



Conceptual Blending Across Ontological Domains—References to Time and Space in Motion Events by Tunisian Arabic Speakers of L2 German

Johannes Gerwien*[†] and Christiane von Stutterheim[†]

Heidelberg University, Institute for German as a Foreign Language Philology, Heidelberg, Germany

OPEN ACCESS

Edited by:

Alberto Hijazo-Gascón,
University of East Anglia,
United Kingdom

Reviewed by:

Sabine De Knop,
Université Saint-Louis
Brussels, Belgium
Benjamin Fagard,
CNRS, France

*Correspondence:

Johannes Gerwien
gerwien@idf.uni-heidelberg.de

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 17 January 2022

Accepted: 08 March 2022

Published: 30 March 2022

Citation:

Gerwien J and von Stutterheim C
(2022) Conceptual Blending Across
Ontological Domains—References to
Time and Space in Motion Events by
Tunisian Arabic Speakers of L2
German. *Front. Commun.* 7:856805.
doi: 10.3389/fcomm.2022.856805

Patterns of information selection and verbal encoding may rely on an interdependence between the spatial and temporal conceptual domain in the context of motion events. This has been shown, e.g., for Tunisian Arabic (TA), a language with a highly differentiated aspectual system. We address the question whether this interdependence can also be observed when L1 speakers of TA describe motion events in their L2 German, a language without grammaticalized aspect. Data obtained in an unscripted language production experiment in which L1 and highly advanced L2 speakers of German describe videos showing different types of motion events (one type showing boundary crossing at a goal like *a woman entering a supermarket*, the other type showing motion along a path with no evident goal such as *a car driving along a road*) suggest that this is indeed the case. The L2 speakers deviate systematically from the L1 speakers of German in the information they select for verbal encoding, but show clearly similar patterns to those used when describing the same scenes in the L1 (TA). The differences can be interpreted as pointing to the high relevance attributed to the temporal dimension of the events shown in the videos by the L2 speakers. The results are placed in the theoretical framework of schema theory. The findings for Arabic speakers of L2 German can be explained by assuming that the same event schema is activated in the context of L2 use as in the context of L1 use.

Keywords: motion events, event schema, crosslinguistic, Tunisian Arabic, German, second language, spatial cognition, language production

INTRODUCTION

With the increasing interest in the relationship between cognitive and linguistic processes, the field of motion events has again come into focus in empirical research. In a large number of studies, motion events, and their expression are used as a window on the possible role of language in pre-verbal information selection and conceptual processing. The findings present a highly diverging picture, however, which applies in particular to the theoretical positions taken in interpreting the results. One line of research follows a paradigm which can be labeled “the cognitive universal approach”. In this framework, spatial cognition is viewed as species-specific and as basically universal across speakers of different languages. Its roots lie in the bodily nature of mankind, in basic categories of our spatial orientation, navigation, reasoning, and language which are common to all human

