Language structure shapes visual cognition: the effect of zoom-in vs. zoom-out presentation on visual preferences

Yoko Sugitani[1]*, Taku Togawa[1] and Minoru Karasawa[2]

[1]Department of Management, Faculty of Economics, Sophia University, Tokyo, Japan, [2]Future Value Creation Research Center, Graduate School of Informatics, Nagoya University, Nagoya, Japan

Introduction: This study provides initial evidence that grammatical structure in language can shape cognitive preferences for sequential visual stimuli. Linguists classify languages as head-initial or head-final based on their syntactic headedness. Building on this typology, we propose two cognitive styles: head-initial or "zoom-out" cognition, which tends to process more specific, detailed information before focusing on broader perspectives, and head-final or "zoom-in" cognition, which focuses on information from comprehensive to specific. We hypothesized that people's cognitive styles (zoom-in vs. zoom-out) are contingent on their language type (zoom-in or zoom-out language), which determines their cognitive preferences for the order of sequential visual stimuli.

Methods: We conducted three experiments ($N = 823$) involving speakers of zoom-in and zoom-out languages to test our hypotheses using a single sequential visual item and questionnaire-based assessments of processing fluency. This design allowed us to isolate the cognitive effect while acknowledging limits on generalizability.

Results: Across studies, speakers of zoom-in (zoom-out) languages tended to experience higher processing fluency and more favorable evaluations when sequential visual stimuli were presented in a zoom-in (zoom-out) style.

Discussion: These findings offer preliminary evidence linking language structure to visual cognition and highlight opportunities for future research on cross-linguistic variation in cognitive style.

KEYWORDS

cognitive style, culture, grammatical structure, language, processing fluency, visual stimuli

1 Introduction

In visual presentations of objects, such as television commercials for consumer products, contrasting temporal sequences in the delivery of information are frequently observed: one that emphasizes detailed, specific features first and shifts to broader, global aspects of the product, and another that follows the reverse order. For instance, an advertisement for a new wristwatch may initially highlight the stylish design of the dial and hands, or alternatively, begin with the overall appearance of the watch. In the present study, we define the approach that progresses from specific, lower-level details to more abstract, higher-level features as a "zoom-out" presentation and the opposite approach, overall to detailed, as a "zoom-in" presentation. This contrast raises an empirical question: which sequence is more effective in attracting viewer attention?

Among the numerous potential factors, we propose in the present study that grammatical structures of language are crucial in determining whether a particular visual sequence aligns or misaligns with consumer preferences. This contention is based on the ample empirical evidence concerning the effect of language on cognitive tendencies (for reviews, see Boroditsky et al., 2003; Lupyan et al., 2020; Regier and Kay, 2009; Wolff and Holmes, 2011). Literature in psycholinguistics has demonstrated that language influences various aspects of cognition, including memory (Fausey and Boroditsky, 2011), color perception (Winawer et al., 2007), and categorization (Lucy, 1992).

A unique feature of the present study lies in its focus on the influence of grammatical structures of languages to direct primary attention either to isolated elements or to the overall context. Specifically, we examined whether language use results in the ease of processing perceptual information when presented in a manner congruent with the corresponding attentional orientation. Previous research has demonstrated that patterns of language use are closely associated with culturally grounded cognition (Kashima and Kashima, 1998; Bettinsoli et al., 2015). For example, in the case of analytic vs. holistic cognition, speakers of Western languages tend to describe individuals using context-independent trait adjectives that remain stable across time and situations, reflecting their analytic orientation. By contrast, speakers of East Asian languages are more likely to focus on situations and employ context-bounded descriptions, which reflects their holistic cognitive tendency (Maass et al., 2006). Moreover, one study has shown that priming individuals into either analytic or holistic cultural framing can elicit language patterns consistent with that orientation (Morris and Mok, 2011). Building on this body of work, the present study investigated whether language expressions highlighting either parts or the whole facilitate the processing of information structured in a manner that aligns with those linguistic cues.

Although numerous studies have examined the relationship between language and visual perception, no study has examined how language influences cognitive preferences regarding *the order in which sequential visual stimuli are presented*. Specifically, whether people tend to adopt zoom-in or zoom-out cognitive styles depending on activated language remains unknown. It is important to fill this gap because the visual stimuli people encounter daily, such as television and videos posted on social media, are not static but dynamic—constantly changing and in motion. Further research is needed to deepen the understanding of the relationship between language and cognitive tendencies in the processing of dynamic visual stimuli. Moreover, although previous studies have shown that Westerners and East Asians are generally characterized by analytic and holistic attentional styles, respectively (e.g., Masuda and Nisbett, 2001; Nisbett et al., 2001), these findings do not suggest that Westerners focus exclusively on focal objects at the expense of contextual information or that East Asians entirely disregard focal objects. This study helps fill a gap in the psychology literature by examining the processing of sequential and dynamic visual stimuli to explore cultural influences on attentional tendencies that cannot be fully explained by the analytic–holistic dichotomy.

This study contributes to the literature in three important ways. First, it introduces two cognitive styles in sequential visual processing, zoom-in and zoom-out, that are shaped by language spoken. Second, it extends existing frameworks, such as the analytic–holistic distinction, by incorporating the dimension of temporal sequencing in dynamic visual contexts, offering a more nuanced understanding of attentional styles. Third, the findings provide practical insights for designing culturally congruent visual communications in fields such as advertising and media in which the order of visual information presentation is critical.

2 Theoretical background

2.1 Language and cognition

Language is at the core of human culture (Schmitt et al., 1994; Lee et al., 2010), and the controversy as to whether language can shape cognitions has been long ongoing (Roberson et al., 2000; Goldin-Meadow et al., 2008; Chen et al., 2014). The Sapir–Whorf hypothesis (Whorf, 1956) assumes that the structure of a language affects its speakers' world view or cognition, whereas the generative grammar theory (Chomsky, 1975) dismisses this idea by emphasizing linguistic universality or universal grammar.

More recent psychological studies have found strong support for the effect of language on cognition by examining the roles of language structure and grammar (Skerrett, 2010; Choi et al., 2016; Altarriba and Basnight-Brown, 2022). For example, Boroditsky (2000) proposed the *Metaphoric Structuring View*, demonstrating that people understand and reason about abstract domains such as time through metaphorical mappings from more concrete domains like space. Maass and Russo (2003) also found that spatial representations are biased in a direction consistent with the writing system of the perceiver's primary language. Specifically, they demonstrated that the direction of writing influences the processing fluency of visual stimuli: a left-to-right bias was observed in Italian speakers whose language is written from left to right but not in Arabic speakers whose language is written from right to left. Furthermore, Kashima and Kashima (1998) found that the "pronoun drop" of languages affects thinking style. Some languages, such as English, always require the use of personal pronouns ("I" and "You") in sentences, while other languages, such as Japanese and Spanish, often drop these pronouns. They argued that in the "pronoun-dropping" languages, the person or subject of the action is less emphasized in the sentence than in languages that require pronouns. Accordingly, people in languages that require personal pronouns tend to have more individualistic thinking than those in pronoun-dropping languages (Kashima and Kashima, 1998). Moreover, other linguistic features, such as the presence of linguistic hypotheticals (Bloom, 1981), differences in classifiers (Zhang and Schmitt, 1998), ideograms or phonograms (Schmitt et al., 1994), and articulation moving inward (Ingendahl et al., 2021), have been identified as influencing speakers' cognitive features and preferences.

