Masterarbeit / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Linguistic Modulation of Visual Attention - The Influence of a Tight-fit or Loose-fit Singleton Distractor on Attention in Korean versus German Speakers"

verfasst von / submitted by

Johan Albert Saelens BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Vienna, 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie UG2002
Betreut von / Supervisor:	Prof. Dr. Ulrich Ansorge

Expression of Thanks

First and foremost, I am very grateful for the support I received from my supervisor Prof. Dr. Ulrich Ansorge. He was always kind and tolerant and yet precise and detailed in his feedback. Furthermore, I want to thank Dr. Florian Goller for his support, especially concerning the practicalities of carrying out the actual experiment but also during the analysis of the results. In addition, I also highly appreciate Daniel Mitic's help in recruiting participants and collaborating with me in conducting the experiment.

Furthermore, I could not have completed this thesis without my parents Eva Onnen-Saelens and Robert Saelens who not only supported me financially but also always had an open ear during the course of finding my way through the master's program and this thesis.

Last but not least, my thanks go my girlfriend Julia Müller who supported me both motivationally and emotionally during the course of the project. Also, without her expertise in proofreading and layout this master's thesis would not look the way it does now.

Table of Contents

Introduction

Historical Background

Languages differ greatly in respect to how they categorize the world in terms of fundamental properties such as color, time or space (Boroditsky, 2001; Winawer et al., 2007). Yet it is a controversial topic of debate to what extent these linguistic differences affect cognition. The debate about the relation of language and thought ranges back into antiquity, where Plato famously expelled poets from his Republic, because of his suspicion of their mastery in manipulating our minds by means of language. Yet not everyone agreed that language has such a profound impact on thought. Philosophers such as the ancient Greek pupil of Plato, Aristotle or the English philosopher John Locke thought of language as simply being a means to express our thoughts, which are independent of the language in which we express them.

Looking at present-day psychology the topic is still open to debate. According to the well-known Sapir–Whorf hypothesis the language of an individual or a society determines their respective world view. The American linguist Benjamin Whorf (Whorf, Carroll, Levinson, & Lee, 2012) reasoned, for instance, that the different verbal tenses in American Indian languages might influence the actual perception of time and punctuality. However, it is debated whether these findings should be seen as proof that human thought is fundamentally relative to language or if these phenomena should be framed as volatile deviations of an otherwise hardwired, universal cognition. While initially there has been a dominance of the theory of a shared universal underlying structure of all languages (Chomsky, 1957) there has been a large increase in empirical work examining differences between various languages and their interplay with cognition in the last two decades (Semenuks, Philips, Dalca, Kim, & Boroditsky, 2017).

Linguistic Relativists

While in present day psychology most authors agree that language shapes cognition to a certain degree, current debate focuses on working out the exact magnitude of the effect of language on cognition. On the one hand, researchers demonstrated quite powerful effects such as the inability to encode exact numerical values when the use of number words is suppressed by a verbal interference task, in which the participants linguistic capacity is depleted by a parallel verbal task

(Frank, Fedorenko, Lai, Saxe, & Gibson, 2012). Along these lines, it has been demonstrated, that pre-linguistic infants and also a number of cultures do not represent exact quantities higher than three but rely on an approximate number system (Cassanto, 2015; Dehaene, 1999). Critically, also homesigners, people who did not learn any manual or oral language, do not have the capacity to match exact numbers greater than three (Spaepen, Coppola, Spelke, Carey, & Goldin-Meadow, 2011). This result hints towards the importance of language rather than culture for learning certain concepts. According to these models, linguistic relativism might amount to an influence of online use of language on high-level cognition. Another strain of research demonstrates even low-level effects of language on cognition, such as the influence of color vocabulary on the discrimination of actual colors. When English speakers are compared to Russian speakers, which have a different color term for light blue and dark blue (*siniy* and *goluboy*), Russian speakers are faster in discriminating colors falling into one of these categories. Furthermore, researchers observed wide-ranging effects of simple grammatical categories, such as female and male, on the attribution of gender-associated qualities to objects with the respective gender. Such an effect could potentially affect our interaction with virtually every object we encounter since gender is applied to every noun and also adjective or verb depending on the language. Gender-based attribution even reaches into art were concepts such as love or death are usually personified according to their gender in the artists respective language (Segel & Boroditsky, 2011). This linguistic effect can also have severe consequences due to lower perceived risk for feminine-named hurricanes (e.g., Alexandra) compared to their male counterparts (e.g., Alexander), leading to an increased death toll due to less preparedness (Jung, Shavitt, Viswanathan, & Hilbe, 2014). Even abstract concepts such as *time* seem to be at least partially influenced by a person's linguistic background and the common reading direction of one's native language. While English speakers describe time in horizontal terms, moving from left to right, Mandarin speakers describe time as vertical, moving from the top to the bottom (Boroditsky, 2001). This leads Mandarin speakers to respond faster to confirm that March comes before April when previously presented with a vertical prime, compared to an advantage for English speakers when a horizontal prime is presented. Furthermore, patients with hemispatial neglect, a neuropsychological condition after damage to the inferior temporal lobe leading to attentional neglection of one side of the visual field, tend to also have difficulties encoding and consequently recalling stimuli falling on the left (associated with past in English speakers) or right (associated with future) mental time line depending on the damaged hemisphere (Saj, Fuhrman, Vuilleumier, & Boroditsky, 2014). When it comes to linguistic influence on attention, Kersten et al. (2010) demonstrated that English speakers are better in picking up the feature *manner of motion* (e.g., running or hopping) of an object compared to Spanish speakers were manner is usually not linguistically marked. Crucially, this is also the case for bilingual Spanish/English speakers when being in an English compared to a Spanish speaking context. Thus, participants automatically attended to features marked in their active language. To conclude, linguistic relativists claim that language has low-level, broad, substantial influences which may lead to real world implications.

Non-relativists

On the other hand, researchers argued for the *modularity of mind*, claiming that the mind consists of several independent cognitive modules, language being just one of them (Fodor, 1983). These modules are all considered to represent independently evolved structures serving a distinct evolutionary purpose. Thus, abilities as diverse as solving logic problems, the ability to attribute mental states, navigating in the environment or appreciating music are all considered to depend on a different module and are thus independent of language (Fedorenko & Varley, 2016). The proponents of this view argue that online effects which often disappear under the condition of verbal interference might represent wide ranging linguistic effects but cannot be considered deep-rooted changes to core cognition. At this point, linguistic relativists are in danger of making a circular argument by providing experiments during which language is used by participants as the sole evidence for permanent linguistic influences on cognition (Pinker, 2003). It has been shown that the effect of many linguistic differences disappears when language use is not required during the experiment (Li, Dunham, & Carey, 2009). Thus, the gold standard for linguistic relativists should be to demonstrate changes in cognition which are not directly related to the online use of language. Furthermore, certain effects, like that of grammatical gender on attribution of stereotypical characteristics or the conception of time in specific spatial terms, proved to be modifiable in the course of a single experimental session (Boroditsky, 2001; Phillips & Boroditsky, 2003). In addition to being easily modifiable, effects like the influence of number words on representation of exact quantities can be bypassed when processed non-linguistically by replacing verbal with visual representations like a mental abacus (Frank & Barner, 2012).

Thus, there seems to be a sophisticated interplay between language and cognition that must be examined closely in each domain separately, rather than a singular explanation of how language shapes cognition that fits all domains. Furthermore, the answer to the question if language influences cognition depends on the respective authors opinion on how language is involved (online or offline) and which changes in cognition are caused (permanent or impermanent). Thus, two researchers might give a different answer to the question if language changes cognition based on their opinion what a substantive linguistic change means, even though they look at the same experimental data (e.g., Winanwer et al., 2007, see the influence of verbal interference on a given linguistic effect as a proof while Gleitman and Papafragou, 2013, see it as a disproof for linguistic relativity). In the case of our experiment we set out to demonstrate an effect on nonlinguistic cognition, namely visual attention, that is due to an offline influence of linguistic categories.

Tight-fit and Loose-fit Distinction

One domain that is characterized by a particularly high cross-linguistic semantic variation is that of spatial terms (Majid, Jordan, & Dunn, 2015). This reflects a general trend for an increased semantic variety in relational terms compared to objects (Gentner & Boroditsky, 2001). There are, for example, more terms to express spatial relations compared to terms for containers, colors or body parts. On the other hand, the domain of space has been taken as a primary example for a cognitive universal that is structuring language (Clark, 1997). Thus, the field of spatial relations proves to be a particularly suiting testing ground for exploring the effect of language on cognition.

