CrossMark

Applied Psycholinguistics **39** (2018), 437–459
doi:10.1017/S014271641700042X

Grounded cognition: Comparing Language × Space interactions in first language and second language

DANIELA KATHARINA AHLBERG, HEIKE BISCHOFF, BARBARA KAUP, DOREEN BRYANT, and JESSICA VANESSA STROZYK
University of Tubingen

Received: July 27, 2016 Accepted for publication: June 1, 2017

ADDRESS FOR CORRESPONDENCE
Jessica Strozyk, Department of Psychology, University of Tübingen, Schleichstrasse 4, 72076, Tübingen, Germany. E-mail: jessica.strozyk@uni-tuebingen.de

ABSTRACT
The experiential traces account claims that language comprehension in one's first language (L1) is based on the reactivation of experiential traces that stem from experiencing the corresponding objects, states, or events. However, it remains unclear to what extent this is transferable to second language (L2) comprehension. In the present study, we compared German L1 speakers with German L2 speakers whose L1 uses similar or different spatial terms as German. In an adaptation of the Stroop paradigm, participants were instructed to respond to the font color of German spatial prepositions (e.g., *auf* "on," *über* "above," and *unter* "under/below") by either an upward or a downward hand movement, resulting in compatible or incompatible responses. We found significant compatibility effects for all speakers, but also clear differences between speakers of different L1s. The results thereby support the assumption that experiential traces built during L1 usage play an important role in L2 processing.

Keywords: embodied cognition; experiential traces; second language processing; spatial prepositions

A particular linguistic skill that seems very challenging to acquire in a second language (L2) is the use of spatial terms, in particular spatial pre- or postpositions (e.g., Alonso, Cadierno, & Jarvis, 2016; Bryant, 2012; Coventry & Garrod, 2004; Coventry, Guijarro-Fuentes, & Valdés, 2012; Grießhaber, 1999; Ijaz, 1986; Lütke, 2008; Munnich & Landau, 2010). While the acquisition of syntax and the case system in the L2 is similar to first language (L1) acquisition, the acquisition of the genus system and prepositions differs between L1 and L2 (Kaltenbacher & Klages, 2006). Especially in the case of spatial prepositions, persistent learning difficulties are observable even for children with an early age of acquisition (AoA; Bryant, 2012). Children tend to use the wrong prepositions in different configurations, as for instance using the word *auf* ("on") instead of *über* ("above") as in the following expression from a 9-year-old Turkish child, who started learning German at the age of 4: "*Die Teekanne ist auf dem Ofen*" ("The teapot is on the oven") in response to a picture showing a teapot on a shelf above an oven (Bryant, 2012). It is also important to note that the later the L2 is learned, the greater the differences between L1 and

© Cambridge University Press 2017 0142-7164/17

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

L2 acquisition of German become, and therefore, L2 learners experience greater difficulties in acquiring the L2 with increasing AoA (Hyltenstam & Abrahamson, 2003; Long, 1990; Rösch, 2001).

Spatial pre- or postpositions are part of the fine-structural level of a language (Talmy, 1983), and they allow speakers to describe where objects are in space, and listeners to find objects based on simple locative descriptions in which one object is located in relation to a reference object (Coventry, Prat-Sala, & Richards, 2001). Communicating by means of these spatial terms can be seen as a basic skill of a competent speaker, as spatial descriptions pervade our lives and occur in a wide range of contexts, from locating objects, to reasoning about the world, to understanding the concept of place (Coventry & Garrod, 2004). Although spatial pre- or postpositions originally exclusively encoded spatial relations between physical entities, over time these spatial terms have also developed a rich set of nonspatial meanings (Tyler & Evans, 2003) and are used in more abstract and metaphorical configurations, for instance, when talking about the temporal domain (e.g., the future lies before us, while the past is behind us), the quantitative domain (e.g., more is up, while less is down), or when talking about mood (e.g., happy is up, while sad is down; Coventry & Garrod, 2004; Cuyckens & Zawada, 2001; Tyler & Evans, 2003).

One possible reason for the learning difficulties of spatial pre- or postpositions may be that the source language and the target language differ with respect to the nature of spatial categorization, and learners therefore run into difficulties when trying to transfer aspects relating to spatial categorization from their L1 to the L2. This concept of transfer or cross-linguistic influence is seen as a multifaceted phenomenon (Jarvis & Pavlenko, 2008) accepted by the majority of linguists. However, some researchers (e.g., Krashen, 1983; Newmark, 1966) claim that influences of L1 on L2 are merely temporary communicative strategies due to knowledge gaps in L1 and not due to interference. Accordingly, they declare this phenomenon to be of marginal interest for L2 acquisition. A main theory in transfer research is the contrastive analysis hypothesis developed by Fries (1945) and Lado (1957) stating that similar structures in L1 and L2 lead to a positive transfer with an easy acquisition, whereas structural differences cause acquisition difficulties (negative transfer). While a strong version of the contrastive analysis hypothesis could not be maintained because of a large amount of empirical counterevidence (e.g., Hernández-Chávez, 1972; Hyltenstam, 1977; Kleinmann, 1977), a weaker version of it is overall well received.

Phenomena of linguistic transfer have been observed in all language areas, including phonology (Leather, 2003), morphology (Collins, 2002), and syntax (Schwartz & Sprouse, 1996; White, 2003). The phenomenon we discuss in this paper relates to language-mediated concepts and belongs therefore to conceptual transfer. As Jarvis and Pavlenko (2008, p. xii) point out, "Languages differ with respect to the aspects of reality that they give prominence to. These differences have implications for how speakers of different languages perform both verbal and nonverbal tasks in either their native or a later-learned language."

Languages often differ in the way in which they map linguistic terms onto spatial relations (Coventry & Garrod, 2004). Thus, the way the L2 carves up the world into spatial categories is often quite different from the way the L1 carves up the

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

world (Becker & Carroll, 1997). According to the *thinking-for-speaking* hypothesis (Slobin, 1996, p. 89), "[. . .] each native language has trained its speakers to pay different kinds of attention to events and experiences when talking about them." The training of spatial categorization and its labeling begins during childhood: at the early age of 14 to 20 months children begin to talk about space, and they do so according to the conventions of the target language (Choi & Bowerman, 1991). According to Slobin (1996), once a category is set, it shows a strong resistance to a posterior restructuring. As a consequence, acquiring L2 prepositions should be more difficult for learners whose L1 exhibits fewer or different spatial categories (Coventry et al., 2012). Languages like German, English, or Russian, for instance, differ from languages like Turkish and Korean with respect to the upper subspace. German, English, and Russian distinguish between locations at the boundary between reference object and relatum (*auf* "on") and locations in the upper subspace (*über* "above"). If a reference object has contact with the upper surface of the relatum, a boundary relation is used (Becker & Caroll, 1997). In contrast, a reference object in Turkish or Korean can be located in the upper subspace regardless of contact (Becker & Carroll, 1997; Munnich, Landau, & Dosher, 2001). Therefore, with regard to the vertical axis, the feature "contact" in those languages is irrelevant. If categories are not set in L1, it is possible that the learners do not notice them in the L2 input. In line with this idea, Bryant (2012) showed that learners of German as L2 with Turkish as L1 detected the frequent *auf* ("on") as the sound string for their L1 concept of "upper subspace" and consequently used *auf* in "on" as well as in "above" configurations.

