjects factor and two between-subjects factors. The within-subjects factor was block (six levels: block 1–6). The between-subjects factors were language group (Turkish vs. English) and condition (manner vs. path). Path categorization was included as a control condition, as both Turkish and English systematically encode path and are therefore not expected to differ significantly in this domain. In addition, age was included as a covariate to control for potential age-related influences on performance. The dependent variable was participants' categorization accuracy (correct vs. incorrect).

3.3.1. Model

Analyses were conducted using mixed-effects models with crossed random effects for subjects and items, fitted with the `glmer()` function from the *lme4* package (version 1.1-37) in R (version 4.4.3; R Core Team, 2023), specifying a binomial family with a logit link function to examine differences in performance between the English and Turkish speakers in the manner and path conditions. Accuracy was modeled as a binomial outcome using the two-column syntax `cbind(item correct, incorrect)`. Each block presented four categories with four trials per category ($4 \times 4 = 16$ trials per block), so for each category, the response was 0–4 correct. Fixed effects were contrast-coded as language (Turkish = -1 vs. English = 1) and condition (path = -1 vs. manner = 1), and block was entered as a centered numeric predictor (1–6). Age (centered) was also included in the model as a covariate. The model included random intercepts for item, as well as random intercepts and random slopes for block by subject (Linck & Cummings, 2015). Following Linck and Cummings (2015), the optimized bobyqa was used to aid convergence. The model formula was `cbind(Item_Correct, Incorrect) ~ Block_c * Language * Condition + Age_c + (1 + Block_c | Subject) + (1 | Item)`.

3.4. Procedure

Instruction was provided in English to the monolingual English speakers and in Turkish to the Turkish monolingual speakers. The instructions in Turkish were translated from English by a native Turkish speaker, who was proficient in English. The procedure follows that of Kersten et al. (2010, Experiment 2), with the following difference: no break was provided at the end of each block during which participants could see their scores. The rationale behind this approach is explained below.

The choice to forego breaks between blocks was deliberate, being supported by the fact that the task was not long enough to cause cognitive fatigue, as well as by considerations related to the specific nature of the task. More specifically, the categorization paradigm lasted approximately 15 minutes and was visually engaging and emotionally neutral. Its structure imposed low intrinsic (inherent task complexity), extraneous (cognitive demands from task instruction) and germane loads. From the perspective of cognitive load theory (Sweller, 1988), tasks with such low overall cognitive load would not be expected to induce cognitive fatigue or performance decline. On the other hand, displaying scores under these circumstances may introduce unnecessary performance pressure or anxiety, potentially affecting individual behavior. Finally, the choice not to introduce breaks at the end of each block was dictated by the nature of the task itself. This supervised classification paradigm was intended to measure implicit, continuous category learning driven by perceptual

feedback, rather than explicit reasoning or verbal hypothesis testing. Introducing breaks provides an opportunity for participants to reflect consciously on possible rules or develop verbal strategies, which would conflict with the task's aim of capturing uninterrupted perceptual learning. To sum up, the decision to present the task as a continuous sequence was made after careful consideration of all of the aspects detailed above.

Participants were informed both orally and through the consent form about the purpose of the study and were told that participation was voluntary, with the option to withdraw at any time during data collection. After providing consent, they read the instructions, presented on a PowerPoint slide, and were given the opportunity to ask clarification questions. Then, they proceeded with the categorization. Testing was performed individually, in a quiet setting, with the administration of the QPTs and background information forms following the experimental task. This study received research ethics approval from the UK (Lancaster University's Research Ethics Committee) and Türkiye (Yildiz Technical University's Research Ethics Committee).

4. Results

The results are depicted in Figures 3 and 4 and Table 2. Scaled residuals indicated adequate model fit ($IQR = -0.84$ to 0.76 ; no residual exceeded ± 5). Random-effects estimates showed substantial subject-level variability in both baseline accuracy ($SD = 1.06$) and improvement across the six blocks ($SD = 0.36$, $r = .85$), suggesting that participants with higher initial accuracy tended to improve more over time. The variability between items was smaller, $SD = 0.43$, indicating relatively consistent difficulty across the four categories.