beings (Miller and Johnson-Laird, 1976; Landau and Jackendoff, 1993; Gleitman and Papafragou, 2005). Differences such as those observed across languages when encoding motion events are rated as superficial from this point of view. There is empirical evidence, however, which does not support this position. The important standpoint, in that respect, is that *language differences imply cognitive differences* (Haun et al., 2011). Numerous studies have investigated the mutual dependency between linguistic categories and cognitive functioning over the last 30 years (Levinson, 1996, 2003; Majid et al., 2004; Gerwien and von Stutterheim, 2018). They span different domains such as language production and comprehension, second as well as first language acquisition (Bylund and Jarvis, 2011; Athanasopoulos et al., 2015; Hijazo-Gascón, 2021; Stutterheim et al., 2021) and sign language (Slobin and Hoiting, 1994; Arik, 2010). Comprehensive overviews have shown that we are far from understanding the intricate interplay between the different roles of cognitive representations, and processes at this level, and those involving linguistic representations (Evans and Chilton, 2007; Vulchanova and van der Zee, 2013). In order to make progress in the current debate, innovative approaches which go beyond the investigation of the well-established spatial concepts selected across typologically different languages are now necessary. In narrowing the focus from spatial cognition in general to the field of motion events in particular, this involves the inclusion of other conceptual domains which go beyond investigations of the use of manner and path verbs. The aim of the present study is to place the focus on those aspects which come to the fore when looking at the interplay between spatial and temporal concepts. Advanced L2 speakers are particularly relevant with regard to the selection of pre-verbal information given the fact that they have two systems of expression at their disposal. The specific linguistic and cognitive structures typically represented in speakers of different languages are represented jointly in the mind of a multilingual speaker. As previous studies have shown, the two levels need not be linked in the process of L2 acquisition (see overview in Bassetti and Cook, 2011). Information selection and conceptualization can follow the principles of either language or principles of a third type, which results from a merge. L2 data can shed light on processes at the level of preverbal conceptualization which are not driven by expressive devices in the sense of *thinking for speaking* (Slobin, 1996). The present study is based on the language combination Tunisian Arabic (TA) L1 with German L2 which is highly relevant for two reasons: In addition to the fact that there are very few studies on the expression of motion events in Arabic languages, this group provides a relevant window on the research domain of motion event representation and encoding in that they allow insight into the relationship between spatial and temporal concepts, given the presence of grammaticalized verbal aspect. Tunisian Arabic has a three-fold aspectual system in which all aspectual categories play specific roles. As shown in a recent study, the aspectual system is highly relevant in motion event encoding for speakers of Arabic varieties (De Knop, 2020). Studies on speakers of an Arabic language, when learning a language in which aspectual categories are not grammaticalized, as in the case of German, present a test case in the identification of potential L1 processes in

information selection and encoding patterns, compared to those of the target language.

The interpretation of the results obtained in the present language production experiment relies on a novel theoretical approach. As will be shown below, the interdependence between the spatial and temporal ¹domain is modeled on the basis of event schemata which initially guide information selection and subsequently verbal encoding. The following section starts with a review of selected examples from the relevant literature with a focus on L2 studies. This is followed by an overview of the typological features of the Tunisian language in the domain of spatial cognition. The empirical section presents the data on Tunisian speakers with German as L2 in comparison to L1 data from German speakers. The discussion section outlines the relevance and implications of the present empirical findings with regard to the debate on language specificity in the conceptualization of motion events relating the findings to the theoretical conception of schema theory (Bartlett, 1932; Brewer and Nakamura, 1984; Hintzman, 1986; Zacks, 2004; Gerwien and von Stutterheim, 2018).