In research on the psychology of language, the theory of experiential traces (Zwaan & Madden, 2005) has received much attention during the last decade. According to this theory, bodily experiences are a crucial part of language comprehension. This account has gained a high degree of support in L1 research, and we now want to investigate whether it also applies to L2 processing. In particular, we believe that this account will contribute to a better understanding of the specific difficulties that L2 learners face when trying to acquire the meanings of German spatial prepositions. The theory of experiential traces (Zwaan & Madden, 2005) is part of the rather new framework of embodied cognition in which language is seen as tightly linked to perception and action (Barsalou, 1999, 2008). According to this theory, during every interaction with the world, a trace resembling the actions and perceptions of this experience is generated. These traces are later reactivated when accessing the same concept. In that sense, language comprehension is the reactivation of experiential traces. More specifically, when linguistic labels often co-occur with certain objects, situations, and events in early childhood, experiential traces stemming from hearing a word get associated with the experiential traces stemming from experiencing its referent. Thus, a word such as *dog* gets associated with experiential traces stemming from hearing, smelling, seeing, or touching a dog. Later on, when the word is processed, these traces become reactivated and presumably provide the basis for comprehension. Similar assumptions can be made for spatial prepositions. When a small child hears a spatial preposition such as *on* or *under*, this will often be in a situation where someone (possibly the child itself) places a reference object into a particular spatial location (e.g., *Look, I am placing the plant on the table*) or draws attention to a particular spatial relation in the

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

context (e.g., *Look, the cat is under the chair*). Thus, with years of co-occurrences of this type, spatial prepositions will become associated with a rich set of experiential traces that stem from experiencing reference situations that can be described by means of these prepositions. These experiential traces then presumably become reactivated when the spatial prepositions are encountered again, and provide the basis for comprehension.

The theory of experiential traces has gained a lot of support in the field of L1 processing (e.g., Glenberg & Kaschak, 2002; Richter, Zwaan, & Hoever, 2009; for an overview, see Jirak, Menz, Buccino, Borghi, & Binkofski, 2010, or the contributions in Coello & Fischer, 2016; Fischer & Coello, 2016). For example, Lachmair, Dudschig, De Filippis, de la Vega, and Kaup (2011) investigated the activation of location information using words whose referents are typically found in the upper vertical space (e.g., *airplane*) or lower vertical space (e.g., *worm*). In one of their experiments, the original Stroop paradigm (Stroop, 1935) was adapted to investigate the processing of these words. In particular, the native German participants in this experiment responded to up and down words with upward or downward movements depending on the font color. Reading the words or accessing their meaning was not required by the experimental task. Nevertheless, a strong compatibility effect could be observed with shorter response times in compatible conditions (e.g., *airplane* and upward response) compared to incompatible conditions (e.g., *airplane* and downward response), suggesting that experiential traces became automatically activated when processing individual words. Spatial compatibility effects of this sort have proven to be extremely stable. They were observed no matter whether the task required lexical access and no matter whether participants noticed the manipulation. They were also observed for verbs (e.g., *rising* vs. *falling*) and for masked stimuli. Finally, these spatial compatibility effects were also observed in tasks focusing on attention shifts to the upper or lower space, as well as in tasks involving the planning of eye movements toward upper or lower locations (for an overview, see Kaup, de la Vega, Strozyk, & Dudschig, 2016). In addition to the above-mentioned studies showing the automatic activation of sensorimotor processes in language comprehension, there are also studies showing that sensorimotor processes can foster comprehension skills. For instance, Glenberg, Brown, and Levin (2007) found that manipulating objects mentioned in a text could improve text comprehension in 6- to 8-year-old children. Compared to children who only reread the text, children who manipulated the relevant objects were able to answer questions about the text with higher accuracy and could even draw inferences. This supports the view that learning through different senses and connecting these learning experiences to linguistic expressions can foster learning and language comprehension abilities.

The results discussed so far focused on processing and acquiring an L1. While this research supports the view that a network of experiential traces is built during L1 acquisition, what happens when people learn and use an L2 remains unclear. However, whether the theory of experiential traces is transferrable to L2 comprehension certainly is an important question. Theoretically, one possibility would be that the terms of the L2 get integrated into the already existing network of experiential traces that was built up during L1 acquisition. Alternatively, the L2 might stay separated from the L1's experiential network and gradually get connected

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

to experiential traces obtained during the use of the L2. These two possibilities seem to fit well with considerations from the literature on L2 acquisition. Some models of L2 acquisition propose that the semantic representations of L1 and L2 are shared across languages (e.g., the revised hierarchical model; Kroll & Stewart, 1994; the bilingual interactive activation plus model; Dijkstra & Van Heuven, 2002). Applied to the experiential traces account, these models would predict that embodiment effects should be observed in both L1 and L2 processing. Other models argue that semantic representations in the L2 are partially overlapping with L1 representations but may be less detailed (see the sense model; Finkbeiner, Forster, Nicol, & Nakamura, 2004). This would lead to the prediction that embodiment effects in L2 speakers are reduced or absent, at least when language proficiency is still rather low and learners have little everyday experience with using the language (De Grauwe, Willems, Rueschemeyer, Lemhöfer, & Schriefers, 2014).

There is a growing body of research investigating embodiment effects in L2 learners (e.g., Bergen, Lau, Narayan, Stojanovic, & Wheeler, 2010; De Grauwe et al., 2014; Dudschig, de la Vega, & Kaup, 2014). De Grauwe et al. (2014) found evidence for embodiment for L1 as well as for L2 speakers on a neural basis. In their functional magnetic resonance imaging study, the participants, Dutch native speakers and German native speakers who learned Dutch as an L2, did a lexical decision task on visually presented Dutch motor and nonmotor verbs. The participants had to respond to pseudowords (i.e., pronounceable words that do not carry a meaning, e.g., *ralmen*) and withhold their responses to actual words allowing the authors to measure brain activity while the participants were passively reading the words. After the main experiment, the participants rated the words for motor-relatedness. Both L2 and L1 speakers showed similar increases of activation in the motor and sensory-motor brain areas for motor compared to nonmotor verbs, for both cognates (orthographically similar words depicting the same meaning in the two languages, e.g., *nemen* in Dutch and *nehmen* in German) and noncognates (orthographically dissimilar words depicting the same meaning in the two languages, e.g., *gooien* in Dutch and *werfen* in German). This result is interesting, as it shows the involvement of the motor cortex during motor word comprehension and thus embodiment of these words in L2 speakers and thereby supports the view of the revised hierarchical and bilingual interactive activation plus models. In addition, Dudschig et al. (2014) were also able to find evidence for embodiment effects in L2 speakers of English with German as L1 in a behavioral paradigm. They used the Stroop paradigm from Lachmair et al. (2011) to investigate whether L2 words, referring to objects typically located in the upper or lower vertical space, would automatically activate location-specific experiential traces and thus facilitate upward or downward responses, respectively. Their participants saw L2 English words like *sun* and *root* presented in different colors and got the instruction to respond only to the font color with the corresponding upward or downward movement. The results revealed a typical compatibility effect, with faster responses for compatible trials (e.g., upward response to *sun*) compared to incompatible trials (e.g., upward response to *root*). This study thereby showed that a reactivation of experiential traces also takes place in L2 processing.

The studies mentioned so far mainly focused on the processing of words that exist in both L1 and L2. However, many words do not have a direct equivalent in

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

different languages (Bowerman, 1996; Stringer, 2010), as discussed above for the spatial prepositions *auf* and *über*. The question now arises whether words of the L2 that have no direct equivalent in the L1 are integrated into the already existing network of experiential traces of L1, and if so, how exactly this is accomplished. In the present study, we investigated how spatial prepositions are processed by learners of German as an L2 as well as by German native speakers. We were particularly interested in comparing embodiment effects related to the processing of spatial prepositions in German native speakers with the embodiment effects potentially observed in different groups of L2 learners. For that reason, we investigated German native speakers and compared them to L2 speakers whose native language uses similar spatial terms as German with respect to the upper part of space (English and Russian) and L2 speakers whose native language uses different spatial terms than German (Turkish and Korean). The English prepositions *on* and *above*,[1] and the Russian prepositions *на* (*na*) and *над*[2] (*nad*) can be seen as near equivalents to *auf* and *über*. In contrast, Turkish uses either *üstünde* or *üzerinde*,[3] which are used interchangeably, and Korean uses only 위[4] (*wi*) for both spatial configurations plus or minus (±) contact in the upper subspace. One additional difference is that, while English and Russian use prepositions, the above-named terms in Turkish and Korean are handled as postpositions (Becker & Carroll, 1997; Munnich et al., 2001).