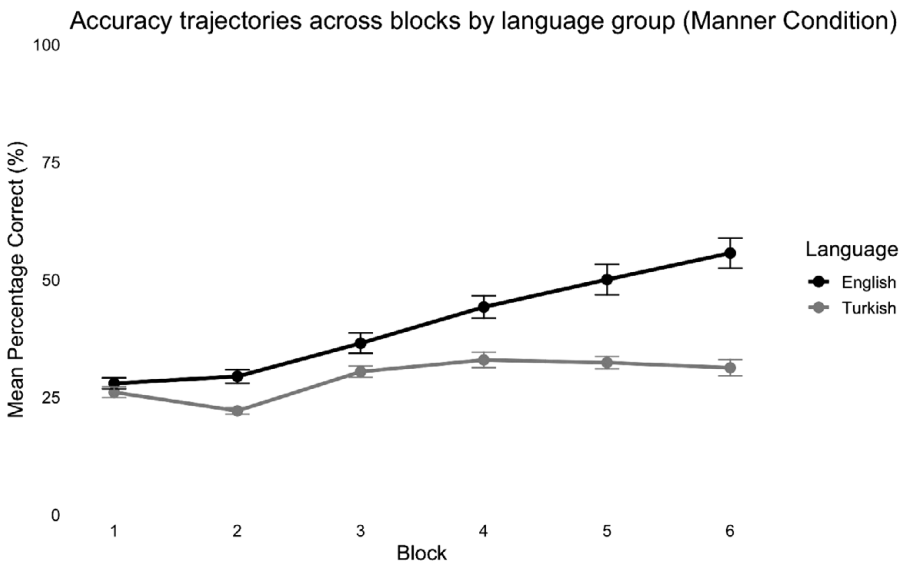


Figure 3. Mean percentage of correct responses across experimental blocks in the manner condition. *Note:* The responses are shown separately for monolingual English speakers (dark gray line) and monolingual Turkish speakers (light gray line). Error bars represent standard errors of the mean. Twenty-five percent reflects chance performance.

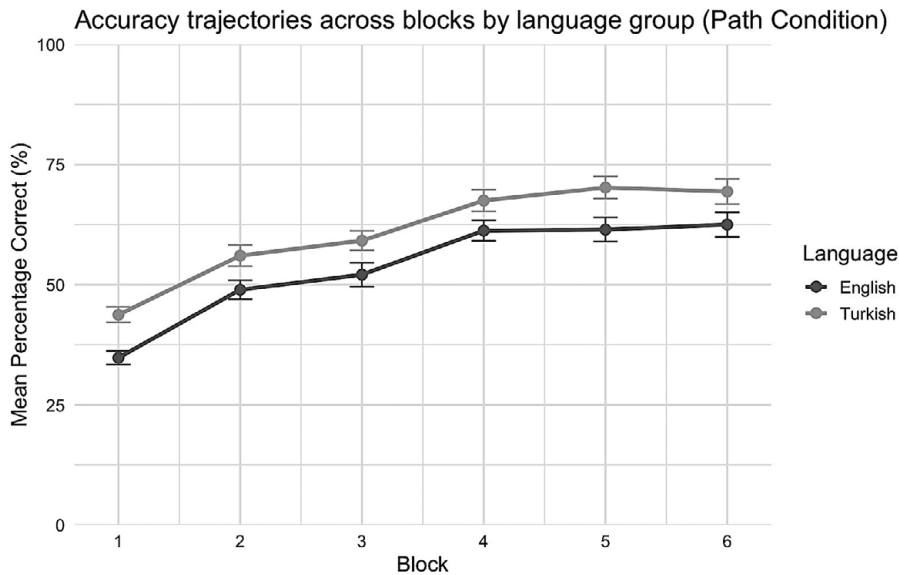


Figure 4. Mean percentage of correct responses across experimental blocks in the path condition. Note: The responses are shown separately for monolingual English speakers (dark gray line) and monolingual Turkish speakers (light gray line). Error bars represent standard errors of the mean. Twenty-five percent reflects chance performance.

The results revealed that accuracy significantly increased across blocks ($estimate = 0.28$, $SE = 0.036$, $z = 7.74$, $p < .001$), and performance was significantly higher in the path condition than in the manner condition ($estimate = -0.56$, $SE = 0.1$, $z = -5.56$, $p < .001$). The main effects of language and age were not significant ($ps > .005$). However, a significant language $\times$ condition interaction was found ($estimate = 0.27$, $SE = 0.100$, $z =$