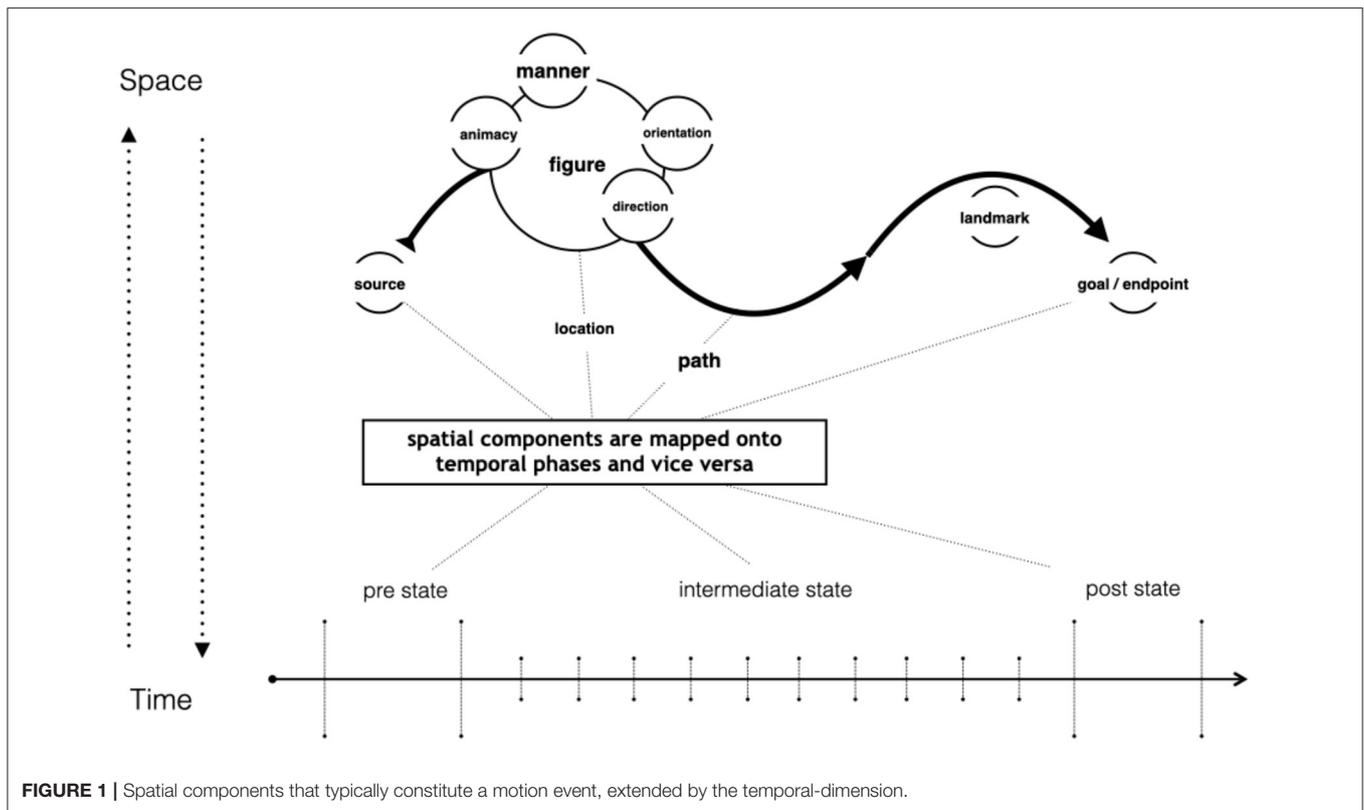
Research Background

Over the last 20 years, the field of motion events has become a test case in empirical research on questions concerning language specificity in processes of conceptualization. Two reasons, both methodological as well as theoretical, have led to this development. Methodological questions relate to specific features of the spatial domain. Since motion in space, and the associated configurations, are visually perceived, this allows controlled experimental situations in which cognitive processing can be clearly linked to specific visual stimuli. This is paired with a well-defined typological framework which provides frames of reference that allow reliable comparisons of studies across different languages. **Figure 1** summarizes the individual conceptual components that typically constitute a motion event, according to the present range of cross-linguistic studies.

The framework for cross-linguistic studies of motion events is based on the seminal studies on spatial cognition carried out by Talmy (see overview in Talmy, 2000). The core components of a motion event are specified as *motion* and *path*. Languages can be categorized as *verb-framed* or *satellite-framed* depending on whether they predominantly express *motion* and *path* together in the verb root, or *path* outside the verb in a “satellite”. If information on the *path* is expressed in a satellite, the verb root typically expresses manner of motion. The distinction between *verb-framed* and *satellite-framed* languages provided the basis for the formulation of hypotheses on a range of factors covering information selection, morpho-syntactic patterns, relevant effects in L1 and L2 language acquisition—to name but a few in a long list. Although the theory has been refined at different levels (Beavers et al., 2010; Croft et al., 2010),

¹We take the notion *temporal* as umbrella term for different subcategories which refer to the global domain of time. Aspect is the subcategory which relates to different phases of an event. The selection of an aspectual category is dependent on a view point. Another subcategory is tense which relates to temporal anchoring on the time line.

OR: Aspect and tense are subcategories under this term.



the critical role of the conceptual categories *path* versus *manner* remains at the core of all approaches in this field.