2.2 Head-initial and head-final language

In linguistics, languages worldwide are generally classified into two types based on their grammatical structure: the head-initial or subject–verb–object type and the head-final or subject–object–verb type (Greenberg, 1963; Dryer, 1992; Dunn et al., 2011). Head-initial

languages, including English, French, Spanish, and German, tend to have specific and important information (e.g., verbs) before the dependent sentence elements, such as objects and complements. By contrast, in head-final languages, such as Japanese, Turkish, and Korean, specific information comes after the dependent contents in sentences. In a single sentence, the verb is the specific and decisive part that determines the meaning of the entire sentence (MacDonald et al., 1994). Thus, the position of the verb, whether it comes earlier or later in the sentence, defines the language as head-initial or head-final.

The following is a sample sentence with English as the head-initial language.

"I visited a newly opened shopping mall located about a 5-minute walk from Tokyo Station yesterday."

In English, the verb "visited" appears at the beginning of the sentence, narrowing the scope of the meaning of subsequent objects ("shopping mall") and other modifiers. However, in Japanese, a head-final language, the verb "visited" comes after "a shopping mall" and "Tokyo Station," leaving the meaning of the object to the end of the sentence. The typical order of the Japanese sentence is "I, yesterday, Tokyo Station, walk from 5-minute located, newly open, shopping mall, visited."

Another example is writing residential addresses. In most countries with head-initial languages, the structure of a written address flows from specific to global perspectives (e.g., "51 Kings Ave., Columbus, OH, U.S."), whereas in countries with head-final languages, the reverse is used (global to specific information, e.g., "U.S. OH, Columbus, Kings Ave., 51").

Accordingly, in head-final languages, the specific and detailed meaning of words in a sentence are characteristically reserved until the final part. By contrast, head-initial languages tend to focus on more specific and detailed elements at the beginning of a sentence, as verbs are presented immediately after the subject (MacDonald et al., 1994).

In this study, we posited that the attentional sequence of head-initial and head-final languages correspond to the zoom-out (i.e., specific to broader) and zoom-in (i.e., broader to specific) orders of visual presentation, respectively.

Notably, studies differ in their approaches to understanding the headedness of languages. For example, some approaches analyze headedness in terms of parameters based on universal structures (Cinque, 2013, 2017) or use machine learning methods to quantify the degree of headedness in a language (Alves et al., 2022). These approaches differ from the binary typology of head-initial vs. head-final proposed by Dryer (1992). Additionally, some studies suggest that the noun–verb ratio varies depending on a language's headedness (Polinsky and Magyar, 2020). The present study adopts Dryer's typology, which focuses on core syntactic structure, specifically, the positional relationship between subjects, verbs, and objects (subject–verb–object or subject–object–verb), rather than on a more comprehensive notion of headedness that also encompasses modifiers such as adjectives and adverbs.

2.3 Zoom-in and zoom-out cognitive style

Because the sequential transition in head-initial (vs. head-final) language, from more specific to broader information, aligns with
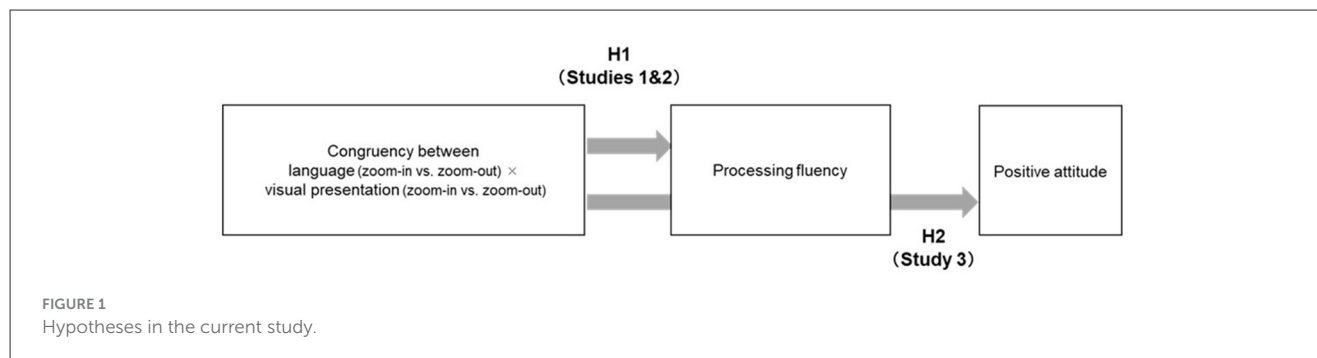
the zoom-out (vs. zoom-in) presentation of stimulus information, we predicted that a match between sequential transitions in languages and visual presentation styles would enhance processing fluency, which in turn would increase cognitive preference.

According to the event knowledge framework (Hare et al., 2009; Elman and McRae, 2019), people acquire "event knowledge" or schemas based on their learning and experiences, which allows them to predict what is likely to happen next and to comprehend their surroundings fluently. In this study, the cognitive tendencies of zoom-in and zoom-out can be understood as a type of event knowledge. We propose that people develop either zoom-in or zoom-out types of event knowledge through the linguistic structures of the language they use in daily communication, and that they perceive greater processing fluency when visual information is presented in a way that is consistent with their own event knowledge.

Human visual processing is strongly influenced by multiple factors, including the nature of the task stimuli, prior experience, and contextual cues. For example, identifying an object becomes easier when one is familiar with the object or has previously encountered it (Maljkovic and Nakayama, 1994; Kristjánsson and Campana, 2010), which is known as *priming*. Research in psycholinguistics has shown that priming can enhance the processing fluency of subsequent linguistic input. For instance, exposure to a sentence in the passive voice increases the likelihood of using the passive voice in subsequent speech, as prior exposure activates specific grammatical structures (structural priming; Van Gompel and Arai, 2018). Amici et al. (2019) also demonstrated that linguistic word order affects non-linguistic working memory, with speakers of left-branching languages remembering early items better and those of right-branching languages recalling later items more accurately. Hence, the linguistic structure of head-initial or head-final languages may be primed when using them to speak, influencing the processing fluency of subsequent visual information. Specifically, when the structure of a head-initial language is primed, attention tends to shift from specific to global information (i.e., zoom-out), whereas when the structure of a head-final language is primed, attention tends to shift from global to specific information (i.e., zoom-in). Thus, processing fluency is likely enhanced when the visual presentation style (zoom-in vs. zoom-out) aligns with the attentional shift associated with these language structures.

Neuroscientific evidence also supports this view, showing that speaking head-initial and head-final languages activates different parts of the brain (Kemmerer, 2012). Therefore, speakers of head-initial or zoom-out languages are assumed to regularly adhere to zoom-out cognition style, whereas speakers of head-final or zoom-in languages could be encouraged to follow zoom-in cognition style. Based on this argument, we hypothesized that zoom-in (zoom-out) language speakers have a zoom-in (zoom-out) cognitive style, resulting in higher processing fluency toward visual stimuli presented in zoom-in (zoom-out) order.

Furthermore, we tested whether the processing fluency of sequential visual stimuli enhances attitudes toward the presented objects. Specifically, we examined the effect of the match between language type (zoom-in vs. zoom-out) and visual presentation (zoom-in vs. zoom-out) on the formation of positive attitudes. Processing fluency refers to the subjective experience of ease or difficulty with which information is processed (Reber et al., 2004).

FIGURE 1
Hypotheses in the current study.

Known as the mere exposure effect, the subjective experience of processing fluency can enhance favorable attitudes (Zajonc, 1968; Fang et al., 2007). Thus, we further hypothesized that the congruency of language (zoom-in vs. zoom-out) and visual presentation (zoom-in vs. zoom-out) would improve participants' attitudes toward the object described in the visual stimuli, as a downstream effect of processing fluency.