One such example for a difference in terms of semantic categorization of spatial relations which has been investigated by Goller, Lee, Ansorge and Choi (2017) is the Korean distinction between tight-fit (*kkita*) and loose-fit (*netha*). While German (and English for that matter) does not directly distinguish between instances of objects which tightly fit around another object (e.g., a mobile phone in a suitable case) and objects which merely fit loosely (e.g., an apple in a bowl), this is indeed the case in Korean. German speakers, on the other hand, distinguish between containment and support – (a) den Kuli IN die Kappe stecken ("put pen in cap"); (b) die Kappe AUF den Kuli stecken ("put cap on pen") – which are collapsed into the Korean *kkita* (tight-fit). Initially, both Korean and English infants differentiate objects along the tight-fit and loose-fit dimension indicated by an increased looking time after a switch from a tight-fit to a loose-fit object, or vice versa, in the *preferential looking paradigm* (Hespos & Spelke, 2004). However, the sensitivity for this spatial distinction is already weakened by the age of 29 months and then disappears gradually (Choi, 2006). While it is still possible for adult German (or English) speakers to distinguish tight-fit and loose-fit objects using adverbs (e.g., eng/tight; locker/loose) it is less common to actually make such a distinction when speaking (Yun & Choi, 2018). Furthermore, this semantic categorization translates into differences in both perception of similarity and looking behavior (Goller et al., 2017). Crucially, Choi and Hattrup (2012) showed that linguistic differences on the tight-fit and loose-fit dimension between Korean and English speakers in a nonverbal spatial categorization task, where participants had to compare a target video to two subsequent videos, could not be eradicated by a parallel verbal interference task. This constitutes evidence for a relatively stable, offline effect of the spatial distinction.

Experimental Design

A particularly intuitive method to investigate the Sapir-Whorf-Hypothesis are comparative studies, either with native speakers of two given languages or with bilinguals. By comparing two separate groups exhibiting a particular linguistic difference one can measure the influence of that linguistic factor on a given variable. In this experiment we will focus on the question of how the linguistic difference between loose-fit and tight-fit affects the basic cognitive process of visual attention in Korean versus German speakers.

A paradigm that has originally been used to demonstrate the pure bottom-up nature of attention is the *additional singleton paradigm* (Theeuwes, 1992). Theeuwes (1992) demonstrates that when participants search for a target only differentiated by a single feature (e.g., green square) from an array of stimuli (e.g., green circles), they are distracted by an additional stimulus that differs from the array by a different feature (e.g., blue circle). The author claims that this is due to attention being allocated sequentially in order of the saliency of each stimulus. This interpretation contrasts with the theory of an initial parallel stage guiding attention in a top-down manner to task-relevant stimulus dimensions (Folk, 2013).

In the additional singleton paradigm, participants usually react to the inclination of a bar inside the target stimulus (horizontal; vertical). There has been evidence for a *congruency effect*, namely a decrease in RT when the inclination of the bar inside the target is congruent with that inside of the distractor (see Figure 1 Block 2 for a congruent condition). Theeuwes (1994) argued that this effect should be seen as evidence for the spatial allocation of attention towards the distractor. When attended, the processing of the distractor's identity and the inclination of the bar inside of it then leads to a priming effect facilitating the response to the actual target. However, Folk (2013) has argued for an explanation in terms of an increased filtering cost during parallel processing of all salient stimuli rather than a sequential allocation of spatial attention.

In addition, it has been observed that when the color of a target is kept constant over time the target of a given trial acts as a prime for consecutive trials, leading to a decrease in RT (Pinto, Olivers, & Theeuwes, 2005). The opposite is true for the distractor color. This effect of intertrial priming has been used as a bottom-up instead of a top-down explanation for why participants are faster when searching for a known target in a given block of trials. According to this explanation, it would be aggregated priming instead of a *contingent capture* effect, according to which stimuli are matched to a target template, explaining the data. However, this explanation only holds true if the target and distractor identity is kept constant during the whole block. Thus, in this experiment we used a mixed block design in which target and distractor color varied within blocks. However, it has been shown recently that intertrial priming cannot account for contingent capture effects entirely (Büsel, Voracek, & Ansorge, 2018). Still, we controlled for intertrial priming between trials.

Hypotheses

Thus, in this experiment we tested if tight-fit and loose-fit stimuli serve as distractors for Korean speakers and not German speakers in the additional singleton paradigm. We compared a shape distractor condition (tight-fit; loose-fit) in the second block with a condition with no distractor in the first and a one with a color distractor (blue) in the third block. We expect an increase in reaction time (RT) in the shape distractor condition in Korean compared to German speakers. A demonstration of such a distractor effect of tight-fit and loose-fit stimuli in Korean speakers and not German speakers would constitute evidence for a linguistically modifiable process guiding visual attention. We did not expect such a difference between German and Korean speakers in both, the condition without a distractor and that with a color distractor. Furthermore, we expected an increased RT when comparing within Korean speakers both the shape distractor condition and the color distractor condition to the no-distractor condition. In German speakers we expected only a difference in RT in the color distractor condition. In addition, in Korean speakers we expected to observe a congruency effect, namely a decrease in RT when the inclination of the bar inside of the target and distractor were congruent, in the case of the shape distractor and color distractor condition. This result would hint towards a spatial allocation of attention to the distractor. We expected to see this effect in German speakers only in the case of the color distractor.

Method

Participants

Thirty-two German speaking psychology students (n = 32, M_{age} = 23.21 years, SD = 2.87 years, 18 females) from Vienna University who received credit for participation and 32 Korean communication sciences students (n = 32, M_{age} = 22.50 years, SD = 2.53 years, 17 females) from Konkuk University, Seoul were tested. Participation was on a voluntary basis and thus an informed consent was given by every participant. Before the experiment, all participants where naïf to the experiments' hypothesis but participants were debriefed after finishing the experiment. All participants were native speakers of their respective language (German in Austria and Korean in Korea). Instructions were given in the native language of the participants, meaning Korean in the case of Korean speakers and German for all German speakers. Normal color vision was ensured by using a four card Ishihara test. All participants had normal or corrected to normal visual acuity.

Participants with an error rate (ER) above 20% in at least one block and trials in which RT was greater than 2.5 standard deviations or smaller than 300 ms (indicating the initiation of a response before seeing the stimulus) were excluded from the analysis of RT. This resulted in one exclusion in the German speaking sample. Overall 4.87% of the trials in Block 1, 7.63% in Block 2 and 5.78% in Block 3 were excluded.

In order to compute our sample size, we conducted an a priori power analysis using *G*Power* (Faul, Erdfelder, Buchner, & Lang, 2009). We assumed a statistical power of 80%, a moderate effect size and an alpha level of .05.

Apparatus

We conducted the experiment using the Python based software package *Psychopy 1.84* (Peirce, 2007) run on a Windows operating system. The experiment was conducted in the computer laboratory of the Institute for Cognitive Psychology at Vienna University. The experiment was conducted in a darkened room using only a dim light two meters behind the participant. Participants were wearing headphones in order to minimalize distraction. During the experiment, participants used a chin and forehead rest keeping their head stable about 57 cm from the computer screen. The experiment was administered on a screen with a resolution of 1,024 x 768 pixels at a vertical refresh rate of 60 Hz. The colors of the stimuli were equiluminant at 54 cd/m^2.

Stimuli and procedure

The experiment was preceded by a practice session in which participants conducted 15 practice trials in the same design as in Block 1 and received feedback of their total accuracy after each trial. No feedback was given during the actual experiment. The experiment consisted of three blocks with a break after each block. A grey fixation point (0.5°; −3.6/−11.9) against a black background (0.5/0.8) was used during the experiment. Stimuli were presented equidistantly with an eccentricity of 5.0° grouped on an imaginary circle. The stimuli were either red (67.9/44.7), green (−63.8/38.4) or blue (49.7/−98.5) and consisted of a central piston (3.3°× 1.2°) surrounded by an either tightly or loosely fitting cylinder (3.3°× 3°). After seeing a central fixation point for 300 ms, an array of eight stimuli, including one target and in Block 2 and 3 one distractor (see Figure 1), was presented until participants made a response. Furthermore, the experiment included two between-subjects conditions. In the first condition all the stimuli but the distractor in Block 2 consisted of the same tight-fit image of a stick surrounded by cylinders, while in the second condition the stimuli were presented in a loose-fit manner. All stimuli were presented as 3D-objects since it has been shown that the tight-fit and loose-fit distinction is uniquely applied to 3D-objects (Bowerman & Choi, 2003). In Block 1, consisting of 128 trials, all stimuli but the target had a uniform color (target red and array green and vice versa) and either a vertical or a horizontal bar was present in the center of the stimuli. Participants were told to respond to the inclination of the bar (horizontal press *f*; vertical press *j*) in the center of the target. Participants were encouraged to respond fast and accurately. In Block 2, consisting of 256 trials, the task was essentially identical but in this case a shape distractor of the opposite fit (loose-fit for tight-fit condition and vice versa) was always present among the stimuli. In Block 3, consisting of 256 trials, the distractor had a different color (blue) than the rest of the stimuli. It took on average half an hour for participants to complete the experiment. We measured RT as well as ER as a dependent variable. Only correct responses were included in the analysis of RT. Stimuli and procedure were identical in the Korean and the Austrian laboratory.