Based on Bryant's (2012) studies and on Slobin's (1996) statement about the rigidity of a category once it is set, our hypothesis was that newly learned words would be connected to the preexisting experiential traces of the L1, and that the nature of the embodiment effects found in the L2 would, therefore, depend on the specific L1 of the participants. More specifically, in our study we focused on the prepositions *auf* and *über*, as we predicted differences in the processing of these spatial terms in German native speakers and L2 learners of German depending on the nature of their L1. We presented these terms in an adaptation of the Stroop paradigm, similar to the one used by Lachmair et al. (2011) and Dudschig et al. (2014), in which participants were required to respond to the words by an upward or downward movement of the arm depending on the font color of the stimuli. This setup resulted in compatible (e.g., upward response to *auf*) and incompatible (e.g., upward response to *unter*) trials. For German native speakers, we expected larger compatibility effects for *über* compared to *auf*, because *über* implies a larger distance between theme and relatum on the vertical axis. In addition, while both *auf* and *über* can be used nonspatially in German, nonspatial usage is more common for *auf* (e.g., *Ich freue mich auf die Party* "I am looking forward to the party"; *aufräumen* "to tidy up"). For L2 speakers of German, we predicted different results depending on their particular L1. We expected the English and Russian L1 speakers to show a similar processing difference between *auf* and *über* as the German native speakers, because the split of the upper subspace into ± contact is also present in English and Russian. For the Turkish and Korean L1 speakers, we predicted a different pattern. These speakers should show stronger effects for *auf* in comparison to *über*. In their L1, the category split of the upper subspace into ± contact is not present, but as *auf* is much more frequent in the German input,[5] we expected them to transfer their experiential traces to this term. This prediction is in line with the above-mentioned finding that Turkish learners of German used *auf* for all instances

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

in which they would have used *üstünde* or *üzerinde* in their native language (Bryant, 2012). If we observe different effects depending on the L1 of the participants, this would also fit well with coactivation accounts of bilingualism (e.g., Blumenfeld & Marian, 2013; Kroll, Bobb, & Hoshino, 2014), which assume that in bilinguals both languages, L1 and L2, are always active during language processing. The speaker needs to decide which language to use and then needs to suppress the not needed language. Applied to our hypothesis, this would mean that the experiential traces related to corresponding L1 terms always get coactivated when participants are processing L2 terms. As a result, different groups of participants with different L1s should show different results during L2 processing.

To summarize, we presented three groups of participants (German native speakers, German L2 speakers with Russian or English as their L1, and German L2 speakers with Korean or Turkish as their L1) with four German words, the three prepositions *auf*, *über*, and *unter* related to our hypotheses and the filler item *ab*. Participants responded to the font color of the words with an upward or downward arm movement, and we expected the compatibility effect to differ in dependence of the native language of the participants.

METHOD

Participants

Overall, 149 people participated in this study in exchange for course credit or financial reimbursement. All participants were students or employees of the University of Tübingen, had normal or corrected-to-normal vision, and gave informed consent prior to participation.

Forty-nine of these participants were German native speakers (38 female, 11 male; age: $M = 22.9$ years, $SD = 4.4$ years). Forty-eight participants had German as an L2 and an L1 with a similar split of the upper subspace into ± contact as German (similar-L1 group). Twenty-four of them had English as their L1 and 24 had Russian as their L1. Three participants of this group needed to be excluded, as they were German/English bilinguals. They had learned German from their parents or other family members before entering kindergarten. Of the remaining 45 participants in this group, 31 were female and 14 were male (age: $M = 25.9$ years, $SD = 4.5$ years). The last group of participants consisted of 52 people with German as an L2 and an L1 that does not split the upper subspace into ± contact (dissimilar-L1 group). Twenty-four of them were native Koreans and 28 had Turkish as their L1. Nine of the Turkish native speakers had to be excluded, as they were German/Turkish bilinguals. They had learned German from their parents or other family members before they entered kindergarten. Of the remaining 43 participants in this group, 35 were female and 8 were male (age: $M = 24.7$ years, $SD = 5.8$ years).

The similar-L1 group started learning German between 8 and 28 years of age (AoA: $M = 16.0$ years, $SD = 5.7$ years); the dissimilar-L1 group started learning German between 3 and 26 years of age (AoA: $M = 14.5$ years, $SD = 6.7$ years). The AoA did not differ significantly between the groups, $t\,(85) = 1.08$, $p = .282$.[6] Most of the L2 participants learned at least one additional language other than

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Table 1. *Distribution of age of acquisition and language proficiency (number of participants per category)*

	Similar-L1 Group	Dissimilar-L1 Group
Age of acquisition		
Early (3–6 years)	0	10
As a child (7–12 years)	16	2
As teenager or adult (>12 years)	28	31
Language proficiency CEFR levels		
A2	2	5
B1	3	7
B2	7	15
C1	16	4
C2	17	12

Note: L1, first language; CEFR, Common European Frame of Reference. The stages of language acquisition were adapted from Klein (1992); Language proficiency was measured on a 6-point scale referring to the CEFR levels (Verhelst et al., 2009). When participants judged themselves as being between two CEFR levels, we categorized them as the higher one.

their native language and German. Except for one Russian-Ukrainian bilingual,[7] none of them started learning the additional languages before the age of 5 years. Therefore, we can safely assume that the system of experiential traces had already been established in the native language before the participants started learning another language.

All participants reported to speak German at least at level A2 of the Common European Frame of Reference (CEFR; Verhelst, Van Avermaet, Takala, Figueras, & North, 2009; for a detailed description of the CEFR, see below). The similar-L1 group (proficiency: $M = 4.9$, $SD = 1.1$) had a higher proficiency than the dissimilar-L1 group (proficiency: $M = 4.2$, $SD = 1.4$), $t\ (86) = 2.64$, $p = .010$; however, on average both groups were categorized as independent language users (cf. description below). For an overview of how many participants fell into each AoA and proficiency category, see Table 1.

Materials and apparatus

Three German words served as stimuli, namely, *über*, *auf*, and *unter*. Hereby, *auf* and *über* served as referents for the upper dimension, and *unter* served as referent for the lower dimension. The preposition *unter* often comes to mind as a counterpart for the preposition *über* (*über und unter*). For counterbalancing purposes, we wanted to have counterparts for both *über* and *auf*. We therefore included the word *ab* ("off/down") as a filler word in our experimental task because this word often comes to mind as a counterpart for the preposition *auf* (*auf und ab*). It is important to note however that the particle *ab* is not a spatial preposition

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

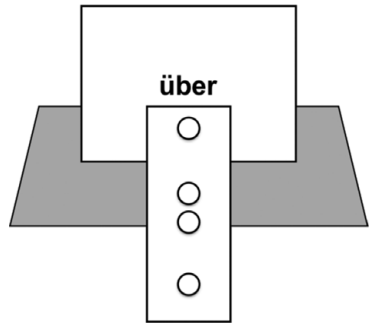


Figure 1. Experimental setup. The keyboard was implemented under a vertical plane in front of the participants. At the beginning of each trial, the participant pressed the two middle keys, one with the left hand and the other with the right hand. A response was made by releasing the correct middle button (upper middle button for an upward response, lower middle button for a downward response), pressing the upper or lower button, and returning back to the released middle button, while the other hand rested on the respective middle button.

but mostly combines productively with verbs without expressing the meaning of "down." Consider, for example, *abwischen* ("to wipe"), *ablehnen* ("to reject"), and *abmachen* ("to arrange"). Its only spatial usage is as part of the directional adverb *abwärts* ("downward"), which probably also explains why it comes to mind as a counterpart of *auf*. We thus do not include the word *ab* in our analyses.