Hypothesis 1 (H1): People who speak head-initial or zoom-out languages will experience higher processing fluency for zoom-out presentations, whereas those who speak head-final or zoom-in languages will experience higher processing fluency for zoom-in presentations.

Hypothesis 2 (H2): The congruence between language type and visual presentation will lead to more favorable attitudes toward presented objects, mediated by increased processing fluency.

3 Study overview

Three experiments were conducted to test these hypotheses (Figure 1). Studies 1 and 2 were designed to examine H1, showing that speakers of zoom-in (or zoom-out) languages reported higher processing fluency in zoom-in (or zoom-out) presentations. In Study 3, as predicted in H2, we demonstrated that people showed more favorable attitudes toward a product presented in a zoom-in or zoom-out style matching their language type.

Furthermore, in all experiments, we examined the possible confounding effect of holistic–analytic cognitive tendencies (Masuda and Nisbett, 2001) and found a non-significant effect on the predicted results (see Supplementary Material for details).

These studies were approved by the Ethics Committee on Research on Human Subjects of Sophia University (approval number: 2018-113). All participants provided written informed consent after being informed of the study's objectives and assured of the confidentiality of their data.

4 Study 1

To test H1, we compared the processing fluency of sequential visual stimuli among participants whose first language was either a zoom-in or zoom-out language. To minimize potential confounds related to cultural or national background, we recruited participants from a variety of home countries within each language group.

4.1 Methods

4.1.1 Participants and design
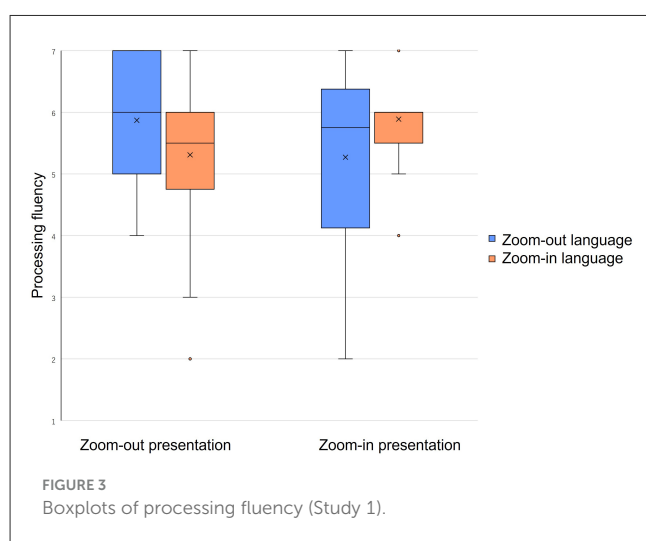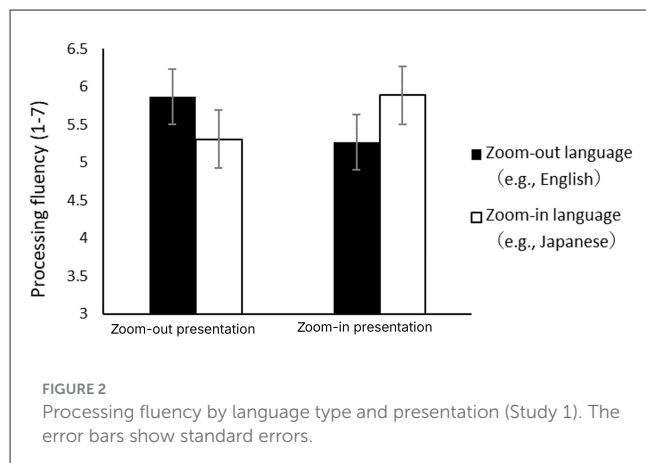
We employed a 2 (language type: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) between-participants design. In total, 86 undergraduate students (M_{age} = 20.41, SD_{age} = 2.60; 28 men, 58 women, and 0 others) participated in the experiment. They were international students from various countries residing in Japan and pursuing their studies in English at a Japanese university. They included 48 students whose first language was a zoom-out language (Arabic, Indonesian, English, French, German, Thai, Chinese, Filipino, or Urdu), and 38 students whose language was a zoom-in language (Japanese, Korean, or Turkish). Further details regarding the sample languages can be found in Section 1.1 of the Supplementary Material.

4.1.2 Procedure

Participants were asked to visit the assigned page on their smartphones and view a series of three photographs of a wristwatch, presented in either a zoom-in or zoom-out order (see Section 1.2 of the Supplementary Material for details). They then responded to measures of processing fluency using two items ("easy to see" and "easy to understand"; Lee and Aaker, 2004), rated on a 7-point scale. The scores were averaged to form the processing fluency (r = 0.76).

4.2 Results

We conducted a 2 (language type: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) analysis of variance (ANOVA) on perceived fluency in information processing revealed that neither the main effect for the language type nor for the visual presentation was significant ($Fs < 1$). Importantly, however, a significant interaction was found ($F(1, 82) = 4.97$, $p = 0.03$, $\eta_p^2 = 0.06$). As Figure 2 illustrates, participants whose first language was zoom-out reported a greater processing fluency for zoom-out presentation ($M = 5.87$, $SD = 1.13$) than for zoom-in presentation ($M = 5.27$, $SD = 1.50$) ($F(1, 82) = 2.84$, $p = 0.096$, $\eta_p^2 = 0.03$). By contrast, participants whose native language was zoom-in did not show the significant difference in processing fluency for zoom-in ($M = 5.89$, $SD = 0.74$) and zoom-out presentation ($M = 5.31$, $SD = 1.28$) ($F(1, 82) = 2.19$, $p = 0.14$, $\eta_p^2 = 0.03$). Boxplots illustrating

FIGURE 2
Processing fluency by language type and presentation (Study 1). The error bars show standard errors.

FIGURE 3
Boxplots of processing fluency (Study 1).

the distribution and variability of processing fluency are provided in Figure 3.

4.3 Discussion

Study 1 showed that the grammatical structure of the native language (zoom-in vs. zoom-out) affects the processing fluency of visual elements, such as sequential pictures, supporting H1. An important next question is whether the observed effect reflects a stable trait associated with each type of language or a temporary state that can be induced by language use.

Prior research on multilinguals has demonstrated that an individual's personality or identity can shift depending on the language they speak (Ramírez-Esparza et al., 2006; Lee et al., 2010; Doucerain et al., 2023). For example, Ramírez-Esparza et al. (2006) found that Spanish–English bilinguals exhibited different levels of extraversion, agreeableness, and conscientiousness when switching languages, reflecting the cultural norms associated with each language. To explore the possibility that the difference between zoom-in and zoom-out languages reflects dynamic cognitive processes rather than static traits, we applied a linguistic priming procedure with multilingual participants and examined the effect

of the two types of visual sequences on perceived fluency observed in Study 1.

5 Study 2

Study 2 expanded on Study 1 with the following additional considerations. First, instead of examining the effect of participants' native language, Study 2 investigated how the *language spoken in the setting* influenced the processing fluency of visual stimuli. Specifically, we recruited undergraduate students who spoke both zoom-in (e.g., Japanese) and zoom-out (e.g., English) languages. Second, in Study 1, participants were presented with zoom-in or zoom-out visual stimuli without any additional context. To enhance the ecological validity of Study 2, the participants were exposed to visual stimuli alongside contextual information. Specifically, they were given the following instruction: "You are reading an article on a website that refers to a nice café."