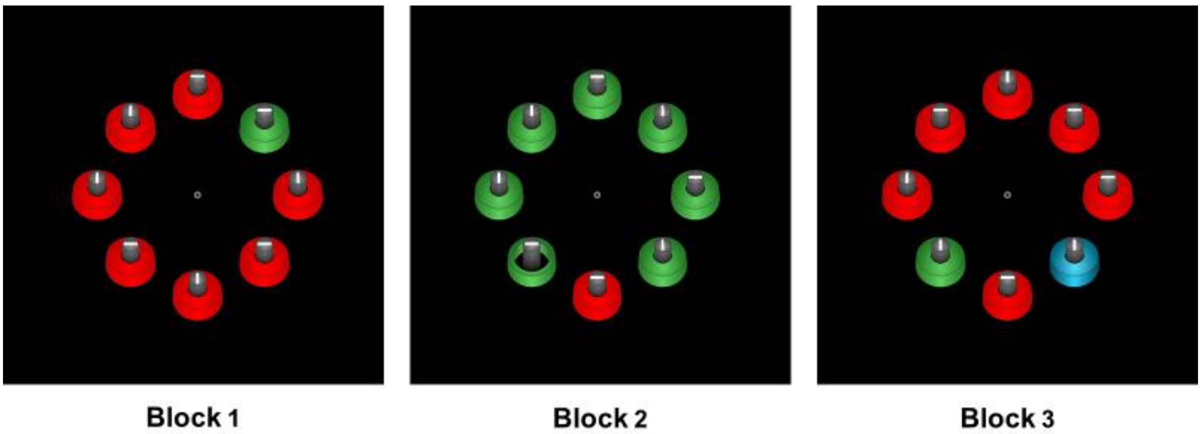

Figure 1. Stimuli display presented in Block 1, Block 2 and Block 3. In Block 1 there is a green target. In Block 2 one can see a red target and a loose-fit shape distractor. Block 3 shows a green target and a blue color distractor. All stimuli but the distractor in Block 2 are presented as tight-fit stimuli. There was also a between-subjects condition where all stimuli were presented as loose-fit stimuli expect for the distractor in Block 2 which was in this case presented as a tight-fit distractor.

Design

An effect of intertrial priming and thus the priming of stimuli with the same identity as the target in the previous trial was accounted for by switching the color of the target, the array, and in Block 2 also the distractor, pseudo-randomly in half of the trials. We further included this factor as a dependent variable in the analysis. The frequency of all target and distractor positions as well as orientations (including target distractor congruence) were evenly distributed in a pseudo-randomized manner so that no information about neither the position nor the identity of the target was hinted by the distractor.

Due to the pseudo-random distribution of vertical and horizontal bars inside the target, the inclination of the bars between target and distractor were respectively congruent and incongruent in half of the trials (see Figure 1 Block 2 for a congruent and Block 3 for an incongruent condition). A decrease of RT to the target in the case of a congruent condition has been taken to indicate an allocation of attention to the distractor (Theeuwes, 1994). This would argue against an allocation of attention guided by an initial parallel stage guiding the attentive stage to the task-relevant stimulus dimension. This combination of search costs due to distractors and compatibility effects has been taken as evidence that spatial attention is allocated in a bottom-up fashion to the salient but irrelevant singletons. Thus, we included the distractor-target congruency as a dependent variable in the analysis.

Results

Reaction Time

Blocks. If the Mauchly test indicated a violation of sphericity, *p* values were adjusted according to the Greenhouse-Geisser correction. An analysis of variance (ANOVA) with repeated measures on mean RT for each participant with blocks (Block 1; Block 2; Block 3) as a within-subject variable and language (German; Korean) as a between-subjects variable revealed a significant difference, $F(2, 124) = 48.59$, $p < .001$, $n_p^2 = .44$, between the conditions. We observed only a trend towards an interaction of block and language, $F(2, 124) = 3.13$, $p = .069$, $n_p^2 = .05$. Post hoc comparisons using Bonferroni correction showed participants were slower in Block 3, $M = 958.87$, $SD = 28.03$, compared to Block 1, $M = 775.35$, $SD = 13.29$, $p < .001$, and Block 2, $M = 781.06$, $SD = 14.46$, p < .001, while no significant difference was found between Block 1 and 2 ($p = 1$). There was no main effect of language, $F(1, 62) = 2.45$, $p = .123$, indicating no general difference in speed between German and Korean speakers. See Figure 2 for an illustration.

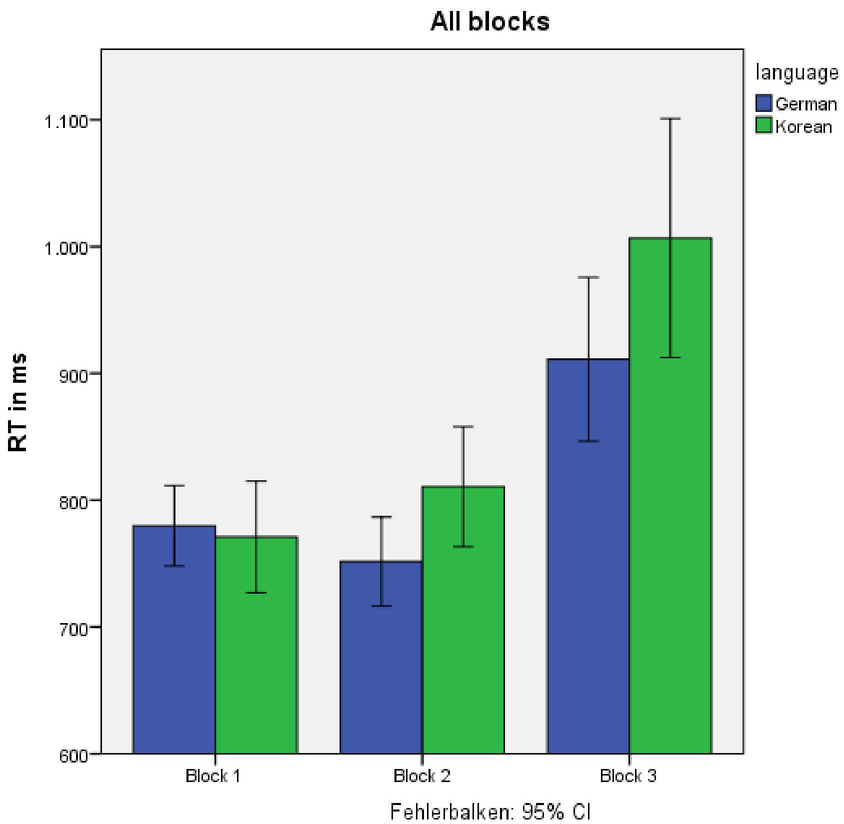

Figure 2. Mean reaction time for Block 1, Block 2 and Block 3 in German and Korean speakers.

When analyzing the German and Korean speaking samples separately in an ANOVA with repeated measures on mean RT, post hoc tests using Bonferroni correction showed only a significant difference between Block 1, M = 779.78, SD = 15.56, and 2, M = 751.51, SD = 17.25, p < .05, in German speakers but not Block 1, M = 770.92, SD = 21.56, and Block 2, M = 810.61, SD = 23.21, p = .154, in Korean speakers. This result indicates a faster RT for Block 2 in German speakers while demonstrating a trend in the opposite direction for Korean speakers.

Language. The analysis with blocks as a within-subject variable was followed by a separate ANOVA with repeated measures of mean RT for each of the three blocks with color (switch; repetition) in Block 1 and color and congruency (congruent; noncongruent) in Block 2 and 3 as a within-subject variable and language (German; Korean) as a between-subjects variable. Critically, in Block 2 the ANOVA revealed a main effect of language, $F(1, 62) = 4.18$, $p < .05$, $n_p^2 = .06$, evidencing faster RT of German speakers compared to Korean speakers. No main effect of language was found in Block 1 or 3, all *F*s < 2.92, all *p*s > .092.