The four words were presented in four different font colors: blue (RGB 0, 0, 255), orange (RGB 255, 128, 0), lilac (RGB 150, 0, 255), and brown (RGB 140, 80, 20). Each word appeared equally often in each color. For each participant, two colors were associated with upward responses, and two were associated with downward responses. By using four colors instead of two, we made the task more complex, which results in overall slower response times. We decided to use this common practice for language experiments using adaptations of the Stroop paradigm (e.g., Lachmair et al., 2011; Dudschig, de la Vega, & Kaup, 2015) because we wanted to make sure that the participants had enough time to process the word before responding. When using only two colors, responses might have been initiated before the meaning of the word had been accessed (cf. the conflict monitoring theory; Botvinick, Braver, Barch, Carter, & Cohen, 2001). Responses were recorded using a PS/2 computer keyboard adapted with a locally constructed overlay (Figure 1).

Participants other than German natives received a short questionnaire after the main experiment assessing their language background and proficiency. They were asked how they learned German (e.g., at school), at what age they started to learn, and to rate their language proficiency on a 6-point scale referring to the levels of the CEFR (Verhelst et al., 2009). According to the CEFR, language proficiency is divided into six categories: A1, A2, B1, B2, C1, and C2.[8] The two A levels refer to the learner as a basic user, who interacts with natives in a very simple way and can use basic expressions and later on phrases or describe routines. The B levels see the learner as an independent user who can understand main points

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

of standard conversations and later on texts. At this stage, he becomes able to interact with natives in a more spontaneous and more fluent way. The C levels describe a proficient user, who can use the language flexibly and effectively for social, academic, and professional purposes. In the latest stage, the user virtually understands everything that is heard or read, and even is able to use idiomatic expressions (Verhelst et al., 2009). The CEFR is widely used in Germany, and language courses are usually categorized according to these levels. Therefore, we can safely assume that participants were able to make this judgement accurately.

Procedure and design

Each trial started with a fixation cross, displayed in the center position of the screen for 1000 ms. Afterward the stimulus was presented at the same position as the fixation cross until the participant responded by releasing one of the middle buttons. After button release, a blank screen was shown until the corresponding upper or lower button was pressed. Between trials, a white screen was shown for 1000 ms.

Participants were instructed to push down the two middle buttons of the response box (see Figure 1) at the beginning of each trial and to keep their hands in this position until responding. One half of the participants used the right hand for the upper middle button and the left hand for the lower middle button. For the other half of the participants, this mapping was reversed. All participants were instructed to respond to the font color of the stimuli, as quickly and accurately as possible, by means of an upward or downward arm movement. That is, participants had to release the respective middle button and press the upper or lower button, depending on the color of the presented word, before returning to the middle button. The upper and the lower buttons were each associated with two of the four possible colors. This mapping of colors to response direction was balanced across participants. All possible color pairs occurred equally often. Unlike in the traditional Stroop task, the response direction that the font colors indicated, and not the font color itself, was defining the compatibility conditions (compatible condition; e.g., upward response to *über*; incompatible condition: e.g., downward response to *über*). This is a common procedure in many Stroop-like experiments nowadays (for an overview, see MacLeod, 1991).

The mapping of colors to response direction was explained to the participants in the beginning of the experiment and was practiced in a short block of 16 stimuli, each presented twice, which were not used during the experimental blocks. In contrast to the experimental blocks, the participants received feedback about response accuracy during this practice block. The instructions, including the information on color assignment, were only given at the beginning of the experiment and were not repeated throughout the experiment.

Following the practice block, participants completed 4 experimental blocks during which each of the words *auf*, *über*, *unter*, and *ab* was repeated 80 times, resulting in a total number of 320 trials. Since *ab* was used as a filler item (see above), this resulted in a 3 (stimulus: *auf*, *über*, *unter*) × 2 (response direction: upward vs. downward) × 3 (group: German natives, similar-L1 group, dissimilar-L1 group) mixed design with stimulus and response-direction as within-subjects factors and

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

group as a between-subjects factor. The dependent variables were the accuracy rate and the release time of the middle button. To analyze the data, we performed mixed analyses of variance (ANOVAs) with repeated measures on the within-subjects factors. In addition, we analyzed the compatibility effect on reaction times for each stimulus and in each group separately.

RESULTS

Three participants, two German natives and one Korean native, were excluded from the analysis due to an error rate[9] above 15%. The remaining participants committed an error on 4.2% of the trials. These trials were excluded from the analysis of response times. Furthermore, release responses faster than 200 ms and slower than 3000 ms were excluded, as well as responses deviating by more than 3 *SD* from the mean of each cell (Participant × Stimulus × Response). This outlier elimination reduced the data set by 1.8%.

Accuracy rates

Upward responses were slightly more accurate (96.1%) than downward responses (95.6%). However, this effect failed to reach significance, $F\,(1, 131) = 3.39$, $p = .068$, $\eta_p^2 = 0.025$. There was a significant interaction between the stimulus and the response direction showing the expected compatibility effect, $F\,(2, 262) = 31.46$, $p < .001$, $\eta_p^2 = 0.194$. For the words *auf* and *über*, upward responses (*auf*: 96.3%, *über*: 96.5%) were more accurate than downward responses (*auf*: 95.4%, *über*: 95.5%), and for the word *unter*, upward responses were less accurate than downward responses (upward: 93.9%, downward: 97.3%). There was no significant three-way interaction, $F\,(4, 262) = 1.04$, $p = .387$, $\eta_p^2 = 0.016$, indicating that the compatibility effect did not differ across groups. All other effects did not reach significance (all ps $> .10$). For the average accuracy rates in each cell (Group × Stimulus × Response direction) see Table 2. We did not include the means for *ab*, since this word was only used as a filler item. However, as expected, accuracy rates for *ab* did not differ between upward (95.8%) and downward responses (96.1%), $t\,(133) = –0.81$, $p = .419$.

Response times

Table 3 shows the average response times for each group, stimulus, and response direction. Just as for the accuracy rate analysis, we did not include the means for *ab*, since it was only used as a filler item. However, as expected, the response times for downward responses (610 ms) and upward responses (617 ms) to *ab* did not differ significantly, $t\,(133) = –1.27$, $p = .206$.

Our ANOVA showed that there was a significant main effect of group, $F\,(2, 131) = 6.56$, $p = .002$, $\eta_p^2 = 0.091$, due to overall faster response times for the German natives (551 ms) than for the other two groups: similar-L1 group, 650 ms, $F\,(1, 90) = 12.68$, $p < .001$, $\eta_p^2 = 0.123$; dissimilar-L1 group, 626 ms, $F\,(1, 87) = 8.25$, $p = .005$, $\eta_p^2 = 0.087$. The two groups with German as an L2 did not differ significantly from each other ($F < 1$). As expected, the interaction between

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Table 2. *Average accuracy rates and corresponding standard deviations in percentages per group*

	German Natives		Similar-L1 Group		Dissimilar-L1 Group	
	M	*SD*	*M*	*SD*	*M*	*SD*
Auf						
Upward response	96.7	3.7	95.9	3.9	96.1	4.1
Downward response	95.3	6.3	95.7	3.9	95.2	5.2
Über						
Upward response	97.1	3.2	96.9	3.1	95.5	4.9
Downward response	95.4	5.2	95.1	4.1	96.1	3.9
Unter						
Upward response	94.7	4.8	93.7	7.0	93.2	6.9
Downward response	97.2	3.5	97.3	2.9	97.5	3.3

Table 3. *Average response times and corresponding standard deviations in milliseconds per group*

	German Natives		Similar-L1 Group		Dissimilar-L1 Group	
	M	*SD*	*M*	*SD*	*M*	*SD*
Auf						
Upward response	548	97	634	152	601	118
Downward response	553	112	667	172	642	166
Über						
Upward response	542	87	633	155	613	128
Downward response	558	111	672	184	635	172
Unter						
Upward response	565	108	674	190	661	176
Downward response	540	112	617	135	606	148

stimulus and response direction was significant, $F\ (2,\ 262) = 47.55$, $p < .001$, $\eta_p^2 = 0.266$, showing the expected compatibility effect of faster response times for upward than downward responses for the words *auf* (upward: 594 ms, downward: 619 ms) and *über* (upward: 595 ms, downward: 620 ms) and the reverse effect for the word *unter* (upward: 632 ms, downward: 586 ms). Most important, and also as predicted, we obtained a significant three-way interaction between group, stimulus, and response direction, $F\ (4,\ 262) = 3.84$, $p = .005$, $\eta_p^2 = 0.055$, indicating that the compatibility effect differed across groups. Therefore, we will analyze this effect in more detail separately for the three groups below. The interaction between group and stimulus did approach but not reach significance, $F\ (4,\ 262) = 2.23$, $p = .067$, $\eta_p^2 = 0.033$. All other effects also did not reach significance (all Fs < 1).