Table 2. Fixed and random effects for manner and path conditions

Parameters	Fixed effects				Random effects	
	<i>Estimate</i>	<i>SE</i>	<i>z</i>	<i>p</i>	By subject	By item
					<i>SD</i>	<i>SD</i>
Intercept	0.088	0.23	0.37	.7	1.06	0.42
Block	0.27	0.035	7.73	< .001	0.35	–
Language	0.09	0.1	0.86	.38	–	–
Condition	–0.55	0.1	–5.5	< .001	–	–
Age	0.009	0.007	–1.32	.18	–	–
Block $\times$ Language	0.049	0.035	1.39	.16	–	–
Block $\times$ Condition	–0.053	0.035	–1.51	.12	–	–
Language $\times$ Condition	0.27	0.1	2.74	.006*	–	–
Block $\times$ Language $\times$ Condition	0.06	0.035	1.79	.07	–	–

Note: Fixed effects are reported with unstandardized estimates, standard errors (*SEs*), Wald *z* statistics and *p*-values. Random effects show standard deviations (*SDs*) for intercepts and slopes by subjects and by items. Block was centered prior to analysis. Language was contrast-coded (English = 1, Turkish = –1), and condition was contrast-coded (manner = 1, path = –1). *p* values marked with an asterisk (*) are significant at $p < .05$. The block $\times$ language $\times$ condition interaction is reported for descriptive purposes only and should not be interpreted in terms of learning rate, which was not assessed in this study.

2.74, $p = .006$), indicating that differences between the manner and path conditions varied by language group.

To further explore the significant main effect of condition, the analysis was conducted separately for the manner and path conditions. The results are depicted in Tables 3 (manner) and 4 (path). Mixed-effects logistic regression was conducted to assess performance in the manner condition, with block (centered), language (English vs. Turkish) and age (centered) as fixed effects. The model included crossed random effects for subjects (random intercepts and slopes for block) and items (random intercepts). Random-effects estimates indicated substantial subject-level variability both in baseline accuracy ($SD = 0.98$) and in performance across the six blocks ($SD = 0.35$), as well as smaller item-level variability ($SD = 0.19$).

Fixed-effects results revealed a significant main effect of block, $estimate = 0.214$, $SE = 0.050$, $z = 4.301$, $p < .001$, indicating that accuracy increased across blocks. The main effect of language was also significant, $estimate = 0.342$, $SE = 0.142$, $z = 2.412$, $p = .016$, with English speakers showing higher accuracy than Turkish speakers. The interaction between block and language was significant, $estimate = 0.110$, $SE = 0.050$, $z = 2.214$, $p = .027$, suggesting that cross-linguistic differences varied as a function of block. Importantly, block is not interpreted as an assessment of learning rate, but in terms of stages of exposure indicating when language-specific differences became

Table 3. Fixed and random effects for the manner condition

Parameters	Fixed effects				Random effects	
	Estimate	SE	z	p	By subjects	By items
					SD	SD
Intercept	0.61	0.16	−3.8	< .001	0.97	0.19
Block	0.21	0.05	4.3	< .001	0.35	—
Language	0.34	0.14	2.41	.016*	—	—
Age	−0.006	0.0094	−0.68	.49	—	—
Block × Language	0.11	0.05	2.21	.027*	—	—

Note: Fixed effects are reported with unstandardized estimates, standard errors (SEs), Wald z statistics and p-values. Random effects show standard deviations (SDs) for intercepts and slopes by subjects and by items. Block and age were centered prior to analysis. Language was contrast-coded (English = 1, Turkish = −1). p values marked with an asterisk (*) are significant at $p < .05$.

Table 4. Fixed and random effects for the path condition

Parameters	Fixed effects				Random effects	
	Estimate	SE	z	p	By subjects	By items
					SD	SD
Intercept	0.48	0.36	1.33	.18	1.17	0.65
Block	0.34	0.052	6.64	< .001	0.36	—
Language	0.34	0.14	2.41	.18	—	—
Age	−0.01	0.01	−1.16	.24	—	—
Block × Language	−0.01	0.05	−0.27	.78	—	—

Note: Fixed effects are reported with unstandardized estimates, standard errors (SEs), Wald z statistics and p-values. Random effects show standard deviations (SDs) for intercepts and slopes by subjects and by items. Block was centered prior to analysis. Language was contrast-coded (English = 1, Turkish = −1). p values marked with an asterisk (*) are significant at $p < .05$.

detectable rather than differences in speed of learning. As illustrated in Figure 3, these differences were most pronounced in the final block ($estimate = -1.23$, $SE = 0.49$, $z = -2.48$, $p = .012$). The effect of age was not significant ($p > .05$), suggesting that age did not reliably predict accuracy.