5.1 Methods

5.1.1 Experimental design

We employed a 2 (language *primed*: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) between-participants design.
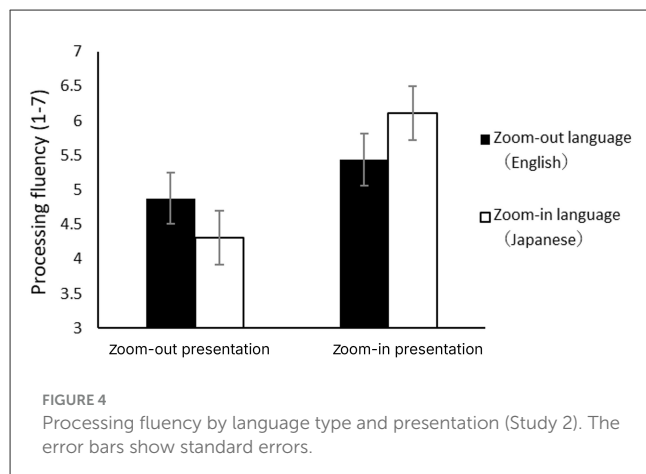
5.1.2 Participants

A total of 115 undergraduate students (M_{age} = 20.56, SD_{age} = 3.09; 51 men, 64 women and 0 others) who could speak both Japanese and English participated in the experiment. Their mother tongues were Japanese (n = 106), Chinese (n = 8), and English (n = 1).

5.1.3 Procedure

The experiment comprised two stages, and the participants were instructed to participate in two independent experiments. In the first section, participants were randomly assigned to either zoom-in or zoom-out language condition and asked to engage in the sentence completion task, which required them to complete sentences by rearranging the order of words from a disjointed word list. Participants in the zoom-in (vs. zoom-out) language condition were given a task in Japanese (vs. English) to prime the zoom-in (vs. zoom-out) language structure (see Section 2.1 of the Supplementary Material for details).

In the second section, participants were asked to use their smartphones to read an article on a fictitious website featuring a café. The website article consisted of three pages, each featuring a photograph: one depicting the café's exterior, one depicting its interior, and one depicting a hamburger served there. Half of the participants were presented with the three photos in a zoom-in order, whereas the other half viewed them in a zoom-out order (see Section 2.2 of Supplementary Material).

Finally, they completed the processing fluency scale (r = 0.75) as in Study 1. We asked them about their familiarity with cafés, as

FIGURE 4
Processing fluency by language type and presentation (Study 2). The error bars show standard errors.

well as holistic and analytic thinking tendencies. These results are reported in Sections 2.3 and 2.4 of Supplementary Material.

5.2 Results

We conducted a 2 (language type: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) ANOVA on processing fluency. The results showed a significant main effect of visual presentation (F (1, 111) = 19.49, $p < 0.001$, $\eta_p^2 = 0.15$). Processing fluency was higher for zoom-in ($M = 5.82$, $SD = 1.35$) than for zoom-out presentations ($M = 4.58$, $SD = 1.53$). This result could have been obtained because most participants were native Japanese speakers (i.e., zoom-in language). The effect of *native (zoom-in) language* might still be salient as the baseline, even when zoom-out language was primed.

Furthermore, as predicted, the results also revealed a significant interaction effect (F (1, 111) = 5.36, $p = 0.02$, $\eta_p^2 = 0.05$; Figure 4). Participants in the zoom-in language (Japanese) condition reported higher processing fluency for the zoom-in presentation ($M = 6.11$, $SD = 1.17$), compared with the zoom-out presentation ($M = 4.31$, $SD = 1.63$) (F (1, 111) = 25.12, $p < 0.001$, $\eta_p^2 = 0.19$). By contrast, participants in the zoom-out language (English) condition showed no significant difference in processing fluency between the zoom-out ($M = 4.88$, $SD = 1.39$) and zoom-in ($M = 5.44$, $SD = 1.51$) presentations (F (1, 111) = 2.01, $p = 0.16$, $\eta_p^2 = 0.02$). Boxplots illustrating the distribution and variability of processing fluency are provided in Supplementary Figure 5.

5.3 Discussion

Study 2 revealed that the structure of the language (zoom-in or zoom-out) *activated at the moment* influences processing fluency. This finding indicates that the observed effect of language on the processing fluency of visual stimuli was not solely attributed to their native languages but was also influenced by the temporarily primed language. Additionally, it rules out the alternative explanation that our results were driven by cultural values or typical advertisement

practices in the participants' country of residence rather than by language.

6 Study 3

Study 3 tested H2 by examining the attitudes toward objects presented in a zoom-in or zoom-out manner as a downstream effect of processing fluency. Furthermore, we conducted the experiment in a more diverse set of countries, recruiting British, German, French, and Chinese participants as zoom-out language speakers, and Japanese and Korean participants as zoom-in language speakers.

6.1 Methods

6.1.1 Experimental design

We employed a 2 (language type: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) between-participants design.

6.1.2 Participants

The experiment was conducted online with 622 participants ($M_{age} = 41.52$, $SD_{age} = 11.16$, 285 men, 336 women, and one other) residing in the UK ($n = 102$) recruited from Prolific and residing in Germany ($n = 104$), France ($n = 104$), China ($n = 104$), Korea ($n = 104$), and Japan ($n = 104$) from the survey panels of Macromill Inc. According to linguistic definitions (Dryer and Haspelmath, 2013), English, German, French, and Chinese were grouped as zoom-out languages, whereas Korean and Japanese were categorized as zoom-in languages.

6.1.3 Procedure

Participants watched a 10-s video showing a wristwatch in either a zoom-in or zoom-out manner (see Section 3.1 of Supplementary Material for details). After viewing the video, they completed scales measuring processing fluency and attitude toward the object (three items: favorable, good, and wise; $\alpha = 0.90$; MacKenzie et al., 1986), using 7-point scales.

All the experiments were conducted in the participants' native language. The questionnaires, which were originally written in English, were translated into German, French, Chinese, Korean, or Japanese. Independent proofreaders then reviewed the translations to ensure consistency of meaning.
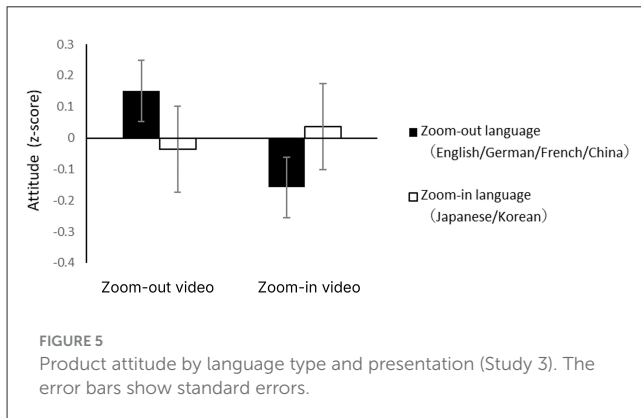
6.2 Results

6.2.1 Attitude toward the object

As cross-cultural studies have found a response bias in that East Asians, including the Japanese, tend to avoid extreme responses compared with North Americans (Chen et al., 1995), z-scores for attitude toward the object were used in the following analysis.