In the initial phases, studies focused on comparisons of the main contrasts across the lexical and structural characteristics of typologically different languages (see overview in Slobin, 2004). The basic discussion then moved on to the question as to where the diverging patterns observed across different languages are “located”. Do they constitute a language-related phenomenon, in the sense that differences in how motion events are verbally expressed are motivated by the respective linguistic system, or are they represented at a non- or pre-verbal level?

There is a further aspect which should be considered in this debate. Empirical results to date point to the fact that language is recruited automatically when solving cognitive tasks (Alderson-Day and Fernyhough, 2015). If it is the case that speakers of different languages are guided by different principles when they perceive and encode motion events in the context of a verbal or a non-verbal task, then this is the level of cognitive processing that counts in information selection and communication. Some of the relevant consequences can be formulated as follows: what does a witness “take into account” when surveying a scene? How do we categorize and cluster objects? What are the principles underlying inferences, implicature, and presuppositions? When you consider this, the discussion whether there are universals in the sense of spatial primitives becomes rather marginal. All these considerations become particularly relevant in multilingual contexts. As mentioned above, balanced bilinguals and speakers of a second language (L2) have more than one language

option when solving non-verbal and verbal tasks: the speakers could, in theory, integrate L1-based cognitive strategies when conceptualizing content with the L2-based linguistic material accessed in the formulation process. This has been shown in a number of studies on motion events to date (Cadierno and Ruiz, 2006; Carroll et al., 2012; Athanasopoulos et al., 2015; Hijazo-Gascón, 2021; Stutterheim et al., 2021; Lambert et al., 2022) which were based to a large extent on Talmy’s distinction between verb-framed and satellite-framed languages. Studies diverge with respect to the language pairs studied (L1 verb-framed—L2 satellite-framed or vice versa) and the elicitation tasks used. The findings were not clear cut. Depending on the level of L2 competence, as well as on the saliency of markers in the L2 and the typological distance, among other factors, L1 patterns were shown to influence how information is selected, structured and expressed in the L2 linguistic forms.

The current approach differs from preceding research in that focus is placed on specific conceptual clusters which go beyond the spatial domain. In some languages, the information selected when encoding a visually presented scene showing a motion event will also include temporal information such as an aspectual component. This stands in a relationship of mutual dependency with the spatial components of the given situation. The question under focus in the present analysis is as follows: Would speakers of a second language show an interdependency between spatial and temporal categories where this constitutes a feature of their L1 (e.g., Tunisian Arabic), but not of the L2 given as target

language (e.g., German)? In studying “cross-domain effects” it should be stressed that the effects of the L1 on information selection and conceptualization can be accounted for on different grounds. Studies to date have mainly focused on the role of spatial concepts in the context of manner- versus path-based patterns and the language-specific constraints involved. It has been rightly argued that one cannot draw conclusions with respect to language specificity at the level of conceptualization on these grounds since both conceptual categories (manner and path) are core components of a motion event across all languages. However, things are different in the case of aspectual differences. The integration of aspectual categories is objectively not required for the conceptualization and encoding of a motion event. In the language pairing given in the present study, the source language TA involves obligatory aspectual marking at the linguistic surface, the target language German does not. If L2 speakers of German show sensitivity to aspectual features when encoding information on motion events then this cannot be due to differences in weighing up specific components, such as *manner* and *path*, that constitute the core domains for the construal of motion events. The attribution of a temporal perspective manifested in the selection of a particular phase of a motion event is motivated by experience with a specific linguistic system (L1 Tunisian in our case). We will come back to this point in the discussion.

Motion Events in Tunisian Arabic Tunisian Arabic

The linguistic situation in Tunisia can be described as diglossic. With Tunisian Arabic as the mother tongue, a further language, Modern Standard Arabic (MSA) is acquired in school and used in formal contexts. In higher education students also learn French. The oral Arabic variety Tunisian Arabic differs from MSA with regard to its lexical, phonological as well as morpho-syntactic features. In the present context the overview on TA will be restricted to those features which are relevant for the verbalization of motion events.