Mixed-effects logistic regression was conducted to assess performance in the path condition (Table 4), with block (centered), language (English vs. Turkish) and age (centered) as fixed effects. The model included crossed random effects for subjects (random intercepts and slopes for block) and items (random intercepts). Random-effects estimates indicated considerable subject-level variability both in baseline accuracy ($SD = 1.18$) and in performance across the six blocks ($SD = 0.36$), as well as moderate item-level variability ($SD = 0.66$). Fixed-effects results revealed only a significant main effect of block, $estimate = 0.34$, $SE = 0.052$, $z = 6.65$, $p < .001$, indicating that learning the path-based categorization pattern occurred regardless of language group. The lack of significance for age suggests that age did not reliably predict accuracy in the path-discrimination task.

Complementary simple effects analyses were computed to assess differences in performance across the manner and path conditions within each language group. Among English speakers, the predicted probability of a correct response was significantly higher for path trials ($M = 0.560$, $SE = 0.072$) than for manner trials ($M = 0.418$, $SE = 0.071$), $z = -2.01$, $p = .044$, $odds\ ratio = 0.57$. Among Turkish speakers, this difference was even more pronounced (path trials: $M = 0.667$, $SE = 0.065$; manner trials: $M = 0.278$, $SE = 0.059$), $z = -5.80$, $p < .001$, $odds\ ratio = 0.19$. These results suggest that path trials were less demanding than manner trials, clarify the source of the language $\times$ condition interaction and support the interpretation that manner of motion posed a greater challenge for Turkish speakers than for English speakers.

5. Discussion

This study investigated whether the typological patterns in path encoding proposed by Talmy (1985) are reflected in nonverbal categorization of motion events. We hypothesized that Turkish speakers would show lower accuracy in the manner classification task when compared to English speakers. The results confirmed this difference in performance in the manner categorization task only, broadly consistent with the pattern found by Kersten et al. (2010) and suggesting that speaking English might provide an advantage in prioritizing manner information during categorization of motion events. Importantly, our results diverge from Kersten et al.'s (2010) findings, as these differences were not observable from the task outset. More precisely, while Kersten et al. (2010) observed an early differentiation in performance in the manner condition between Spanish and English speakers, in this study, the difference emerges later, starting from block 3. One explanation for this discrepancy lies in the structure of the task.

In Kersten et al., breaks at the end of each block, during which participants viewed their scores, may have facilitated rule discovery, amplifying language-specific biases early. In contrast, the continuous design used here reduced opportunities for explicit strategy formation, allowing typological effects to appear through cumulative exposure. Differences in categorization between the two groups emerging later in the task suggest that language does not constrain initial perception of manner but may influence how readily manner becomes a stable feature for categorization over time. An alternative explanation for these late-emerging effects would be that, within the predictive-

processing framework, manner information may compete for Turkish speakers with more strongly reinforced path representations. When combined with the need to maintain manner as the relevant dimension in working memory across trials, this mismatch may have increased competition from path information, making manner-based categorization progressively more difficult for Turkish speakers.

A further explanation for the divergence between the results of this current study and those reported by Kersten et al. (2010) might be linguistic. Although both Spanish and Turkish are verb-framed languages, they may vary in how systematically manner information is included in motion utterances, potentially contributing to the different performance trajectories observed. On the other hand, in both studies, participants performed equally well in the path-discrimination task, ruling out the idea that the advantage English speakers had in the manner-discrimination task was caused by overall performance disparities between the two language groups (Kersten et al., 2010). The consistent improvement in the path condition when compared to the manner condition suggests that path may be more predictable across trials. This aligns with Skordos et al.'s (2020) findings, which suggest that path may be more memorable compared to manner. This phenomenon is attributed to the status of path as a core element of motion (Talmy, 2000), facilitating stronger top-down predictions. The participants may have formed a stable model early on, leading to reduced prediction errors and forming a steep learning curve. The finding that path discrimination was easier than manner-based discrimination for both language groups is in line with Kersten et al.'s (2010) finding in Experiment 1 of their study. They attributed it to the intentional use of verb labels, which may draw attention to path, as suggested by previous research (e.g., Kersten, 1998; Kersten & Billman, 1997). Regarding the present study, although the word *species* was not used to discriminate among the four categories of bugs, each type was numerically labeled (e.g., 1, 2, 3, 4), and these labels can be considered nouns. Thus, it may be inferred that, despite the use of noun labels in our study, both English and Turkish speakers found path-based classification easier compared to manner-based discrimination. Obtaining the same pattern of results with numerical labels as Kersten et al. (2010) did with verbal labels indicates that numerical identifiers may represent subtle cues in the perception of the path dimension, highlighting the need for further research on motion perception in the presence of lexical or symbolic identifiers.