Color. Regarding Block 1 the ANOVA revealed a main effect of color, $F(1, 62) = 101.23$, $p < .001$, $n_p^2 = .62$. Concerning Block 2, again the ANOVA revealed a main effect for switching the color of the target, $F(1, 62) = 128.86$, $p < .001$, $n_p^2 = .74$. Similarly to Block 2, in Block 3 there was a main effect for switching the color of the distractor, $F(1, 62) = 178.11$, $p < .001$, $n_p^2 = .742$, translating into a faster RT when the target color did not change.

Congruency. Furthermore, we found a main effect for congruency in Block 3, $F(1, 62) = 6.08$, $p < .05$, $n_p^2 = .09$, indicating faster RT in the congruent condition. No such effect was found in Block 2, $F(1, 62) = .27$, $p = .605$. This was also not the case when just looking at the Korean speaking sample, $F(1, 31) = .32$, $p = .579$. See Figure 3 for an illustration

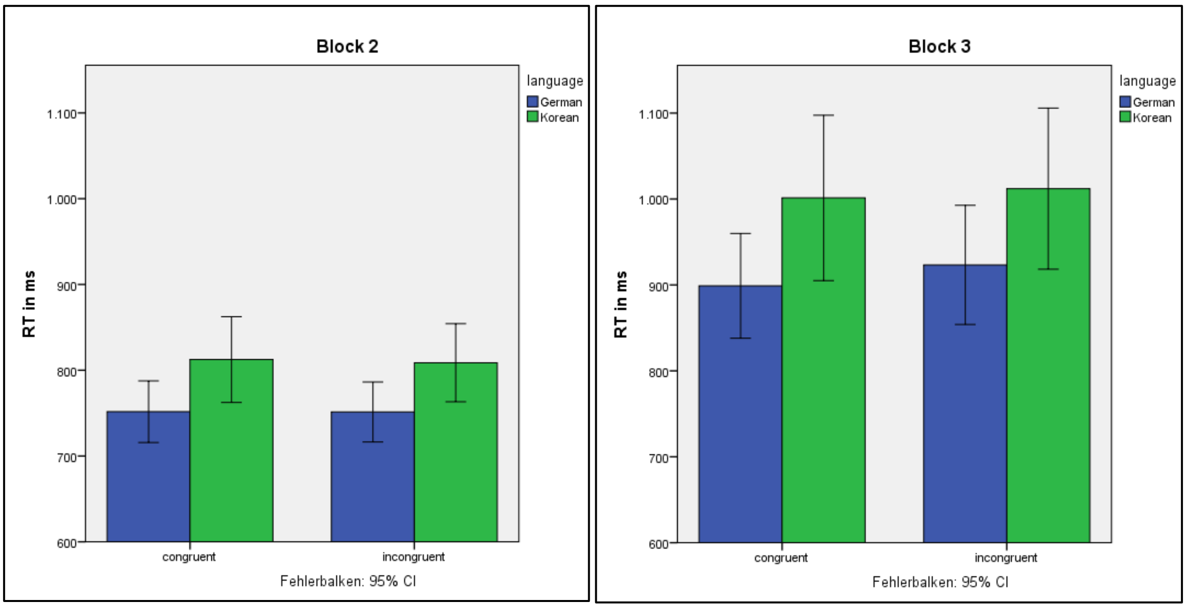

Figure 3. Mean reaction time for the congruent and incongruent condition in Block 2 and Block 3 in German and Korean speakers.

Error Rates

Blocks. Again, we conducted an ANOVA with repeated measures on mean ERs for each participant with blocks (Block 1; Block 2; Block 3) as a within-subject variable and language (German; Korean) as a between-subjects variable. Similarly, if the Mauchly test indicated a violation of sphericity, *p* values were adjusted according to the Greenhouse-Geisser correction. We used arcsine-transformed ERs during this part of the analysis. The analysis revealed no significant effect, all *F*s < 12.78, all *p*s > .10. We found only a main effect of language, $F(1, 62) = 30.68$, $p < .001$, $n_p^2 = .33$, indicating a higher ER in German (3.08%) than Korean speakers (0.94%).

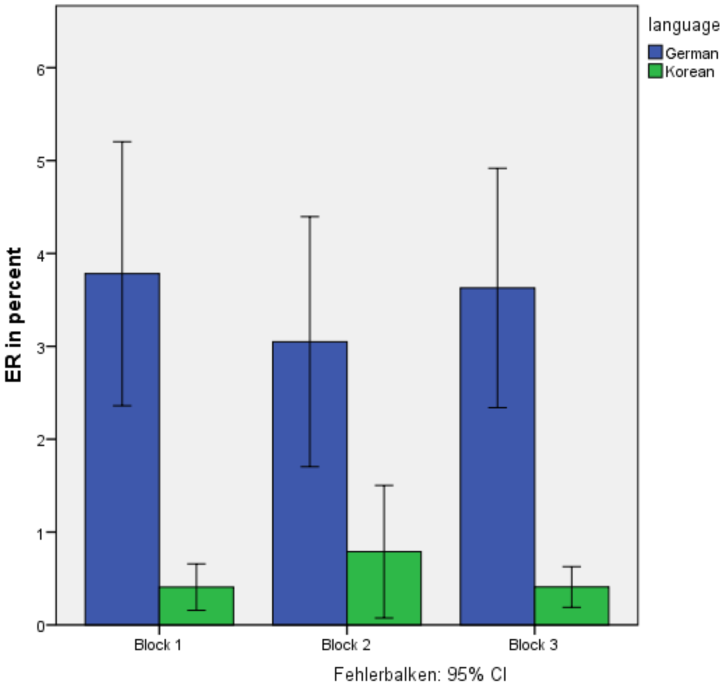

Figure 4. Mean error rates for Block 1, Block 2 and Block 3 in German and Korean speakers.

Language. Similarly, as for RT we conducted an ANOVA with language (German; Korean) as a between-subjects variable and color (switch; repetition) in Block 1 and color plus congruency (congruent, noncongruent) in Block 2 and 3 as a within-subject variable on ERs. The ANOVA revealed a main effect for language in Block 1, $F(1, 62) = 32.48$, $p < .001$, $n_p^2 = .34$, Block 2, $F(1, 62) = 16.04$, $p < .001$, $n_p^2 = .21$, and Block 3, $F(1, 62) = 41.04$, $p < .001$, $n_p^2 = .40$. Korean speakers in all blocks had a lower ER compared to German speakers. See Figure 4 for an illustration.

Color. The ANOVA revealed a main effect for color in Block 3, $F(1, 62) = 4.68$, $p < .05$, $n_p^2 = .07$, resulting in a lower ER if the color stayed the same. No main effect of color was observed in the other blocks, all *F*s < .50, all *p*s > .47.

Congruency. We observed a main effect for congruency in both Block 2, $F(1, 62) = 4.06$, $p < .05$, $n_p^2 = .06$, and Block 3, $F(1, 62) = 14.39$, $p < .001$, $n_p^2 = .19$. In both blocks congruency between target and distractor led to a lower ER. However, when analyzed separately Korean speakers in Block 2 showed a lower ER in congruent (0.36%) compared to incongruent (1.81%) conditions, $t(31) = -2.04$, $p = .050$, $d = 0.36$, while German speakers did not differ in the congruent (3.03%) compared to the incongruent condition (3.46%), $t(31) = -.86$, $p = .394$, $d = 0.15$. On the other hand, when comparing the congruent and the incongruent condition in Block 3 both Korean speakers (0.28% vs. 0.72%), $t(31) = -4.08$, $p < .001$, $d = 0.72$ and German speakers (2.80% vs. 4.91%), $t(31) = -2.23$, $p < .05$, $d = 0.39$, differed significantly between these conditions. See Figure 5 for an illustration.