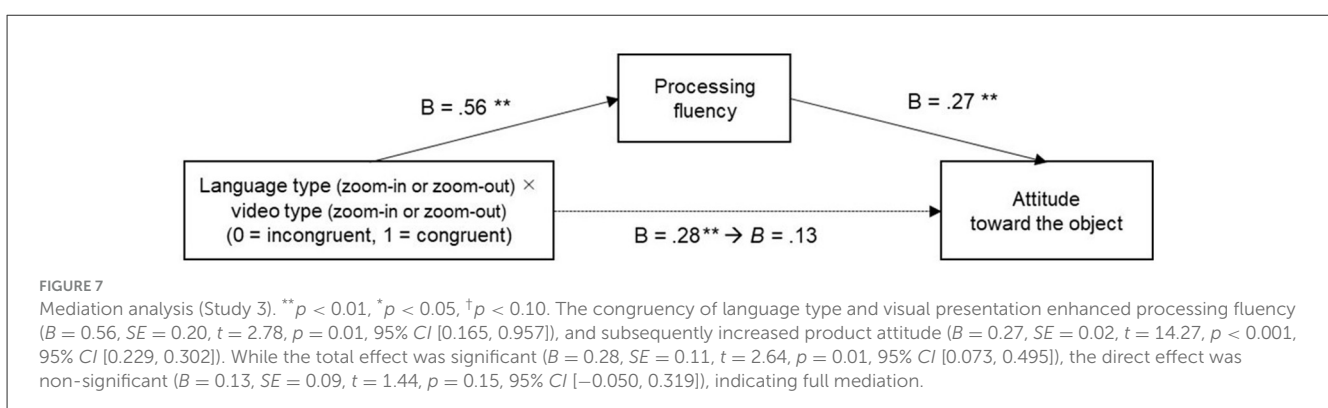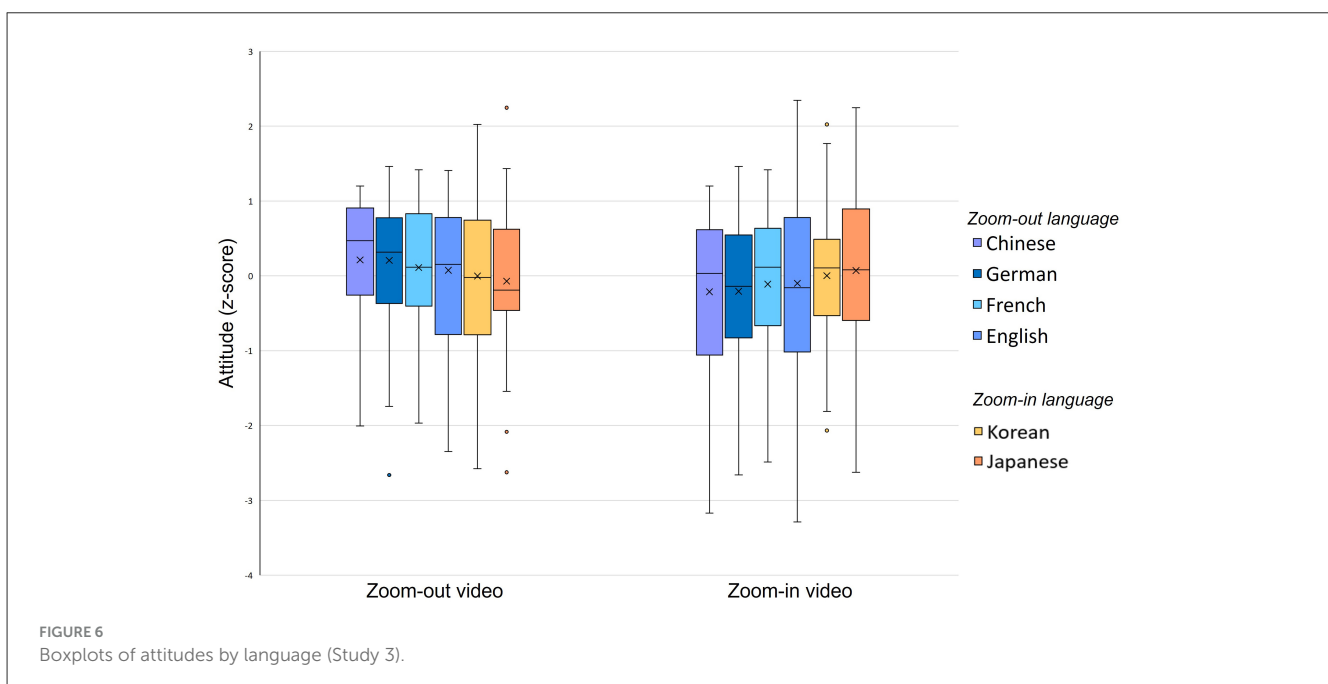We ran a 2 (language type: zoom-in vs. zoom-out) × 2 (visual presentation: zoom-in vs. zoom-out) ANOVA on attitude

toward the object. The results showed a significant interaction effect (F (1, 618) $= 5.09$, $p = 0.02$, $\eta_p^2 = 0.01$) (Figure 5). Zoom-out language speakers showed more positive attitudes toward the object presented in the zoom-out video ($M = 0.15$, $SD = 0.90$) than in the zoom-in video ($M = -0.16$, $SD = 1.07$)

FIGURE 5
Product attitude by language type and presentation (Study 3). The error bars show standard errors.

(F (1, 618) $= 9.98$, $p < 0.001$, $\eta_p^2 = 0.02$). The boxplots in Figure 6 present attitude distributions for each of the six languages (instead of zoom-in/zoom-out categories), highlighting cross-linguistic variability. Additional language-specific analyses are reported in Section 3.3, *Language-Specific Analysis,* of the Supplementary Material.

6.2.2 Mediation analysis

To test H2, we ran a mediation analysis (Hayes, 2022; SPSS PROCESS Model 4, 10,000 bootstrapping resamples) with attitude toward the object as the dependent variable and processing fluency as the mediator. As the independent variable, we created a new index of congruency between language and video types [1 = congruent between language and video types (e.g., zoom-in language and zoom-in video), 0 = incongruent between language and video type (e.g., zoom-in language and zoom-out video)]. The result showed that the indirect effect was significant ($B = 0.15$, $SE = 0.06$, 95% CI [0.044, 0.262]; Figure 7), supporting H2.

FIGURE 6
Boxplots of attitudes by language (Study 3).

FIGURE 7
Mediation analysis (Study 3). $**p < 0.01$, $*p < 0.05$, $†p < 0.10$. The congruency of language type and visual presentation enhanced processing fluency ($B = 0.56$, $SE = 0.20$, $t = 2.78$, $p = 0.01$, 95% CI [0.165, 0.957]), and subsequently increased product attitude ($B = 0.27$, $SE = 0.02$, $t = 14.27$, $p < 0.001$, 95% CI [0.229, 0.302]). While the total effect was significant ($B = 0.28$, $SE = 0.11$, $t = 2.64$, $p = 0.01$, 95% CI [0.073, 0.495]), the direct effect was non-significant ($B = 0.13$, $SE = 0.09$, $t = 1.44$, $p = 0.15$, 95% CI [−0.050, 0.319]), indicating full mediation.

7 General discussion

This study explored how the grammatical structure of a language (head-initial or head-final) influences speakers' cognitive preferences for dynamic and sequential visual stimuli. Through three experiments, we demonstrated that head-initial (vs. head-final) language speakers exhibited higher processing fluency for zoom-out (vs. zoom-in) presentations, resulting in a more positive attitude toward objects presented in the corresponding style. These results have provided support for our proposition that people adopt one of two cognitive styles, zoom-in or zoom-out, depending on their language.