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

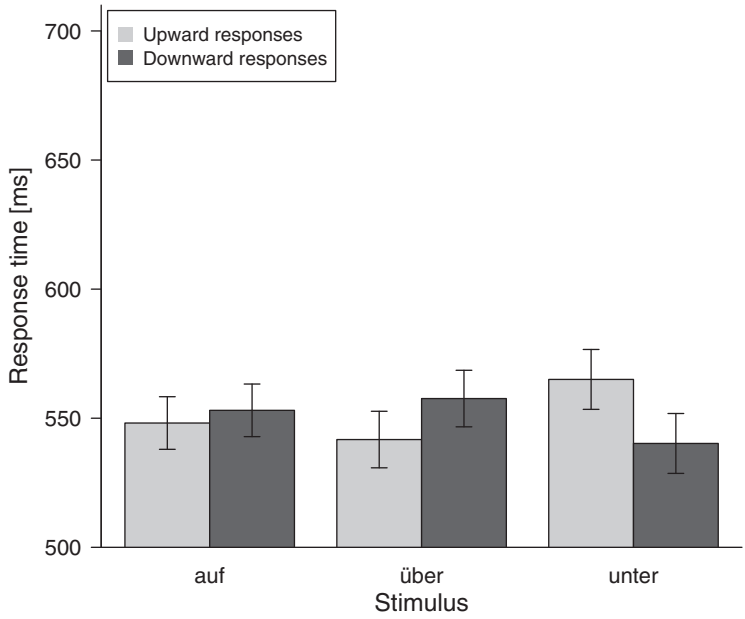


Figure 2. Mean response times of German natives for correct responses as a function of response direction and stimulus. Error bars represent 95% confidence intervals for the pairwise comparisons (according to Loftus & Masson, 1994).

German natives. The average response times for the German native participants are displayed in Figure 2. A separate ANOVA for just this group with stimulus and response direction as within-subjects factors showed that there was a significant compatibility effect, $F\,(2, 92) = 14.38$, $p < .001$, $\eta_p^2 = 0.238$. For the word *über*, upward responses were faster than downward responses, $F\,(1, 46) = 4.23$, $p = .045$, $\eta_p^2 = 0.085$; for the word *unter*, the reverse was true, $F\,(1, 46) = 9.26$, $p = .004$, $\eta_p^2 = 0.168$. There was, however, no significant difference between upward and downward responses for the word *auf*, $F\,(1, 46) = 0.47$, $p = .496$, $\eta_p^2 = 0.010$.

Similar-L1 group. As can be seen in Figure 3, there was a significant compatibility effect in this group, too, represented by a significant interaction between stimulus and response direction, $F\,(2, 88) = 18.71$, $p < .001$, $\eta_p^2 = 0.298$. Separate analyses for the three stimuli revealed significant compatibility effects for all three words with shorter response times for upward compared to downward responses for *auf*, $F\,(1, 44) = 16.28$, $p < .001$, $\eta_p^2 = 0.270$, and *über*, $F\,(1, 44) = 8.96$, $p = .005$, $\eta_p^2 = 0.169$, and shorter response times for downward compared to upward responses for the word *unter*, $F\,(1, 44) = 15.32$, $p < .001$, $\eta_p^2 = 0.258$.[10]

Dissimilar-L1 group. This group also showed a compatibility effect represented in a significant interaction between stimulus and response direction, $F\,(2, 82) = 18.30$, $p < .001$, $\eta_p^2 = 0.309$ (see Figure 4). There were significant compatibility effects

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

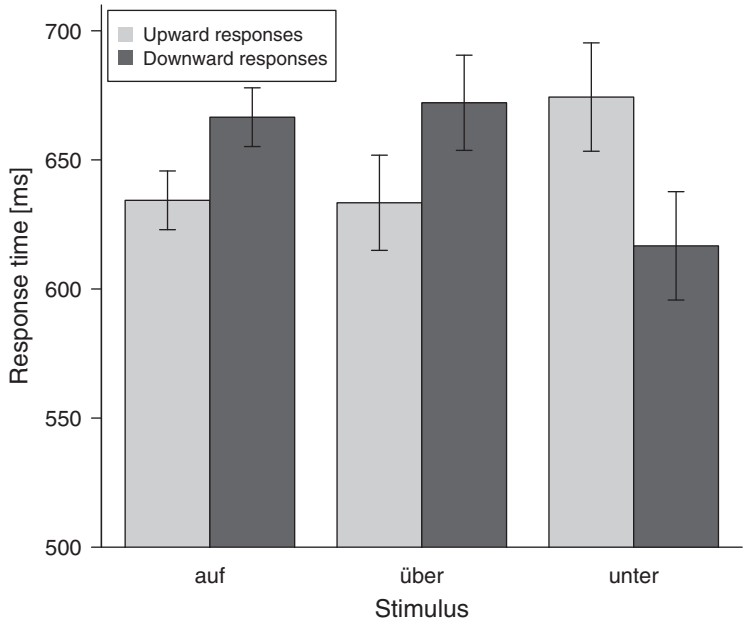


Figure 3. Mean response times of the similar-first language group as a function of response direction and stimulus. Error bars represent 95% confidence intervals for the pairwise comparisons (as per Loftus & Masson, 1994).

for the word *auf* with shorter response times for upward compared to downward responses, $F\,(1, 41) = 7.96$, $p = .007$, $\eta_p^2 = 0.163$, and for the word *unter*, with shorter response times for downward compared to upward responses, $F\,(1, 41) = 12.96$, $p < .001$, $\eta_p^2 = 0.240$. No significant compatibility effect was obtained for *über*, $F\,(1, 41) = 2.56$, $p = .117$, $\eta_p^2 = 0.059$.

DISCUSSION

In this study, we investigated whether embodiment plays a role in L2 comprehension. In particular, we were interested in whether processing German spatial prepositions such as *auf*, *über*, and *unter* would influence subsequent upward or downward directed response movements not only for German native speakers but also for nonnative speakers learning German as an L2. If so, this would support the view that experiential traces (Zwaan & Madden, 2005) are not only built and reactivated in L1 processing but also play a role in L2 processing.