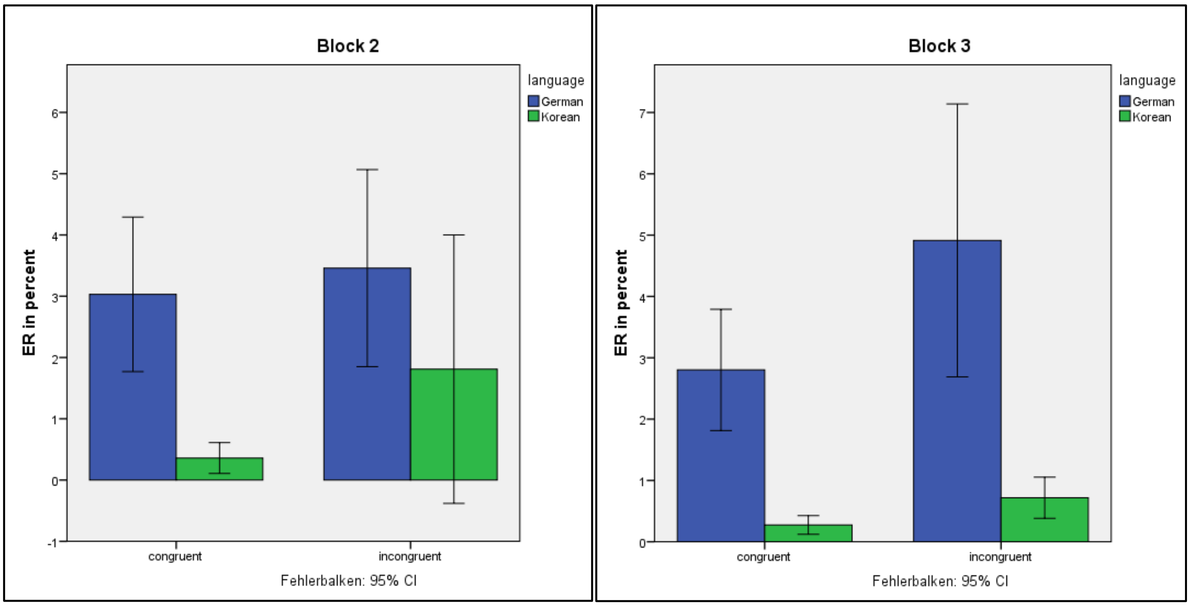

Figure 5. Mean error rates for the congruent and incongruent condition in Block 2 and Block 3 in German and Korean speakers.

Discussion

Considering the current evidence, a distraction by shape singletons, based on the Korean tight-fit and loose-fit distinction, for Korean but not German speaking participants, has been observed. However, Korean speakers showed only a trend towards being distracted when compared to Block 1 where no distractor was present. This could be due to the combination of a relatively small effect, compared to the color distractor, and a learning effect counteracting a possible distraction. Furthermore, we found an effect of intertrial priming for color, indicating a faster RT when color stayed the same. This was expected (Pinto et al., 2005) but does not affect the remaining effects, since we mixed trials with color repetition and non-repetition. A congruency effect was found only in Block 3 where a color distractor was present. This was true both for Korean speakers and German speakers. These results hint towards an allocation of spatial attention towards the color distractor but not shape distractor. It might be that this is due to the generally weaker distraction by shape compared to color singletons (Theeuwes, 1992).

The analysis of ERs did only partially confirm the results of the analysis of RT. However, one should be cautious in interpreting these results since we observed a rather low overall percentage of ERs (Desender, Van Lierde, & Van den Bussche, 2013). In further study one could increase the ER by limiting the time participants are permitted to respond to investigate if the linguistic influence on attention can also be observed in the ERs. As predicted we did find a main effect in congruency both in Block 2 and Block 3. Crucially, in Block 2 only Korean speakers but not German speakers showed a lower ER in congruent compared to incongruent conditions. This can be taken as a hint towards the allocation of attention to the shape distractor (Theeuwes, 1994). Alternatively, one might interpret this result in terms of an increased filtering cost due to the shape distractor (Folk, 2013). There was only a main effect of switching colors in Block 3. It could be that we did not observe an effect in Block 1 and Block 2 due to the low mean ER, possibly leading to a ceiling effect. Finally, we observed a main effect of language in all conditions, indicating that Korean speakers made less errors than German speakers in all blocks. However, this should not affect the main effect of a linguistic distractor on RT, since there was no interaction of block and ER.

Linguistic Effects on Cognition

How can one explain the neuronal and temporal properties of such a linguistic effect on cognition? Concerning the temporal properties of such linguistic influences, it has been demonstrated using electroencephalography (EEG) that the mismatch negativity, an event-related potential (ERP) in response to odd stimuli, in response to linguistic differences can arise as early as 100 ms after stimulus onset (Mo, Xu, Kay, & Tan, 2011). This indicates a quite early linguistic influence on cognition. Regarding the neuronal properties Kumar, Federmeier, Fei-Fei and Beck (2017) demonstrated shared neuronal activity patterns in the occipitotemporal, posterior parietal and frontal cortices in response to verbal and visual stimuli. This activation indicates a common semantic code for pictures and words which would suggest a direct linguistic effect on visual representations. Another strain of research that manages without imaging techniques revealed a possible lateralization of linguistic effects to the left hemisphere. For the case of color Roberson, Pak and Hanley (2008) demonstrated a stronger effect of language specific categorical perception of colors presented in the right visual field for colors that belong to different categories in Korean but not English. This was especially the case for fast responders, since information is subsequently transferred between both hemispheres. This evidence has been further complemented by Mo et al. (2011) showing that the mismatch negativity in response to a color that transcends a language specific color category is considerably smaller when the stimulus is presented in the left visual field compared to the right visual field.

The debate about the temporal and neural properties of linguistic effects is further linked to the debate about the level at which these effects take place. It is controversial if those effects are explainable by influences on perception (Lupyan, 2012), visual representations (Kan, Barsalou, Olseth Solomon, Minor, & Thompson-Schill, 2003; Kumar et al., 2017), working memory (Klemfuss, Prinzmetal, & Ivry, 2012) or a subset of those possibilities.

Lupyan (2017) aims to demonstrate effects of language on early perception by showing that perception of change in blurriness is more accurate in meaningful words compared to meaningless letter strings. According to the researcher, language facilitates more accurate expectations about perception which in turn lead to more precise predictions and enables thus a more exact comparisons between stimuli. This culminates in Lupyan’s label-feedback-hypothesis which claims a general fast and automatic online modulating effect of language on perception (Lupyan, 2012).

Klemfuss et al. (2012) on the other hand argue that verbal labels facilitate retention of a given object in working memory. A retention in working memory would in turn affect the late stage of response and not early perception. This is evidenced by the authors demonstration that the advantage of linguistic labels in finding the oddball in a set of four stimuli is abolished when providing a cue representing the target. Similarly, Slobin (1996) has argued that participants covertly verbalize during the experiment which activates respective representations in working memory. In accordance with this theory, Papafragou, Hulbert and Trueswell (2008) demonstrated that participant’s eye movement pattern follows linguistic distinctions when participants perform a verbal description task but not a memorization task. While the influence of verbalization is further evidenced by studies in which the linguistic effect disappears in a verbal interference condition (Boroditsky et al., 2002; Phillips & Boroditsky, 2003), the theory is challenged by examples of verbal interference having no effect (Choi & Hattrup, 2012; Frank et al., 2012).

This hypothesis is complemented by Kan et al. (2003) who state that conceptual knowledge is generally organized visually and thus grounded in the perceptual system. Therefore, an activation of a semantic category would automatically represent an activation of the corresponding perceptual features of that category, evidenced by an activation of the left fusiform gyrus and area usually associated with mental imagery. This could in turn explain differences in color discrimination or deployment of attention.

However, these hypotheses are not necessarily mutually exclusive, and it may be the case that the nature of the effect differs according to the task and cognitive capacity (attention, discrimination, representation etc.) that is affected by language.

Attentional Mechanisms

Granted that there is a stable effect of language on cognition it is still open to debate what mechanism guides attention towards the linguistic difference. One possibility linguistic differences might be incorporated into cognition is by acting as guiding attributes of attention (Wolfe & Horowitz, 2017). Guiding attributes are those attributes which may be processed preattentively in parallel and thus consequently guide attention allocation both in bottom-up and top-down guidance. While features like color, motion, orientation and size have largely proven to be guiding attributes, the features of shape and semantics, which are of interest in this study, are currently listed as probably guiding and probably not guiding attributes respectively. When it comes to shape, a satisfactory description of what a relevant difference in shape amounts to seems to be missing (Wolfe, 2015). However, some differences in shape, like that between a circle and a square, are fairly uncontroversial guiding features. Concerning the factor semantics, so far studies have primarily investigated intralinguistic search for categories like animal versus artifact (Levin, Takarae, Mine, & Keil, 2001), while our study deals with semantic differences between languages. Furthermore, it might be the case that linguistic differences, like tight-fit and loose-fit, merely modulate the allocation of attention towards certain pre-existing differences in shape. In this case, semantic categories would not guide attention themselves but would help activating certain characteristic differences in shape or other basic features which would in turn guide attention allocation. Similarly, in the case of color, visual search seems to be mediated by color categories and not just differences on the chromatic scale (Yokoi & Uchikawa, 2005). Other proposals suggest that the difference in action affordances (e.g., chair; sitting down) between semantic categories resulting in different cerebral activation is used to distinguish objects belonging to two semantic categories (Levin et al., 2001).