7.1 Theoretical contribution

This study offers a new perspective on how language use influences cognition. First, it proposes the possibility of linguistic categorization, zoom-in (i.e., head-final) and zoom-out (i.e., head-initial), to explain cultural differences in cognitive styles. Since the Sapir–Wolf hypothesis was proposed, many scholars have investigated the language relativity of cognition (Goldin-Meadow et al., 2008; Nicoladis and Foursha-Stevenson, 2012; Meir et al., 2017; Altarriba and Basnight-Brown, 2022; Doucerain et al., 2023). While some of these studies categorized languages according to prominent cultural studies such as Hofstede's national culture (Hofstede et al., 2010) and cultural self-construals (Markus and Kitayama, 1991), others have focused on differences in language grammar, such as pronoun drop (Kashima and Kashima, 1998), writing direction (Maass and Russo, 2003), human/animacy entities (Meir et al., 2017), classifiers (Zhang and Schmitt, 1998), and ideograms/phonograms (Schmitt et al., 1994). However, few studies have addressed word order in grammar (Meir et al., 2017). The present study sheds light on the linguistic typology of head-initial and head-final languages and proposed that these two languages would align with zoom-out and zoom-in visual presentations, respectively. To the best of our knowledge, this was the first attempt to use this linguistic typology to discuss differences in the cognitive processing of sequential visual stimuli. This framework of "zoom-in and zoom-out congruency" could be applied to explain various cultural differences in other important domains, such as making effective messages in persuasive communication, global advertising strategy in marketing, and efficient team building in organizational behavior research.

Second, we proposed a new perspective on the zoom-in and zoom-out cognitive styles, which includes preferences for the temporal transition of visual stimuli. Prior research on languages and cognition has examined various cognitive variables such as memory (Amici et al., 2019; Fausey and Boroditsky, 2011; Schmitt et al., 1994), categorization (Zhang and Schmitt, 1998), color perception (Roberson et al., 2000; Winawer et al., 2007), left-to-right bias (Maass and Russo, 2003), and in–out effect (Ingendahl et al., 2021), as well as social cognition including extroversion (Ramírez-Esparza et al., 2006), self-enhancement (Lee et al., 2010), identification and subjective value (Doucerain et al., 2023), and individualism (Kashima and Kashima, 1998). However, all these studies examined cognition *at a single point in time.* Although

people are often exposed to dynamic and sequential images, no previous research has addressed which images people prefer to view first and in what sequence. Thus, the current study focuses on the *transition of stimuli*, specifically the order in which visual stimuli are presented in a zoom-in or zoom-out manner.

Although this argument is similar to the distinction between holistic and analytic thinking styles (Masuda and Nisbett, 2001), such as classifying people's attention based on global or specific perspectives, the two have distinct focuses. First, research on holistic and analytic thinking has examined cultural differences in focal attention in relation to cultural values (collectivism vs. individualism); by contrast, the current study addresses grammatical structures in language. Second, as previously mentioned, holistic and analytic thinking focuses on focal attention within a single image, whereas our study examined the effect of culture (language) on processing videos or multiple stimuli presented sequentially. To rule out the possibility that our findings could be explained by differences in holistic vs. analytic thinking rather than zoom-in and zoom-out cognition, we measured the participants' holistic and analytic thinking tendencies using the locus of attention scale in all three experiments, thereby eliminating this alternative explanation (see Sections 1.3, 2.3, and 3.2 in the Supplementary Material).

Third, our study demonstrated that the relationship between language and cognition is not solely determined by one's native language. The results of Study 2 indicated that the cognitive processing of visual elements in multilinguals is also influenced by the language they speak at a given moment (i.e., the primed language). Given that previous research has primarily addressed this relationship in the context of native languages (e.g., Schmitt et al., 1994), our study provides new insights into the relativity of cognition in relation to primed language.

Finally, this study provides novel evidence that aligns with prior neuroscientific findings on the relationship between language and the cognitive processing of visual elements. Roberson et al. (2000) argued that whether the brain can distinguish certain colors depends on whether the language in which the person speaks differentiates them. Our findings have clarified which visual elements are more likely to be processed first, depending on the language being spoken, supporting the neuroscientific perspective on the relativity of visual processing to language.

Thus, this study provides new evidence in the historical debate surrounding the Sapir–Whorf hypothesis, which posits that language shapes how we see the world (Whorf, 1956).

7.2 Limitations and future directions

This study had some limitations. First, the effect sizes obtained from the three experiments were not large ($\eta_p^2 = 0.01$–0.19), which limits their generalizability. Our findings suggest that speakers of head-initial (head-final) languages may not consistently prefer zoom-out (zoom-in) visual presentations. One possibility is that the participants' evaluations of visual stimuli, such as videos and photos, are influenced not only by their dominant language but also by their goals and motivations. For example, people may prefer zoom-in presentations when viewing large landscapes,

whereas zoom-out presentations may be consistently favored when photographing small, complex objects such as food or stationery. Thus, preferences can vary depending on the content of the visual stimuli. Additionally, this effect may be situational. For instance, in emergencies, even speakers of zoom-in languages might shout "Run away!", which aligns with the zoom-out perspective. Therefore, we acknowledge that the priming effect of language does not always dominate everyday life.

Second, we should remain cautious when interpreting the extent to which the study results support our hypotheses. In Study 3, the interaction effect between visual presentation and language type reached significance; however, the simple main effect of visual presentation was not significant for head-final language speakers. This pattern may reflect preference flexibility among head-final language speakers, whereas head-initial language speakers may exhibit a more directional preference. Therefore, based solely on the present findings, we cannot conclude that our hypotheses are fully supported. Future work would benefit from examining preference outcomes through the lens of flexibility vs. directionality.

Third, the languages included in the present study do not represent an exhaustive set. In terms of speaker population, head-final or zoom-out languages dominate, whereas zoom-in languages are relatively rare (Dryer and Haspelmath, 2013; Murayama et al., 2016). Among the 52 languages with more than 10 million speakers, 33 are of the head-initial type and 19 are of the head-final type (Dryer and Haspelmath, 2013). Furthermore, among the languages with the 10 largest speaker populations, head-initial type languages are dominant: only three languages (Hindi, Bengali, and Japanese) are of the head-final type, whereas the remaining seven (Chinese, English, Russian, Spanish, German, French, and Italian) are all head-initial languages. As a result, we were able to recruit only Japanese and Korean participants as native speakers of the zoom-in languages in this study. Future studies should incorporate participants representing a wider range of language backgrounds. Furthermore, as noted above, headedness can be classified in multiple ways (e.g., Cinque, 2013, 2017; Alves et al., 2022), including Dryer's typology, which we adopted in the present studies. Previous work also indicates that headedness may influence a broader range of syntactic structures (Alves et al., 2022). If our hypotheses are valid, similar patterns should be observable even when adopting alternative typological frameworks. Therefore, to establish more robust evidence that processing fluency increases when visual presentation aligns with headedness, future research should evaluate this effect using multiple classification systems. This need is underscored by a key limitation of the current work, namely, that it tested only a restricted set of languages.

Fourth, we measured processing fluency using self-report items following prior research (Lee and Aaker, 2004), as this approach has been widely used in studies examining preference formation (Graf et al., 2018; Kostyk et al., 2021). Offline, self-report measures are designed to capture the subjective experience of ease in processing information, which reflects a metacognitive judgment regarding how fluently a stimulus is processed by the perceiver (Alter and Oppenheimer, 2009; Schwarz, 2004; Graf et al., 2018). In contrast, online behavioral indices such as reaction time or eye-tracking capture earlier attentional stages of processing and show

only partial convergence with subjective fluency (Kostyk et al., 2021). Prior research further demonstrates that although objective fluency manipulations can affect processing speed, subjective fluency ratings tend to predict evaluative judgments such as liking more strongly than online measures including reaction time (Forster et al., 2013). Taken together, these findings suggest that the effect of language headedness observed in our data may not emerge during initial perceptual encoding but instead at a later evaluative stage in which metacognitive assessments shape preference. However, because we did not collect online behavioral measures, this interpretation remains provisional. Future research incorporating both online and offline metrics will be required to clarify the temporal dynamics through which headedness influences preference.