In summary, the present findings suggest that Turkish and English speakers differ in how they utilize motion information during categorization in ways that align with typological tendencies described in the literature. While both groups performed comparably in the path-discrimination task, differences emerged when manner was the relevant basis for categorization. These results indicate that language-specific patterns of motion encoding may guide attention during nonverbal categorization, providing further evidence for context-dependent effects of language on cognition.

6. Limitations

Several limitations should be considered when interpreting the present findings. First, the English-speaking participants were on average older than the Turkish-speaking participants. Age-related differences in cognitive performance could in principle influence categorization accuracy; however, this factor is unlikely to account for the observed pattern, as no cross-linguistic differences were found in the path-

discrimination task, which was performed comparably by both groups. Any general age-related advantage would be expected to affect performance across conditions rather than selectively in the manner task.

Second, supervised categorization paradigms, which provide immediate feedback, may increase competition between the path and manner dimensions compared to more naturalistic or unsupervised tasks, potentially enhancing language-specific biases. However, this type of task was well suited to the aims of this study, which sought to identify the contexts in which cross-linguistic differences are more likely to affect perception.

Another limitation is that, despite the consistency between these results and typological tendencies in motion encoding, the present design does not allow us to localize the effect to a specific level of processing. In particular, the late emergence of the language difference suggests that typology does not constrain initial access to manner information but rather influences how motion dimensions are prioritized during learning, attention, or decision-making.

More specifically, because the task required repeated exposure to a fixed set of motion events, it remains unclear to what extent the observed effects generalize beyond learning contexts to more spontaneous forms of motion perception. Future studies using multiple approaches, including online measures of attention or learning, could help clarify how predictive processing and working memory each contribute to language-related effects in motion categorization.

Data availability statement. The dataset analyzed during the current study is available at the following link: <https://osf.io/8eg2v/files/sv8cy> Videos depicting one trial of each of the two categorization tasks are available at the following link: https://osf.io/8b4c6/overview?view_only=16ecd0c689b4b62889cb6d0996caba6

Competing interests. The author(s) declare none.

References

- Ada, O. (2018). *Bug-like creatures*. https://osf.io/z78w5/?view_only=d0de51d469bb487fb3fbb896d365c53d
- Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A cross-linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37, 286–309.
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., Riches, N., & Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of instruction. *Psychological Science*, 26, 518–526.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2).
- Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189–208.
- Beavers, J., Levin, B., & Wei Tham, S. (2010). The typology of motion expressions revisited. *Journal of Linguistics*, 46(2), 331–377. [10.1017/S0022226709990272](https://doi.org/10.1017/S0022226709990272).
- Bylund, E., & Athanasopoulos, P. (2013). Language and thought in a multilingual context: The case of isiXhosa. *Bilingualism: Language and Cognition*, 17(2), 431–441. <https://doi.org/10.1017/s1366728913000503>.
- Bylund, E., Athanasopoulos, P., & Oostendorp, M. (2013). Motion event cognition and grammatical aspect: Evidence from Afrikaans. *Linguistics*, 51(5), 929–955. <https://doi.org/10.1515/ling-2013-0033>.
- Cardini, F. (2010). Evidence against Whorfian effects in motion conceptualisation. *Journal of Pragmatics*, 42, 1442–1459.
- Casasanto, D. (2008). Who's afraid of the big bad Whorf? Crosslinguistic differences in temporal language and thought. *Language Learning*, 58(1), 63–79.