The other way around one might argue that this natural category of fit is suppressed in cultures where the pre-existing spatial differences between tight-fit and loose-fit are collapsed into one category (Hespos & Spelke, 2004). In this case, children would start with a large number of pre-existing concepts, which are then adjusted according the cultural and linguistic environment. This parallels the learning of phonemes, were children initially start with the capability to distinguish phonemes of all languages and then, starting from 6 months on, become increasingly insensitive to phonemes of languages other than their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). However, while there is a critical period for phoneme learning, this does not seem to be case for learning concepts like tight-fit and loose-fit (Yun & Choi, 2018). Crucially, this model of linguistic influence on cognition would go against the claim of a static set of natural categories that can influence the allocation of attention and open the possibility to culturally learned or suppressed concepts (Levin et al., 2001). Nonetheless, it is up to debate if the distraction by the shape singleton should be explained in terms of bottom-up or top-down processing.

Bottom-up. On the one hand, a proponent of a bottom-up approach might argue that shape is integrated into a priority map and then guides attention pre-attentively due to an increased saliency in this area. Even though proponents of the bottom-up approach do not explicitly mention linguistic effects they grant at least a certain modifiability of attention allocation by prior history such as association with reward or fear, priming or statistical learning (Theeuwes, 2018).

When it comes to contingent capture effects, where targets are seemingly matched to a search-template rather than evaluated according to saliency, proponents of the bottom-up approach usually provide two different explanations. First, the results of an increased RT to a known stimulus is explained by intertrial priming rather than specific feature search. Because the target defining feature is typically held constant across trials, priming might selectively enhance processing of the respective feature (Belopolsky, Schreij, & Theeuwes, 2010). However, we did vary the target feature across trials and thus balancing possible intertrial priming effects.

Secondly, proponents of a pure capture perspective usually explain results in terms of the *rapid disengagement hypothesis* (Theeuwes, Atchley, & Kramer, 2000; Theeuwes, Olivers, & Belopolsky, 2010). This hypothesis claims that attention is rapidly disengaged according to saliency instead of deployed only after parallel processing. Thus, a decrease in RT after a target-matching, but not non-matching cue presented previously to the array of stimuli, is explained by a slower disengagement when the stimulus matches the target. Thus, when the cue does not match the target, attention can rapidly disengage from the cue and no cueing effect is observed. On the other hand, when the cue does match the target, attention disengages slower and consequently stays at the attention of the target, leading to a cueing effect. The congruency effect, present when the inclination of the bar inside the distractor matched that inside the target, is taken as evidence for such an initial allocation of attention to the distractor. However, we found only partial evidence for such a rapid disengagement, since no congruency effect was found in the shape condition.

In addition, Theeuwes and Godijn (2002) used the phenomenon of *inhibition of return* (IOR), the reluctance of attention to revisit a location previously already scanned, to demonstrate involuntary attention allocation. The authors found that participants were slower to detect targets at the location where a singleton distractor had previously been presented. This was explained by a previous shift of attention to the distractor and a consequent IOR.

Top-down. On the other hand, a proponent of a top-down approach might argue that stimuli are processed according to the search goals of a person. In this case, distractor costs in the additional singleton paradigm have been explained in terms of filtering costs. Folk (2013) for example has proposed that, when only a small number of salient stimuli is present, their identity is processed in parallel and checked for the target. It has been shown that under low-load conditions with few salient stimuli identity can be processed in parallel (Lavie, 1995). Thus, the congruency effect would be due to processing the identity of the distractor in low-load conditions. This is demonstrated by the finding that the congruency effect but not the general search costs due to the distractor disappears under high-load conditions. Thus, the effect of the distractor on RT and further the congruency effect would be due to filtering costs of parallel processing of all salient stimuli rather than sequential deployment of attention. This explanation of the congruency effect in terms of filtering costs also questions the validity of this effect as an indirect measure of attention allocation. Further evidence stems from an EEG study that demonstrated that salient distractors elicited a P_D, an ERP associated with attentional suppression, rather than a N2pc, associated with attention allocation (McDonald, Green, Jannati, & Di Lollo, 2013). Still, in both the pure capture and the filtering cost explanation it would have to be granted that saliency of the shape singleton is different in the Korean compared to the German speaking condition and thus linguistically modulated. On the other hand, Horstmann, Becker and Ernst (2017) have argued for a serial account, according to which target similar distractors are visited until the actual target is found. Following this account, search efficiency is not only determined by guidance to the target but also dwelling time, skipping and revisiting of the distractor. The more a distractor resembles the target the more participants dwell on it and revisit it after initial rejection. Still other authors have argued for a hybrid approach of serial and parallel processing (Wolfe, 2003). According to his proposal stimuli might enter processing serially (e.g., every 50 ms) but can then be processed at the same time (e.g., total processing takes 300 ms, thus up to 6 stimuli could be processed in parallel). However, in our experiment Koreans were less efficient in the presence of a linguistic shape singleton, even though the linguistic distinction results in a distractor less similar to the target and thus potentially easier to reject for Korean speakers. In addition, participants even tend to skip distractors without attending them when they are considerably dissimilar to the target.

Still, proponents of a contingent capture approach, which stresses the importance of a search template for visual search, could argue that participants might have adopted a singleton search mode (Bacon & Egeth, 1994). Thus, participants would have searched for singletons in general, which would encompass both shape and color singletons, instead of adopting a feature search mode restricted to the two colors red and green. However, this is improbable because participants were explicitly instructed to search only for the colors red and green and there was no switch of target features within a given block. Still, future experiments could test this by forcing participants to adapt a feature search mode rather than a singleton search mode (e.g., by including multiple targets in some trials).

Also, Moores, Laiti, & Chelazzi (2003) demonstrated that stimuli that are semantically associated to the target are also more often recalled and recognized than an unrelated distractor. It is argued that the top-down activation of the target's feature representation automatically also activates the representation of semantically associated stimuli. Hence, the association between the two concepts in long-term memory would translate into an activation in working memory. Thus, in our experiment, it might be the case that the semantic association in Korean between tight-fit and loose-fit stimuli leads to an activation of the associated representation whenever a tight-fit or loose-fit target is searched. This activation then might lead to an increased distraction by distractors with the associated composite features. Nonetheless, it might also be the other way around since tight-fit and loose-fit represent opposite categories and the cited evidence normally stems from non-opposite categories (e.g., motorbike; helmet). Furthermore, in the case of color search, there is no evidence for the facilitation of target search by opposite color cues (Ansorge & Becker, 2014). Thus, to exclude this alternative hypothesis one could introduce a verbal interference condition to inhibit an activation of the spatial distinction in working memory and thus also a consequent activation of any associated concepts. However, this would not exclude the alternative hypothesis that these top-down settings influence directly the waiting of lower level feature detectors without relying on templets present in working memory (Bundesen, 1990).

Future Studies

The spatial properties of the linguistic distraction of attention might be tested in a second version of the experiment involving eye tracking. In this case the potential effect of a distractor on the saccades could be seen in the form of a deviation away from the distractor (Van der Stigchel & Theeuwes, 2005). Even though eye-tracking does not directly measure attention (Posner, 1980), fixation can be taken as a strong indicator for attention. This might further clarify the spatial nature of the effect. Alternatively, one might measure the N2pc, an ERP measured via EEG, as an indicator for covert shifts of attention (Eimer, 1996). This ERP, which occurs roughly 200 ms after stimulus onset, could help clarifying if covert attention, which may act independently of eye movements, shifts to the distractor. These approaches could also help to rule out the possibility that participants visit stimuli other than the target occasionally on a random basis and then dwell longer on the distractor stimulus in the case of Korean speakers. Nonetheless, this explanation is challenged by evidence that participants can suppress overt shifts of attention to a salient distractor (Gaspelin, Leonard, & Luck, 2017).