A further limitation is that each experiment employed a single stimulus item: a watch in Experiments 1 and 3, and a café in Experiment 2. Also, in Experiment 2, the item was presented in one of two languages, and in Experiments 1–3, items were presented with zoomed-in or zoomed-out perspectives. While this design allowed us to rigorously test our hypotheses under tightly controlled experimental conditions, reliance on a single stimulus per experiment limits the generalizability of the findings. Future research should examine a broader range of items to test the robustness and generalizability of the observed effects.

Another consideration is that in Study 2, participants responded in a second language, which may have influenced the degree of affective processing (Costa et al., 2014a,b). While Study 3, in which all participants responded in their first language, showed a consistent association between zoom-in (zoom-out) language and corresponding cognitive style, the potential effects of second-language use warrant further investigation. Future research could examine how using a second language may modulate cognitive and affective responses in similar tasks.

Finally, to extend the present findings and strengthen their generalizability, future work should address the limitations noted above. One priority is to expand the stimulus set beyond a single visual item to evaluate whether the observed pattern generalizes across different scenes and event types. It will also be important to examine more diverse linguistic populations to determine whether the effects hold across cultural and typological contexts. Methodologically, incorporating online processing measures, such as eye-tracking or reaction-time paradigms, may help reveal the temporal dynamics that underlie visual sequence preferences. Furthermore, rather than treating headedness as a binary feature, a gradient conceptualization may allow for more fine-grained predictions and deepen our understanding of how linguistic structure relates to visual cognition.

Beyond this, an important avenue is to investigate whether zoom-in and zoom-out cognitive tendencies extend to domains other than visual processing fluency. For instance, prior work shows linguistic context can modulate personality expression, with speakers demonstrating greater self-enhancement in English than in Chinese (Lee et al., 2010). This raises the question of whether zoom-in or zoom-out cognition may similarly shape other psychological outcomes, such as memory, self-perception, causal attribution, and stereotyping.

Data availability statement

Ethics statement

Author contributions

YS: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition. TT: Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing. MK: Conceptualization, Methodology, Supervision, Writing – review & editing.

Funding

Acknowledgments

Conflict of interest

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/flang.2025.1637387/full#supplementary-material

References

Altarriba, J., and Basnight-Brown, D. (2022). The psychology of communication: the interplay between language and culture through time. *J. Cross Cult. Psychol.* 53, 860–874. doi: 10.1177/00220221221114046

Alter, A. L., and Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Pers. Soc. Psychol. Rev.* 13, 219–235. doi: 10.1177/1088868309341564

Alves, D., Tadić, M., and Bekavac, B. (2022). "Multilingual comparative analysis of deep-learning dependency parsing results using parallel corpora," in *Proceedings of the BUCC Workshop within LREC 2022*, 33–42.

Amici, F., Sánchez-Amaro, A., Sebastián-Enesco, C., Cacchione, T., Allritz, M., Salazar-Bonet, J., et al. (2019). The word order of languages predicts native speakers' working memory. *Sci. Rep.* 9:1124. doi: 10.1038/s41598-018-37654-9

Bettinsoli, M. L., Maass, A., Kashima, Y., and Suitner, C. (2015). Word-order and causal inference: the temporal attribution bias. *J. Exp. Soc. Psychol.* 60, 144–149. doi: 10.1016/j.jesp.2015.05.011

Bloom, A. H. (1981). *The Linguistic Shaping of Thought.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition* 75, 1–28. doi: 10.1016/S0010-0277(99)00073-6

Boroditsky, L., Schmidt, L. A., and Phillips, W. (2003). "Sex, syntax and semantics," in *Language in Mind: Advances in the Study of Language and Thought,* eds. D. Gentner, and S. Goldin-Meadow (Boston Review), 61–79. doi: 10.7551/mitpress/4117.003.0010

Chen, C., Lee, S. Y., and Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychol. Sci.* 6, 170–175. doi: 10.1111/j.1467-9280.1995.tb00327.x

Chen, S. X., Benet-Martínez, V., and Ng, J. C. K. (2014). Does language affect personality perception? A functional approach to testing the Whorfian hypothesis. *J. Pers.* 82, 130–143. doi: 10.1111/jopy.12040

Choi, H., Connor, C. B., Wason, S. E., and Kahan, T. A. (2016). The effects of interdependent and independent priming on Western participants'

ability to perceive changes in visual scenes. *J. Cross Cult. Psychol.* 47, 97–108. doi: 10.1177/0022022115605384

Chomsky, N. (1975). *The Logical Structure of Linguistic Theory*. New York, NY: Springer.

Cinque, G. (2013). Cognition, universal grammar, and typological generalizations. *Lingua* 130, 50–65. doi: 10.1016/j.lingua.2012.10.007

Cinque, G. (2017). A microparametric approach to the head-initial/head-final parameter. *Linguist. Anal.* 41, 309–366.

Costa, A., Foucart, A., Arnon, I., Aparici, M., and Apesteguia, J. (2014a). "Piensa" twice: on the foreign language effect in decision making. *Cognition* 130, 236–254. doi: 10.1016/j.cognition.2013.11.010

Costa, A., Foucart, A., Hayakawa, S., Aparici, M., Apesteguia, J., Heafner, J., et al. (2014b). Your morals depend on language. *PLoS ONE* 9:e94842. doi: 10.1371/journal.pone.0094842

Doucerain, M. M., Medvetskaya, A., Moldoveanu, D., and Ryder, A. G. (2023). Who are you—right now? Cultural orientations and language used as antecedents of situational cultural identification. *J. Cross. Cult. Psychol.* 54, 784–807. doi: 10.1177/00220221231193148

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language (Baltim)* 68, 81–138. doi: 10.1353/lan.1992.0028

Dryer, M. S., and Haspelmath, M. (2013). "Order of subject, object and verb," in *The World Atlas of Language Structures Online*. Oxford: Oxford University Press.

Dunn, M., Greenhill, S. J., Levinson, S. C., and Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature* 473, 79–82. doi: 10.1038/nature09923

Elman, J. L., and McRae, K. (2019). A model of event knowledge. *Psychol. Rev.* 126, 252–291. doi: 10.1037/rev0000133

Fang, X., Singh, S., and Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *J. Consum. Res.* 34, 97–103. doi: 10.1086/513050

Fausey, C. M., and Boroditsky, L. (2011). Who dunnit? Cross-linguistic differences in eye-witness memory. *Psychon. Bull. Rev.* 18, 150–157. doi: 10.3758/s13423-010-0021-5

Forster, M., Leder, H., and Ansorge, U. (2013). It felt fluent, and I liked it: subjective feeling of fluency rather than objective fluency determines liking. *Emotion* 13, 280–289. doi: 10.1037/a0030115

Goldin-Meadow, S., So, W. C., Özyürek, A., and Mylander, C. (2008). The natural order of events: how speakers of different languages represent events nonverbally. *Proc. Natl. Acad. Sci. USA* 105, 9163–9168. doi: 10.1073/pnas.0710060105

Graf, L. K. M., Mayer, S., and Landwehr, J. R. (2018). Measuring processing fluency: one versus five items. *J. Consum. Psychol.* 28, 393–411. doi: 10.1002/jcpy.1021

Greenberg, J. H. (1963). "Some universals of grammar with particular reference to the order of meaningful elements," in *Universals of Human Language*, ed. J. H. Greenberg (Cambridge: MIT Press), 73–113.