The Encoding of Spatial Concepts

The present analysis builds on earlier studies by the Heidelberg research group (Carroll et al., 2012; von Stutterheim et al., 2012, 2017; Flecken et al., 2014; von Stutterheim and Gerwien, in press) as well as Saidi (2006) and Louhichi (2018). The two latter studies focus on the spatial-typological classification of TA. Semitic languages in general were classified as verb-framed by Talmy (1985). Saidi and Louhichi examined this classification in detail, taking into account the critical features put forward in Talmy’s typological framework.

- a. Verbal lexicon: TA has a limited number of path verbs and no verbs for the expression of goal-oriented directionality (such as *to advance*, *to head for*). The number of manner verbs is not as high as in typical satellite-framed languages.
- b. Boundary crossing constraint: TA shows the boundary crossing constraint. This means that manner verbs cannot be

- combined with boundary crossing adjuncts (such as *the girl runs into the garden shed*, as typical in English for example)².
- c. Multiple-ground-information constraint: TA complies with this constraint. Manner verbs cannot combine with multiple ground adjuncts when providing information on the path (unlike English: *they run out of the house across the lawn into the woods*).

Louhichi (2018) points to a structure which he claims has remained unnoticed in previous studies. Speakers of TA break down the motion event into two subevents, with one referring to the manner of motion and the other to the path: *the boy runs and enters the house*. Given this pattern in TA, Louhichi argues that, this constitutes a specific subtype of the verb-framed category (2018:358). This pattern, however, can also be found in other “typical” verb-framed languages such as French (Gerwien and von Stutterheim, 2018; von Stutterheim and Gerwien, in press). Based on these findings, the categorization of TA as verb-framed seems to be warranted. However, the two studies by Saidi (2006) and Louhichi (2018) do not extend the question to the other relevant domain, the presence of temporal aspect.

The Interrelationship Between Temporal and Spatial Conceptual Categories

The study by von Stutterheim et al. (2017) shows an interdependency between spatial and temporal components in the verbalization of motion events which had not been described before. TA is a language in which verbal aspect is encoded morphologically in the verb. As in MSA, there is the opposition between an imperfective (prefixed) and a perfective (suffixed) as well as a progressive form (semi-finite participle active). Unlike TA, this does not have a grammaticalized status in MSA. The use of the perfective requires a point of change, *mša* PF.3SGM “to move”, “go away” (starting point, change from being at location *x* to leaving *x*) or *dxal* PF.3SGM “to enter” (goal point, change from being outside to inside of an object *y*). A path verb marked by the perfective expresses “goal reached” and the state following the boundary crossing is thereby asserted:

- (1) *qaʔos dxal li l-bīt*
cat enter-PF.3SGM to DEF-room
'a cat entered the room (and is now in the room)'

By contrast, verbs in the progressive form refer to a motion event where the figure is on a path that leads incrementally to a goal, or away from a point at a source. In this case, reference is made to a dynamic situation which does not include the state at the goal point after boundary crossing.

- (2) *rāgel dāxel li s-supermarché*
man enter-PART.3SGM to DEF-supermarket
'a man is entering the supermarket (but is not yet in the supermarket)'

²We are aware of the fact that there are exceptions to this constraint. First of all there are differences across languages which are considered to be of the type verb-framed. In Turkish, for example, manner verbs combine with boundary crossing constraints. French, which is more restrictive in this respect, nevertheless has exceptions to this principle (Aurnague, 2011; Sarda, 2019).

Use of the imperfective form cannot occur with references to goal-oriented motion since the perspective established by the imperfective relates to a figure in motion without an envisaged goal. It can combine with forms which refer to a location or those with no spatial adjunct.