In addition, one might consider using a shape singleton in Block 3. It has been shown that color categorization for blue and green differs in Korean compared to English. Korean entails the color *pureu-da* as a shared category between green and blue, possibly reducing the distraction of a blue singleton in an array of green stimuli (Roberson et al., 2008). This effect might be less severe in our experiment, since the effect seems to be more prominent in older people, which were not recruited in this experiment (Tyson, 1998). In any case, this should not affect Block 2, since green and red are clearly distinct categories in Korean. However, one might consider an alternative color or shape singleton for Block 3. In any case, comparing the effect of loose-fit and tight-fit singletons with that of another shape singleton might help to ensure that a possible effect is not due to general sensitivity to shape in Koreans.

Even though we did not explicitly instruct participants to verbalize during the experiment, nor was there an obvious advantage of verbalization in the task, it is still possible that participants nonetheless indulged in verbalization while solving the task. A methodological possibility to tackle this problem could be the method of verbal interference. In this method a speaker's verbal capacities are down-regulated by a simultaneous verbal task, such as the shadowing of meaningless syllables, while the participant is completing the main task. Thus, one can observe if certain differences in cognition rely on online-employment of language or are due to deeper cognitive changes (Winawer et al., 2007). It has been shown that the influence of verbal interference on language related tasks differs according to the linguistic difference investigated (Frank et al., 2012; Phillips & Boroditsky, 2003). However, there seems to be an effect resistant to verbal interference in the case of similarity judgments of spatial categories in German and Korean speakers (Choi & Hattrup, 2012). A further approach, taken by Phillips and Boroditsky (2003) is to teach German speakers spatial terms such as *kkita* and *netha* and observe if the effect of spatial language on attention still holds. Thus, by demonstrating the causal effect of language, one might grapple with the inherent difficulty of research in psycholinguistics to disentangle the effect of culture and language respectively.

Implications

When it comes to practical implications, research on visual attention can help to understand attention capture in crucial cases such as medical image perception. Furthermore, research can help informing decisions on how to design an object in order to capture attention. This might be of interest when designing machine interfaces that help guide the user's attention in a beneficial way. Furthermore, this strand of research can help informing the conception of warnings on medicine, road signs or dangerous substances or warning clothing that is maximally visible. In addition, research on visual attention might also be used in marketing to increase visibility of a given product. However, our research shows the potential pitfall of linguistic modifiability of attention. Thus, when designing a product, one has to consider the attentional set of the potential costumer. This might lead to different optimal decisions to attract attention according to the respective language context.

Conclusion

Similar as in other areas of cognition, one's native language appears to also have a significant effect on people's visual attention. People who share a certain concept in their native language automatically pick up on stimuli that share the features of the respective concept. This is the case even if verbalization of the distinction is not explicitly required. However, it rests unclear if there is an actual shift in visual attention since we did not examine a congruency effect in the shape condition. Thus, the spatial nature of the effect still has to be investigated, possibly using eye-tracking as a more direct measure of attention allocation. Furthermore, it is debated if linguistic effects are due to an early effect on perception or later online effects on working memory or the decision stage. Multiple techniques like EEG or verbal interference might help to further clarify these questions. If in fact earlier stages of perception are affected, it is still debated if bottom-up or top-down explanations should account for this effect. However, both an explanation in terms of filtering costs and an effect in terms of pure capture would grant an increased saliency of the shape singleton. On the other hand, an explanation in terms of contingent capture seems unlikely, since the effect occurred spontaneously and no search for shape was instructed. In any case, the results demonstrate a certain degree of linguistic relativism in the deployment of visual attention.

References

Ansorge, U., & Becker, S. I. (2014). Contingent capture in cueing: the role of color search templates and cue-target color relations. *Psychological Research, 78*, 209-221. https://doi.org/10.1007/s00426-013-0497-5

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*, 485-496. https://doi.org/10.3758/bf03205306

Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture?. *Attention, Perception, & Psychophysics, 72*, 326-341. https://doi.org/10.3758/app.72.2.326

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1-22. https://doi.org/10.1006/cogp.2001.0748

Bowerman, M., & Choi, S. (2003). *Space under construction: Language-specific spatial categorization in first language acquisition.* In D. Gentner & S. Goldin-Meadow (Eds.), Language in mind: Advances in the study of language and thought, (pp. 387-427). Cambridge, MA: MIT Press.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*, 523-547. https://doi.org/10.1037//0033-295x.97.4.523

Büsel, C., Voracek, M., & Ansorge, U. (2018). A meta-analysis of contingent-capture effects. *Psychological Research, 82*, 1-26. https://doi.org/10.1007/s00426-018-1087-3

Cassanto, D. (2015). *Linguistic relativity.* In N. Riemer (Ed.), The Routledge handbook of semantics (pp. 158-174). London, England: Routledge. https://doi.org/10.4324/9781315685533

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language, 26*, 207-232. https://doi.org/10.1177/0142723706060748

Choi, S., & Hattrup, K. (2012). Relative contribution of perception/cognition and language on spatial categorization. *Cognitive Science, 36*, 102-129. https://doi.org/10.1111/j.1551-6709.2011.01201.x

Chomsky, N. (1957). *Syntactic structures.* The Hague, Netherlands: Mouton.

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

Dehaene, S. 1999. *The number sense: How the mind creates mathematics.* Oxford, England: Oxford University Press.

Desender, K., Van Lierde, E., & Van den Bussche, E. (2013). Comparing conscious and unconscious conflict adaptation. *PLoS One, 8*, e55976. https://doi.org/10.1371/journal.pone.0055976

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology, 99*, 225-234. https://doi.org/10.1016/s0921-884x(96)95711-2

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. https://doi.org/10.3758/brm.41.4.1149

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences, 1369*, 132-153. https://doi.org/10.1111/nyas.13046

Fodor, J. A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Folk, C. L. (2013). Dissociating compatibility effects and distractor costs in the additional singleton paradigm. *Frontiers in Psychology, 4*:434, 1-10. https://doi.org/10.3389/fpsyg.2013.00434

Frank, M. C., & Barner, D. (2012). Representing exact number visually using mental abacus. *Journal of Experimental Psychology: General, 141*, 134-149. https://doi.org/10.1037/a0024427

Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., & Gibson, E. (2012). Verbal interference suppresses exact numerical representation. *Cognitive Psychology, 64,* 74-92. https://doi.org/10.1016/j.cogpsych.2011.10.004

Gaspelin, N., Leonard, C. J., & Luck, S. J. (2017). Suppression of overt attentional capture by salient-but-irrelevant color singletons. *Attention, Perception, & Psychophysics, 79*, 45-62. https://doi.org/10.3758/s13414-016-1209-1

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language Acquisition and Conceptual Development, 3*, 215-256. https://doi.org/10.1017/cbo9780511620669.010

Gleitman, L., & Papafragou, A. (2013). *32 relations between language and thought.* In D. Reisberg (Ed.), The Oxford handbook of cognitive psychology (pp. 504-523). Oxford, England: Oxford University Press. 10.1093/oxfordhb/9780195376746.013.0032

Goller, F., Lee, D., Ansorge, U., & Choi, S. (2017). Effects of language background on gaze behavior: A crosslinguistic comparison between Korean and German speakers. *Advances in Cognitive Psychology, 13*, 267-279. https://doi.org/10.5709/acp-0227-z

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature, 430*, 453-456. https://doi.org/10.1038/nature02634

Horstmann, G., Becker, S., & Ernst, D. (2017). Dwelling, rescanning, and skipping of distractors explain search efficiency in difficult search better than guidance by the target. *Visual Cognition, 25*, 291-305. https://doi.org/10.1080/13506285.2017.1347591

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 451-468. https://doi.org/10.1037//0096-1523.21.3.451

Jung, K., Shavitt, S., Viswanathan, M., & Hilbe, J. M. (2014). Female hurricanes are deadlier than male hurricanes. *Proceedings of the National Academy of Sciences, 111*, 8782-8787. https://doi.org/10.1073/pnas.1402786111

Kan, I. P., Barsalou, L. W., Olseth Solomon, K., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology, 20*, 525-540. https://doi.org/10.1080/02643290244000257

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General, 139*, 638-653. https://doi.org/10.1037/a0020507

Klemfuss, N., Prinzmetal, B., & Ivry, R. B. (2012). How does language change perception: A cautionary note. *Frontiers in Psychology, 3*:78, 1-6. https://doi.org/10.3389/fpsyg.2012.00078

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606-608. https://doi.org/10.1126/science.1736364

Kumar, M., Federmeier, K. D., Fei-Fei, L., & Beck, D. M. (2017). Evidence for similar patterns of neural activity elicited by picture-and word-based representations of natural scenes. *Neuroimage, 155*, 422-436. https://doi.org/10.1016/j.neuroimage.2017.03.037