Three groups of participants (German native speakers, German L2 speakers with an L1 with a similar split of the upper subspace as German, and German L2 speakers with an L1 without that split of the upper subspace) processed German prepositions in an adaptation of the Stroop paradigm in which they responded with an upward or downward directed response movement depending on the font color. We found a significant compatibility effect, reflected in an interaction between the

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

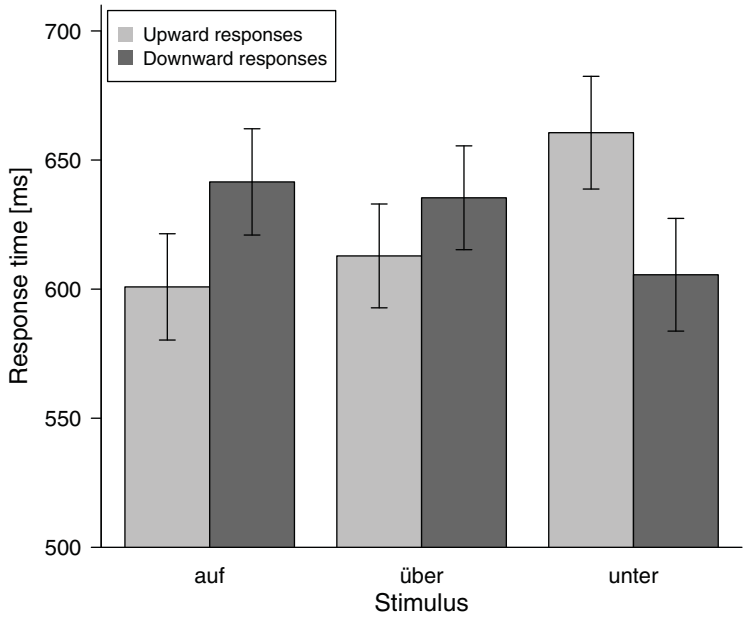


Figure 4. Mean response times the dissimilar-first language group as a function of response direction and stimulus. Error bars represent 95% confidence intervals for the pairwise comparisons (according to Loftus & Masson, 1994).

meaning of the spatial prepositions and the direction of the response movement, for all three groups. Overall, responses were faster and more accurate in conditions in which the meaning of the preposition was compatible to the response direction (e.g., upward movement for *über*; downward movement for *unter*) compared to conditions in which the two were incompatible (e.g., upward movement for *unter*; downward movement for *über*). This result shows that the Stroop paradigm is suitable for the investigation of spatial prepositions, as the observed compatibility effect is similar to the one obtained with nouns referring to entities with a typical location in vertical space (*airplane* vs. *worm*, e.g., Lachmair et al., 2011). In addition, this result implies that experiential traces already get activated when people read small units like prepositions, which can be considered further evidence for embodied theories of comprehension. This finding is also in line with previous research using subliminal presentations of spatial prepositions (Ansorge, Kiefer, Khalid, Grassl, & Koenig, 2010) and with research on the Simon effect in response to spatial words (Khalid & Ansorge, 2013). One might argue that these results could also be explained by assuming that participants internally verbalized the response direction when planning the response in the present paradigm (i.e., verbalizing "upward" when detecting a particular font color) and that this causes the compatibility effect with the meaning of the presented preposition. However, a recent study conducted in our lab indicated that language–space compatibility effects are observed even in a modified Stroop paradigm where there is no stable

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

mapping of colors to response directions, and inner speech can therefore be ruled out as the main contributing factor (Dudschig & Kaup, 2016).[11] We therefore can be quite sure that the compatibility effect observed in this experiment (interaction of stimulus and response direction) constitutes an embodiment effect and reflects the automatic activation of experiential traces during word processing.

Furthermore, that we found compatibility effects not only for the German natives but also for the two groups of German L2 speakers supports the view that experiential traces are being reactivated during L2 processing. Even the dissimilar-L1 group showed a compatibility effect, even though the spatial terms of their L1 cannot be matched directly to the German system during L2 learning. This finding suggests that the reactivation of experiential traces plays an important role not only in L1 processing but also in the processing of an L2, even if this is quite dissimilar to the participants' native language. A relevant question now is whether these experiential traces rely on the L1 or represent a new network of traces associated with the L2. More information with respect to this question can be obtained when comparing the observed pattern of results for the nonnative speakers with that observed for German native speakers.

As predicted, German native speakers processed *auf* and *über* differently, even though both words refer to the upper dimension. It seems that *über* is more strongly connected with spatial experiential traces than *auf*, since for *auf* the attributed spatial dimension neither interfered with nor facilitated the response. One possibility is that these differences reflect variations in the way these prepositions carve up the upper space. While *über* refers to objects that do not touch the relatum, *auf* refers to objects that are in direct contact with the relatum. Therefore, we can expect to find stronger compatibility effects for *über* than for *auf*. It is however surprising, that there was no compatibility effect for *auf* at all.

We had initially predicted that the pattern of results in the similar-L1 group should be the same as for German natives. This was not the case. Whereas German native speakers showed no effect for *auf*, the participants in the similar-L1 group did, suggesting that they processed *auf* spatially. One possible reason for this difference in comparison to German native speakers might be that German L1 speakers first learn *auf* as spatially related, but, as mentioned above, the word *auf* is also frequently used in various contexts in which the spatial meaning is no longer obvious (e.g., *aufhören* "to stop/terminate," *aufmachen* "to open sth," *sich auf etwas freuen* "to look forward to sth."). It therefore appears possible that the spatial meaning fades away throughout the lifetime as people gain more nonspatial experiences with this word, which explains why there was no compatibility effect for *auf* in the German native speakers. However, this hypothesis needs further testing before definite conclusions can be drawn. One possibility would be to compare adult language processing with that of children who are just starting to learn these words and, therefore, have different linguistic experiences than adults. Another possibility would be a control study where the spatial meaning would be clearly triggered, for example, by using prepositional phrases. A possible reason why the L2 speakers with a Russian or English background showed compatibility effects for both *auf* and *über* is that the equivalent for *auf* in their native languages is used spatially to a stronger degree than *auf* in German. If they transferred the experiential traces of their L1 onto the respective German preposition, this might

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

explain the significant compatibility effect for *auf*. This interpretation would also be in line with the coactivation account (Blumenfeld & Marian, 2013; Kroll et al., 2014) according to which bilinguals always activate both languages (see above). However, at least in English *on* also has multiple nonspatial usages. Thus, clearly more research is needed that investigates this post hoc explanation of the results before any definite conclusions can be drawn.

For the dissimilar-L1 group, we had expected a different processing pattern than for German L1 speakers, since their native languages, Korean and Turkish, do not distinguish between contact and noncontact along the vertical axis, and are therefore not directly connectable to German words when learned. We found clear processing differences between this group and the other two groups. As expected, Korean and Turkish native speakers did not show a larger effect for *über* compared to *auf*, which presumably reflects that their L1s do not distinguish between the two corresponding spatial relations, and they, therefore, do not retrace this distinction in the German input. These speakers not only did not show a larger effect for *über* compared to *auf* (as the German natives did), but quite the contrary was the case: they showed a significant compatibility effect for *auf* but not for *über*. We consider it likely that this difference is related to *auf* being much more frequent compared to *über* (Quasthoff, Fiedler, & Hallsteinsdóttir, 2011) and also acquired much earlier (Bryant, 2012; Grimm, 1975). Concerning the different word formation products with the respective morphemes, many more words with the prefix *auf-* (5,591 word form entries) exist than with the prefix *über-* (4,065 word form entries; see Quasthoff et al., 2011). These frequency differences might account for the reversal of the effect, as the L2 speakers search for the equivalent of their spatial terms related to the upper dimension and the first spatial term they learn is *auf*, which is then taken as the equivalent. *Über*, which is learned later, is then harder to connect, in particular because the category of spatial relations needs a split-up in this case. Korean and Turkish native speakers probably simply transferred all their experiential traces related to the upper vertical space to the more frequent term *auf*, which they learned first, and more or less ignored the less frequent *über*. In that case, *über* would not be processed in an embodied way, at least not at early stages of learning. If this assumption is actually true, one might expect to find differences between participants with higher and lower language proficiency in German. Possibly, *über* leads to compatibility effects only for highly proficient L2 speakers of German because they either transfer the experiential traces of their L1 also to *über* or with enough experiences they build new traces for their L2. To tentatively investigate this question, we included language proficiency as an additional factor in the analysis of this group. We categorized the participants into two groups according to their language proficiency. Speakers with a CERF level of A2, B1, or B2 were considered as mid- to low- proficient; speakers with a CERF level of C1 or C2 were considered highly proficient. The analysis revealed a significant interaction between stimulus, response direction, and language proficiency, $F(2, 78) = 3.35$, $p = .040$, $\eta_p^2 = 0.077$, indicating that a difference exists between the high and mid- to low-proficiency groups. In order to gain more information with respect to our hypotheses concerning the role of language proficiency, we analyzed the compatibility effects for the three stimuli and the two proficiency groups separately. For the mid- to low-proficiency group, *auf* showed a signifi-

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

cant compatibility effect, t (26) = –2.26, p = .016, one-tailed, while for *über* no significant compatibility effect was obtained, t (26) = 1.06, p = .149, one-tailed. In the high-proficiency group, *auf* as well as *über* showed significant compatibility effects: *auf*, t (14) = –1.77, p = .049, one-tailed; *über*, t (14) = –1.91, p = .038, one-tailed. These findings fit with the assumption that *über* is acquired later than *auf* also in L2 acquisition. However, these results have to be interpreted with caution, since they were obtained with a small sample of low power leading only to significant results when tested one-tailed, and because this explanation is quite post hoc. More research is needed to investigate these points in detail.