Hare, M., Jones, M., Thomson, C., Kelly, S., and McRae, K. (2009). Activating event knowledge. *Cognition* 111, 151–167. doi: 10.1016/j.cognition.2009.01.009

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach, 3rd Edn.* New York, NY: The Guilford Press.

Hofstede, G., Hofstede, G. J., and Minkov, M. (2010). *Cultures and Organizations: Software of the Mind, 3rd Edn.* New York, NY: McGraw-Hill Education.

Ingendahl, M., Schöne, T., Wänke, M., and Vogel, T. (2021). Fluency in the in-out effect: the role of structural mere exposure effects. *J. Exp. Soc. Psychol.* 92:104079. doi: 10.1016/j.jesp.2020.104079

Kashima, E. S., and Kashima, Y. (1998). Culture and language: the case of cultural dimensions and personal pronoun use. *J. Cross Cult. Psychol.* 29, 461–486. doi: 10.1177/0022022198293005

Kemmerer, D. (2012). The cross-linguistic prevalence of SOV and SVO word orders reflects the sequential and hierarchical representation of action in Broca's area. *Linguist. Lang. Compass* 6, 50–66. doi: 10.1002/lnc3.322

Kostyk, A., Leonhardt, J. M., and Niculescu, M. (2021). Processing fluency scale development for consumer research. *Int. J. Mark. Res.* 63, 353–367. doi: 10.1177/1470785319877137

Kristjánsson, Á., and Campana, G. (2010). Where perception meets memory: a review of repetition priming in visual search tasks. *Atten. Percept. Psychophys.* 72, 5–18. doi: 10.3758/APP.72.1.5

Lee, A. Y., and Aaker, J. L. (2004). Bringing the frame into focus: the influence of regulatory fit on processing fluency and persuasion. *J. Pers. Soc. Psychol.* 86, 205–218. doi: 10.1037/0022-3514.86.2.205

Lee, S. W. S., Oyserman, D., and Bond, M. H. (2010). Am I doing better than you? That depends on whether you ask me in English or Chinese: Self-enhancement effects of language as a cultural mindset prime. *J. Exp. Soc. Psychol.* 46, 785–791. doi: 10.1016/j.jesp.2010.04.005

Lucy, J. A. (1992). *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511620713

Lupyan, G., Abdel Rahman, R., Boroditsky, L., and Clark, A. (2020). Effects of language on visual perception. *Trends. Cogn. Sci.* 24, 930–944. doi: 10.1016/j.tics.2020.08.005

Maass, A., Karasawa, M., Politi, F., and Suga, S. (2006). Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person representation. *J. Pers. Soc. Psychol.* 90, 734–750. doi: 10.1037/0022-3514.90.5.734

Maass, A., and Russo, A. (2003). Directional bias in the mental representation of spatial events: nature or culture? *Psychol. Sci.* 14, 296–301. doi: 10.1111/1467-9280.14421

MacDonald, M. C., Pearlmutter, N. J., and Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychol. Rev.* 101, 676–703. doi: 10.1037/0033-295X.101.4.676

MacKenzie, S. B., Lutz, R. J., and Belch, G. E. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: a test of competing explanations. *J. Mark. Res.* 23:130. doi: 10.1177/002224378602300205

Maljkovic, V., and Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Mem. Cognit.* 22, 657–672. doi: 10.3758/BF03209251

Markus, H. R., and Kitayama, S. (1991). Culture and the self: implications for cognition, emotion, and motivation. *Psychol. Rev.* 98, 224–253. doi: 10.1037/0033-295X.98.2.224

Masuda, T., and Nisbett, R. E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *J. Pers. Soc. Psychol.* 81, 922–934. doi: 10.1037/0022-3514.81.5.922

Meir, I., Aronoff, M., Börstell, C., Hwang, S. -O., Ilkbasaran, D., Kastner, I., et al. (2017). The effect of being human and the basis of grammatical word order: insights from novel communication systems and young sign languages. *Cognition* 158, 189–207. doi: 10.1016/j.cognition.2016.10.011

Morris, M. W., and Mok, A. (2011). Isolating effects of cultural schemas: cultural priming shifts Asian-Americans' biases in social description and memory. *J. Exp. Soc. Psychol.* 47, 117–126. doi: 10.1016/j.jesp.2010.08.019

Murayama, K., Izuma, K., Aoki, R., and Matsumoto, K. (2016). "'Your choice' motivates you in the brain: the emergence of autonomy neuroscience," in *Recent Developments in Neuroscience Research on Human Motivation*, eds. S. Kim, J. Reeve, and M. Bong (Bingley: Emerald Publishing Limited), 95–125. doi: 10.1108/S0749-742320160000019004

Nicoladis, E., and Foursha-Stevenson, C. (2012). Language and culture effects on gender classification of objects. *J. Cross. Cult. Psychol.* 43, 1095–1109. doi: 10.1177/0022022111420144

Nisbett, R. E., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychol. Rev.* 108, 291–310. doi: 10.1037/0033-295X.108.2.291

Polinsky, M., and Magyar, L. (2020). Headedness and the lexicon: the case of verb-to-noun ratios. *Languages* 5:9. doi: 10.3390/languages5010009

Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J. P., and Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *J. Res. Pers.* 40, 99–120. doi: 10.1016/j.jrp.2004.09.001

Reber, R., Schwarz, N., and Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Pers. Soc. Psychol. Rev.* 8, 364–382. doi: 10.1207/s15327957pspr0804_3

Regier, T., and Kay, P. (2009). Language, thought, and color: Whorf was half right. *Trends Cogn. Sci.* 13, 439–446. doi: 10.1016/j.tics.2009.07.001

Roberson, D., Davies, I., and Davidoff, J. (2000). Color categories are not universal: replications and new evidence from a stone-age culture. *J. Exp. Psychol. Gen.* 129, 369–398. doi: 10.1037/0096-3445.129.3.369

Schmitt, B. H., Pan, Y., and Tavassoli, N. T. (1994). Language and consumer memory: the impact of linguistic differences between Chinese and English. *J. Consum. Res.* 21:419. doi: 10.1086/209408

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *J. Consum. Psychol.* 14, 332–348. doi: 10.1207/s15327663jcp1404_2

Skerrett, D. M. (2010). Can the Sapir-Whorf hypothesis save the planet? Lessons from cross-cultural psychology for critical language policy. *Curr. Iss. Lang. Plan.* 11, 331–340. doi: 10.1080/14664208.2010.534236

Van Gompel, R. P. G., and Arai, M. (2018). Structural priming in bilinguals. *Biling: Lang. Cog.* 21, 448–455. doi: 10.1017/S1366728917000542

Whorf, B. L. (1956). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. New York, NY: Wiley.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., and Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proc. Natl. Acad. Sci. USA* 104, 7780–7785. doi: 10.1073/pnas.0701644104

Wolff, P., and Holmes, K. J. (2011). Linguistic relativity. *WIREs Cogn. Sci.* 2, 253–265. doi: 10.1002/wcs.104

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* 9, 1–27. doi: 10.1037/h0025848

Zhang, S., and Schmitt, B. (1998). Language-dependent classification: the mental representation of classifiers in cognition, memory, and ad evaluations. *J. Exp. Psychol. Appl.* 4, 375–385. doi: 10.1037/1076-898X.4.4.375