Levin, D. T., Takarae, Y., Miner, A. G., & Keil, F. (2001). Efficient visual search by category: Specifying the features that mark the difference between artifacts and animals in preattentive vision. *Perception & Psychophysics, 63*, 676-697. https://doi.org/10.3758/bf03194429

Li, P., Dunham, Y., & Carey, S. (2009). Of substance: The nature of language effects on entity construal. *Cognitive Psychology, 58*, 487-524. https://doi.org/10.1016/j.cogpsych.2008.12.001

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 3*:54, 1-13. https://doi.org/10.3389/fpsyg.2012.00054

Lupyan, G. (2017). Objective effects of knowledge on visual perception. Journal of Experimental Psychology: *Human Perception and Performance, 43*, 794-806. https://doi.org/10.1037/xhp0000343

Majid, A., Jordan, F., & Dunn, M. (2015). Semantic systems in closely related languages. *Language Sciences, 49*, 1-18. https://doi.org/10.1016/j.langsci.2014.11.002

McDonald, J. J., Green, J. J., Jannati, A., & Di Lollo, V. (2013). On the electrophysiological evidence for the capture of visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 849-860. https://doi.org/10.1037/a0030510

Mo, L., Xu, G., Kay, P., & Tan, L. H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 14026-14030. https://doi.org/10.1073/pnas.1111860108

Moores, E., Laiti, L., & Chelazzi, L. (2003). Associative knowledge controls deployment of visual selective attention. *Nature Neuroscience, 6*, 182-189. https://doi.org/10.1038/nn996

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition, 108*, 155-184. https://doi.org/10.1016/j.cognition.2008.02.007

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods, 162*, 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Phillips, W., & Boroditsky, L. (2003). Can quirks of grammar affect the way you think? Grammatical gender and object concepts. *Proceedings of the Annual Meeting of the Cognitive Science Society, 25*, 928-933.

Pinker, S. (2003). *The language instinct: How the mind creates language.* London, England: Penguin.

Pinto, Y., Olivers, C. L., & Theeuwes, J. (2005). Target uncertainty does not lead to more distraction by singletons: Intertrial priming does. *Attention, Perception, & Psychophysics, 67*, 1354-1361. https://doi.org/10.3758/bf03193640

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*A, 3-25. https://doi.org/10.1080/00335558008248231

Roberson, D., Pak, H., & Hanley, J. R. (2008). Categorical perception of colour in the left and right visual field is verbally mediated: Evidence from Korean. *Cognition, 107*, 752-762. https://doi.org/10.1016/j.cognition.2007.09.001

Saj, A., Fuhrman, O., Vuilleumier, P., & Boroditsky, L. (2014). Patients with left spatial neglect also neglect the “left side” of time. *Psychological Science, 25*, 207-214. https://doi.org/10.1177/0956797612475222

Segel, E., & Boroditsky, L. (2011). Grammar in art. *Frontiers in Psychology, 1*:244, 1-3. https://doi.org/10.3389/fpsyg.2010.00244

Semenuks, A., Phillips, W., Dalca, I., Kim, C., & Boroditsky, L. (2017). Effects of grammatical gender on object description. *Proceedings of the 39th Annual Conference of the Cognitive Science Society,* 1060-1065.

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

Spaepen, E., Coppola, M., Spelke, E. S., Carey, S. E., & Goldin-Meadow, S. (2011). Number without a language model. *Proceedings of the National Academy of Sciences, 108*, 3163-3168. https://doi.org/10.1073/pnas.1015975108

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Attention, Perception, & Psychophysics, 51*, 599-606. https://doi.org/10.3758/bf03211656

Theeuwes, J. (1994). *Perceptual selectivity for color and form: On the nature of the interference effect.* In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), Converging operations in the study of visual selective attention (pp. 297–314). Washington, DC: American Psychological Association.

Theeuwes, J. (2018). Visual selection: usually fast and automatic; seldom slow and volitional. *Journal of Cognition, 1*, 1-4. https://doi.org/10.5334/joc.13

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). *On the time course of top-down and bottom-up control of visual attention.* In S. M. J. Driver (Ed.), Attention & performance (pp. 105-125). Cambridge, MA: MIT Press.

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & Psychophysics, 64*, 764-770. https://doi.org/10.3758/bf03194743

Theeuwes, J., Olivers, C. N., & Belopolsky, A. (2010). Stimulus-driven capture and contingent capture. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 872-881. https://doi.org/10.1002/wcs.83

Tyson, R. E (1998). Color naming and color categorization in Korean. *Japanese/Korean Linguistics, 7*, 177-196.

Van der Stigchel, S., & Theeuwes, J. (2005). Relation between saccade trajectories and spatial distractor locations. *Cognitive Brain Research, 25*, 579-582. https://doi.org/10.1016/j.cogbrainres.2005.08.001

Whorf, B. L., Carroll, J. B., Levinson, S. C., & Lee, P. (2012). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf.* Cambridge, MA: MIT Press. https://doi.org/10.2307/411163

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 7780-7785. https://doi.org/10.1073/pnas.0701644104

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*, 70-76.

Wolfe, J. M. (2015). *Visual search.* In J. Fawcett, E. Risko, & A. Kingstone (Eds.), The handbook of attention (pp. 27-56). Cambridge, MA: MIT Press.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*, 1-8. https://doi.org/10.1038/s41562-017-0058

Yokoi, K., & Uchikawa, K. (2005). Color category influences heterogeneous visual search for color. *Journal of the Optical Society of America A, 22*, 2309-2317. https://doi.org/10.1364/josaa.22.002309

Yun, H., & Choi, S. (2018). Spatial semantics, cognition, and their interaction: A comparative study of spatial categorization in English and Korean. *Cognitive Science, 42*, 1-41. https://doi.org/10.1111/cogs.12622

List of Figures

List of Abbreviations

ANOVA	Analysis of Variance
EEG	Electroencephalography
ER	Error Rate
ERP	Event Related Potential
IOR	Inhibition of Return
RT	Reaction Time

Appendix

Abstract (English)

What is the nature of linguistic effects on cognition? We set out to tackle this ancient question by testing if linguistic differences in spatial categorization translate into differences in visual attention allocation. While Korean speakers employ the spatial distinction of loose-fit (*netha*) and tight-fit (*kkita*), German speakers semantically collapse this distinction into a marker for containment (*in*). Here, we used this distinction by examining if Korean speakers get more distracted by a loose-fit or tight-fit shape singleton in the additional singleton paradigm compared to German speakers. In this paradigm participants have to respond to the inclination of a bar (horizontal; vertical) inside of a color-defined target (red; green), while a shape-defined distractor (tight-fit; loose-fit) is present. Indeed, we found that Korean speakers compared to German speakers show an increased RT when a tight-fit or loose-fit shape singleton is present. This result indicates an effect of linguistic differences on the allocation of visual attention. However, it remains subject to speculation if these effects represent shallow online effects or more deep-rooted changes of cognition.

Abstract (Deutsch)

Auf welche Weise prägt Sprache Kognitionen? In dieser Studie beschäftigen wir uns mit dieser historischen Frage, indem wir den Effekt von sprachlich unterschiedlichen räumlichen Einteilungen auf die visuelle Aufmerksamkeit untersuchen. Während im Koreanischen zwischen enger Passung (*kkita*) und loser Passung (*netha*) sprachlich unterschieden wird, geht diese Differenzierung im Deutschen in einer gemeinsamen Präposition auf (*in*). Hier untersuchen wir die Auswirkungen dieser semantischen Unterscheidung, indem wir beobachten, ob sich Koreanischsprachler im Vergleich zu Deutschsprachlern mehr durch einen Ablenkungsreiz in Form eines Stimulus enger bzw. loser Passung im Additional-Singleton-Paradigma ablenken lassen. In diesem Paradigma werden Teilnehmer dazu aufgefordert auf die Orientierung eines Balkens (horizontal; vertikal) in einem farbcodierten Stimulus (grün; rot) zu reagieren während ebenfalls ein formcodierter Ablenkungsreiz (lose Passung; enge Passung) anwesend ist. In der Tat zeigen unsere Ergebnisse, dass Koreanischsprachler in Anwesenheit des formcodierten Ablenkungsreizes, im

Vergleich zu Deutschsprachlern, eine längere Antwortzeit benötigen. Dieses Ergebnis legt einen sprachlichen Einfluss auf visuelle Aufmerksamkeit nahe. Allerdings bleibt zu untersuchen inwiefern dieser Effekte nur einen flüchtigen, oberflächlichen Einfluss oder eine grundlegende, langfristige Veränderung menschlicher Kognitionen darstellt.