Taken together, we found evidence for embodiment in L1 and in L2 processing. Experiential traces were reactivated in all tested groups. Furthermore, our results indicated that the observed processing differences between the groups depend on the learned native language, as hypothesized by Slobin (1996). In addition, our results can also be taken as evidence for and an extension of the coactivation account (Blumenfeld & Marian, 2013; Kroll et al., 2014), by showing that not only phonological or orthographical information relating to the L1 is activated during L2 processing but also the experiential knowledge related to the L1 as well.

Our study provides a good starting point for further investigations into the nature of embodiment effects during L2 processing. Future studies are needed to find out whether other factors besides the nature of the native language are relevant. We found first evidence that language proficiency might be important, as was also suggested by De Grauwe et al. (2014) and Vukovic (2013). Further variables that could be relevant but that were not tested in the current study are AoA, language contact, and conditions of language use. Furthermore, it would be interesting to investigate the effects of language proficiency using a more objective measure compared to the rather subjective indication of language proficiency that we used in the present study. Once we know more about how experiential traces are built and used during L2 acquisition and which moderating factors are at work, we might be able to gain a better understanding of the precise difficulties to learn spatial prepositions in an L2.

ACKNOWLEDGMENTS

This research was supported by the LEAD Graduate School, a project of the Excellence Initiative of the German federal and state governments under Grant GSC1028. This study was part of the doctoral dissertation of the first author conducted at the LEAD Graduate School.

NOTES

1. With regard to frequency, *on* occupies rank 17, *above* occupies rank 896. An alternative expression of *above* is *over* with rank 124 (Davies, 2008).
2. *На* (*na*) occupies frequency rank 4; *над* (*nad*) occupies rank 181 (Sharoff, Umanskaya, & Wilson, 2013).
3. Although both words are used interchangeably, *üstünde* with Rank 647 is less frequent than *üzerinde* with Rank 92 (Aksan et al., 2012).

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

4. 위 (*wi*) occupies Rank 119 in word frequency. Although in the above stated literature *wi* is seen as a postposition, it is also referred to as a common noun (National Institute of Korean Language, 2005). An alternative expression would be the verb *nohta* ("put on"; Choi & Hattrup, 2012).
5. *Auf* occupies Rank 11 of all words (frequency = 8.0%) whereas *über* is at Rank 45 with a frequency of 2.7%. *Auf* is used four times more often than *über*. Considering only prepositions, *auf* occupies Rank 5 and *über* Rank 12 (see Quasthoff et al., 2011).
6. One participant in the similar-L1 group did not provide a year or an age at which she started learning German and is therefore not included in this analysis. She reported to have started learning German in school and does therefore meet our inclusion criterion of not being bilingual.
7. Russian and Ukrainian both use *на* (*na*) and *над* (nad) for "above" and "on," which is why we did not exclude that person.
8. We coded these levels from 1 to 6 in order to be able to analyze group differences. For participants who rated themselves as being in between two levels, we used the average score of these two levels (e.g., someone rating himself as being between B1 and B2 was coded as 3.5).
9. We coded a trial as being incorrect when the incorrect middle button was released and/or when the incorrect upper or lower button was pressed.
10. Because the English *under* and the German *unter* are cognates, while the Russian *под* and German *unter* are noncognates, it would have been possible that Russian L1 speakers and English L1 speakers differ in their responses to this word. However, there were no significant differences between the English and Russian L1 speakers with $F < 1$ in the general analysis as well as in a separate analysis for *unter* only.
11. More specifically, in this experiment, participants saw words in the center of the screen in four different font colors, as well as four colored rectangles, located above, below, left, and right of the word stimulus. The participants' task was to make a movement in the direction of the rectangle that matched the font color of the word stimulus, whereby the location of the colored rectangles randomly changed from trial to trial. Participants could therefore not facilitate their response by memorizing a rule such as "red is upward," and it is therefore highly unlikely that inner speech is responsible for the observed compatibility effects.

REFERENCES

Aksan, Y., Aksan, M., Koltuksuz, A., Sezer, T., Mersinli, Ü., Demirhan, U. U., . . . Yıldız, i. (2012). *Construction of the Turkish National Corpus (TNC)*. Paper presented *at* the 8th International Conference on Language Resources and Evaluations, Istanbul. Retrieved from http://www.lrec-conf.org/proceedings/lrec2012/pdf/991_Paper.pdf

Alonso, R., Cadierno, T., & Jarvis, S. (2016). Cross-linguistic influence in the acquisition of spatial prepositions in English as a L2. In R. Alonso (Ed.), *Cross-linguistic influence in SLA* (pp. 93–120). Bristol: Multilingual Matters.

Ansorge, U., Kiefer, M., Khalid, S., Grassl, S. I., & Koenig, P. (2010). Testing the theory of embodied cognition with subliminal words. *Cognition*, *116*, 303–320. doi:10.1016/j.cognition.2010.05.010

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577–660. doi:10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Becker, A., & Carroll, M. (1997). *The acquisition of spatial relations in a second language*. Amsterdam: Benjamins.

Bergen, B., Lau, T. T. C., Narayan, S., Stojanovic, D., & Wheeler, K. (2010). Body part representations in verbal semantics. *Memory & Cognition*, *38*, 969–981. doi:10.3758/MC.38.7.969

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology*, *25*, 37–41. doi:10.1080/20445911.2013.812093.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. doi:10.1037/0033-295X.108.3.624.

Bowerman, M. (1996). Learning how to structure space for language. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space* (pp. 385–436). Cambridge, MA: MIT Press.

Bryant, D. (2012). *Lokalisierungsausdrucke im Erst-und Zweitspracherwerb: Typologische, ontogenetische und kognitionspsychologische Überlegungen zur Sprachförderung in DaZ* (Thema Sprache–Wissenschaft für den Unterricht, 2). Baltmannsweiler, Germany: Schneider Verlag Hohengehren.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition, 41*, 83–121. doi:10.1016/0010-0277(91)90033-Z

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

Coello, Y., & Fischer, M. H. (2016). *Foundations of embodied cognition: Vol. 1. Perceptual and emotional embodiment*. London: Routledge.

Collins, L. (2002). The roles of L1 influence and lexical aspect in the acquisition of temporal morphology. *Language Learning, 52*, 43–94. doi:10.1111/1467-9922.00177

Coventry, K., Guijarro-Fuentes, P., & Valdés, B. (2012). On the first and second language acquisition of spatial language. *Spatial Cognition & Computation*, *12*, 219–230. doi:10.1080/13875868.2012.713058

Coventry, K. R., & Garrod, S. C. (2004). *Saying, seeing and acting: The psychological semantics of spatial prepositions*. New York: Psychology Press.