- Czechowska, N., & Ewert, A. (2011). Perception of motion by Polish-English bilinguals. In V. Cook & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 287–314). Psychology Press.
- Eminoğlu, F. H. (2019). *Supervised categorization task* [Animated game using HTML, CSS, and Javascript]. https://osf.io/uzvbg/?view_only=16ecdd0c689b4b62889cb6d0996caba6
- Feinmann, D. (2020). Language and thought in the motion domain: Methodological consideration and new empirical evidence. *Journal of Psycholinguistic Research*, 49, 1–29. <https://doi.org/10.1007/s10936-019-09668-5>.
- 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, 466–485.
- Finkbeiner, M., Nicol, J., Greth, D., & Nakamura, K. (2002). The role of language in memory for actions. *Journal of Psycholinguistic Research*, 31, 447–457.
- Gennari, S., Sloman, S., Malt, B., & Fitch, W. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79.
- Hohenstein, J. M. (2005). Language-related motion event similarities in English- and Spanish-speaking children. *Journal of Cognition and Development*, 6, 403–425.
- Kersten, A. W. (1998). An examination of the distinction between nouns and verbs: Associations with two different kinds of motion. *Memory & Cognition*, 26, 1214–1232.
- Kersten, A. W., & Billman, D. O. (1997). Event category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 638–658.
- Kersten, A. W., Lechuga, J., Albrechtsen, J., Meissner, C. A., Schwartz, B. L., & 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, 638–653.
- Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed effects models in second language research. *Language Learning*, 65, 185–207. <https://doi.org/10.1111/lang.12117>.
- Lupyan, G. (2012). Linguistically modulated perception and cognition: The label- feedback hypothesis. *Frontiers in Psychology*, 3, 1–13. <https://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. <https://doi.org/10.1177/0963721415570732>.
- Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science*, 18(12), 1077–1083. <https://doi.org/10.1111/j.1467-9280.2007.02028.x>.
- Lupyan, G., & Thompson-Schill, S. L. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General*, 141 (1), 170–186. <https://doi.org/10.1037/a0024904>.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967. [https://doi.org/10.1044/1092-4388\(2007\)067](https://doi.org/10.1044/1092-4388(2007)067).
- Montero-Melis, G., & Bylund, E. (2016). Getting the ball rolling: The cross-linguistic conceptualization of caused motion. *Language and Cognition*, 9(3), 446–472. <https://doi.org/10.1017/langcog.2016.22>.
- Özçalışkan, Ş. (2005). Metaphor meets typology: Ways of moving metaphorically in English and Turkish. *Cognitive Linguistics*, 16(1), 207–246.
- Özçalışkan, Ş., & Slobin, D. I. (2003). Codability effects on the expression of manner of motion in English and Turkish. In A. S. Özsoy, D. Akar, M. Nakipoglu-Demiralp, E. Erguvanlı-Taylan, & A. Aksu-Koç (Eds.), *Studies in Turkish linguistics* (pp. 259–270). Bogaziçi University Press.
- Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures? Insights from crosslinguistic variations and similarities. *Gesture*, 5, 219–240.
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108, 155–184.
- Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, ‘n’ roll: The representation of motion in language and cognition. *Cognition*, 84, 189–219.
- Papafragou, A., & Selimis, S. (2010). Event categorisation and language: A cross- linguistic study of motion. *Language and Cognitive Processes*, 25, 224–260.
- Quick Placement Test. (2001). Oxford University Press.

- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sato, S., & Athanasopoulos, P. (2018). Grammatical gender affects gender perception: Evidence for the structural-feedback hypothesis. *Cognition*, 176, 220–231. <https://doi.org/10.1016/j.cognition.2018.03.014>.
- Skordos, D., Bungler, A., Richards, C., Selimis, S., Trueswell, J., & Papafragou, A. (2020). Motion verbs and memory for motion events. *Cognitive Neuropsychology*, 37(5–6), 254–270. <https://doi.org/10.1080/02643294.2019.1685480>.
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In L. Verhoeven & S. Stromqvist (Eds.), *Relating events in narrative: Typological and contextual perspectives* (Vol. 2, 1st ed., pp. 219–257). Psychology Press. <https://doi.org/10.4324/9781410609694>.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7).
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description: Volume III, grammatical categories and the Lexicon* (pp. 57–149). Cambridge University Press.
- Talmy, L. (1991). *Path to realization: A typology of event integration*. Berkeley Linguistics Society.
- Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring* (Vol. 3). MIT Press.
- Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations. *Language Learning*, 66, 554–580. <https://doi.org/10.1111/lang.12188>.
- Ünal, E., Manhardt, F., & Özyürek, A. (2022). Speaking and gesturing guide event perception during message conceptualization: Evidence from eye movements. *Cognition*, 225, 105127. <https://doi.org/10.1016/j.cognition.2022.105127>.
- Vanek, N., & Fu, X. (2022). Low-level visual processing of motion events as a window into language-specific effects on perception. *International Review of Applied Linguistics in Language Teaching*, 61(1), 61–78. <https://doi.org/10.1515/iral-2022-0048>.
- Zlatev, J., & Yangklang, P. (2004). A third way to travel: The place of Thai in motion event typology. In S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative, vol. 2: Typological and contextual perspectives* (pp. 159–190). Lawrence Erlbaum Associates.