Coventry, K. R., Prat-Sala, M., & Richards, L. (2001). The interplay between geometry and function in the comprehension of over, under, above, and below. *Journal of Memory and Language, 44*, 376–398. doi:10.1006/jmla.2000.2742

Cuyckens, H., & Zawada, B. (Eds.). (2001). *Polysemy in cognitive linguistics: Selected papers from the Fifth International Cognitive Linguistics Conference*. Amsterdam: John Benjamins.

Davies, M. (2008) *Word frequency data—Corpus of Contemporary American English*. Retrieved May 4, 2016, from http://www.wordfrequency.info/

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first- and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia*, *56*, 334–349. doi:10.1016/j.neuropsychologia.2014.02.003

Dijkstra, T., & Van Heuven, W. J. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, *5*, 175–197. doi:10.1017/S1366728902003012

Dudschig, C., de la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain & Language, 132*, 14–21. doi:10.1016/j.bandl.2014.02.002

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Dudschig, C., de la Vega, I., & Kaup, B. (2015). What's up? Emotion-specific activation of vertical space during language processing. *Acta Psychologica*, *156*, 143–155. doi:10.1016/j.actpsy.2014.09.015

Dudschig, C., & Kaup, B. (2016). Is it all task-specific? The role of binary responses, verbal mediation and saliency for eliciting language-space associations. *Journal of Experimental Psychology*. Advance online publication. doi:10.1037/xlm0000297

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language, 51*, 1–22. doi:10.1016/j.jml.2004.01.004

Fischer, M. H., & Coello, Y. (2016). *Foundations of embodied cognition: Vol. 2. Conceptual and interactive embodiment*. London: Routledge.

Fries, C. (1945). *Teaching and learning English as a foreign language*. Ann Arbor, MI: University of Michigan Press.

Glenberg, A. M., Brown, M., & Levin, J. R. (2007). Enhancing comprehension in small reading groups using a manipulation strategy. *Contemporary Educational Psychology, 32*, 389–399. doi:10.1016/j.cedpsych.2006.03.001

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*, 558–565. doi:10.3758/BF03196313

Grießhaber, W. (1999). *Die relationierende Prozedur. Zu Grammatik und Pragmatik lokaler Präpositionen und ihrer Verwendung durch türkische Deutschlerner*. Münster, Germany: Waxmann.

Grimm, H. (1975). On the child's acquisition of semantic structure underlying the wordfield of prepositions. *Language and Speech, 18*, 97–119. doi:10.1177/002383097501800201

Hernández-Chávez, E. (1972). *Early code separation in the second language speech of Spanish-speaking children*. Paper presented at the Stanford Child Language Research Forum, Stanford University.

Hyltenstam, K. (1977). Implicational patterns in interlanguage syntax variation. *Language Learning, 27*, 383–411.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in second language acquisition. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 539–588). Oxford: Blackwell.

Ijaz, I. H. (1986). Linguistic and cognitive determinants of lexical acquisition in a second language. *Language Learning*, *36*, 401–451. doi:10.1111/j.1467-1770.1986.tb01034.x

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.

Jirak, D., Menz, M. M., Buccino, G., Borghi, A. M., & Binkofski, F. (2010). Grasping language—A short story on embodiment. *Consciousness and Cognition*, *19*, 711–720. doi:10.1016/j.concog.2010.06.020

Kaltenbacher, E., & Klages, H. (2006). Sprachprofil und Sprachförderung bei Vorschulkindern mit Migrationshintergrund. In B. Ahrenholz (Ed.), *Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten* (pp, 80–97). Freiburg, Germany: Fillibach.

Kaup, B., de la Vega, I., Strozyk, J., & Dudschig, C. (2016). The role of sensorimotor processes in meaning composition. In M. H. Fischer & Y. Coello (Eds.), *Foundations of embodied cognition: Vol. 2. Conceptual and interactive embodiment* (pp. 46–70). London: Routledge.

Khalid, S., & Ansorge, U. (2013). The Simon effect of spatial words in eye movements: Comparison of vertical and horizontal effects and of eye and finger responses. *Vision Research*, *86*, 6–14. doi:10.1016/j.visres.2013.04.001

Klein, W. (1992). *Zweitspracherwerb* (Studienbuch Linguistik, 5th ed.). Frankfurt am Main, Germany: Anton Hain.

Kleinmann, H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning, 27*, 93–107.

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Krashen, S. (1983). Newmark's "ignorance hypothesis" and current second language acquisition theory. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 135–153). Rowley, MA: Newbury House.

Kroll, J. F., Bobb, S. C., & Hoshino, N. (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science, 23*, 159–163. doi:10.1177/0963721414528511

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*, 149–174. doi:10.1006/jmla.1994.1008

Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., & Kaup, B. (2011). Root versus roof: Automatic activation of location information during word processing. *Psychonomic Bulletin & Review, 18*, 1180–1188. doi:10.3758/s13423-011-0158-x

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.

Leather, J. (2003). Phonological acquisition in multilingualism. In M. García Mayo & M. García Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 23–58). Bristol: Multilingual Matters.

Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review, 1*, 476–490. doi:10.3758/bf03210951

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*, 251–285. doi:10.1017/S0272263100009165

Lütke, B. (2008). Beobachtungen zur Raumreferenz in mündlichen Erzählungen von Grundschulkindern mit Deutsch als Erst-und Zweitsprache. In B. Ahrenholz (Ed.), *Zweitspracherwerb. Diagnosen, Verlagno, Voraussetzungen* (pp. 151–170). Freiburg im Breisgau, Germany: Fillibach.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203. doi:10.1037/0033-2909.109.2.163

Munnich, E., & Landau, B. (2010). Developmental decline in the acquisition of spatial language. *Language Learning and Development*, *6*, 32–59. doi:10.1080/15475440903249979

Munnich, E., Landau, B., & Dosher, B. (2001). Spatial language and spatial representation: A cross-linguistic comparison. *Cognition, 81*, 171–208. doi:10.1016/S0010-0277(01)00127-5

National Institute of Korean Language. (2005). Research of frequency of modern Korean usage (Hyeondae Gugeo Sayong Bindo Josa 2). Retrieved May 16, 2016, from http://korean.go.kr/front/reportData/reportDataView.do?mn_id=45&report_seq=1&pageIndex=1)

Newmark, L. (1966). How not to interfere with language learning. Language learning: The individual and the process. *International Journal of American Linguistics, 40*, 77–83.

Quasthoff, U., Fiedler, S., & Hallsteinsdóttir, E. (2011). *Frequency dictionary German: Häufigkeitswörterbuch Deutsch*. Leipzig, Germany: Leipziger Uni-Verlag.

Richter, T., Zwaan, R. A., & Hoever, I. (2009). Acquiring experiential traces in word-referent learning. *Memory & Cognition, 37*, 1187–1196. doi:10.3758/MC.37.8.1187

Rösch, H. (2001). *Deutsch als Zweit- und Fremdsprache*. Berlin: Akademie Verlag.

Schwartz, B. D., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research, 12*, 40–72. doi:10.1177/026765839601200103

Sharoff, S., Umanskaya, E., & Wilson, J. (2013). *A frequency dictionary of Russian: Core vocabulary for learners*. London: Routledge.

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

Stringer, D. (2010). The gloss trap. In Z.-H. Han & T. Cadierno (Eds.), *Linguistic relativity in SLA: Thinking for speaking* (pp. 102–124). Clevedon: Multilingual Matters.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi:10.1037/h0054651

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225–282). Boston: Springer.

Tyler, A., & Evans, V. (2003). *The semantics of English prepositions: Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.

Verhelst, N., Van Avermaet, P., Takala, S., Figueras, N., & North, B. (2009). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Vukovic, N. (2013). When words get physical: Evidence for proficiency-modulated somatotopic motor interference during second language comprehension. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 3711–3716). Austin, TX: Cognitive Science Society.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. In D. Pecher & R. A. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 224–245). Cambridge: Cambridge University Press.

https://doi.org/10.1017/S014271641700042X Published online by Cambridge University Press