- (3) *rāgel yhawwes fi š-šāri*
man go-out-IPF.3SGM in DEF-street
'a man is going out for a walk in the street'
- (4) *mra tegri*
woman run-IPF.3SG.F
'a woman is running'

The implications that arise between spatial and temporal concepts in the representation of motion events in TA lead to specific constraints that are not found in languages that do not have a comparable aspectual system. The perfective combines with a two-state predicate and asserts the second state (post state) which refers to the position of the figure following the boundary crossing. The progressive combines with the same types of predicates but refers to the initial state before the boundary crossing. The opposition between the perfective and the progressive, when used in the context of motion events, has implications for information selected for expression in the spatial domain. The imperfective is confined to one-state predicates. It therefore blocks access to a goal point as well as to forms of spatial reference which imply directionality. The relevant factors when encoding information on motion events in TA in the three templates can be summarized as follows (see also von Stutterheim et al., 2017, 2020).

- Change-of-state / path verbs + PROG/PF → events expressing directed motion
- Change-of-state/ path verbs + PF → events expressing the state following a boundary crossing
- No change-of-state/ manner verbs+ IPF → situations expressing activities/ states relating to the location of figure or those with no spatial information

German

Since motion events in German have been studied in depth (see overview De Knop, 2020), its status as a typical representative of the type *satellite-framed language* is well-established. German possesses a rich repertoire of manner verbs, with very few path verbs that are rarely used. Motion events are thus expressed by manner verbs and particles³ or prepositional adjuncts when providing information on the path of motion. Speakers of German typically combine manner verbs with directional adjuncts when referring to dynamic motion events as in *eine Frau läuft in den Supermarkt* "a woman walks into the supermarket" (cf. Carroll and von Stutterheim, 2011; Carroll, 2000). Note,

³Verb particles of the type *hinein/herein* (into), *raus* (out), *drüber* (across), which as particles are only roughly translated as preposition in English, form part of the verb morphologically and constitute a specific property of the German system. Particles can be separated from the verb in finite form but are merged with the verb in non-finite form: *Peter kommt in das Zimmer rein.. Peter ist in das Zimmer reingekommen* (Peter comes into the room to-in; Peter is into the room to-in-come).

however, that in most cases, specific information on the boundary crossing remains underspecified: In contrast to a language such as TA, the description in German leaves the question open as to whether the woman is still underway to the supermarket or whether the supermarket has actually been entered. In cases where directed motion would not include reference to a source or a goal point, speakers of German provide information on the path taken by the figure by reference to the ground traversed (*eine Frau läuft eine Straße entlang* "a woman walks a road along"). Since manner verbs are a core feature in the expression of motion events in German which means that they have to be complemented by information on the path, a sentence such as *Ein Kind läuft auf der Straße* (a child walks on the street) does express a different type of situation.. It focusses on the manner "walk-on-the-street" as a kind of property of the figure in contrast to somewhere else (e.g., *walk-on-the side path*). The data show that German speakers rarely use locational adjuncts when referring to events involving directed motion (Flecken et al., 2015). This means that the relevant categories in the encoding of motion events in German differ markedly from those described for TA. Furthermore, aspectual categories do not play a role in German. The specific criteria thus merge into one template in German.

- manner verb + spatial information (source, goal, spatial features of the path)

As will be shown below, the analysis of TA and German shows that speakers select and encode different categories of spatial and temporal information when referring to the same set of situations.

MATERIALS AND METHODS

The present investigation is a follow-up on the comparative studies mentioned above (Carroll et al., 2012; Flecken et al., 2015; von Stutterheim et al., 2017; von Stutterheim and Gerwien, in press). The design of the present study, and the material and procedure are identical, in order to ensure comparability at this level. Focus is placed on L2 speakers of German with Tunisian as their L1. A new data set was collected for the L1 speakers of German.

Stimuli

The stimulus set was the same as that used in previous studies conducted by the research group. It covers a total of 70 real-world video clips showing different types of situations. All videos are six seconds in length, with motion events that cover three different categories with 10 videos for each case: Type A stimuli show a figure moving along a path with an evident goal point, which, however, was not reached before the video ended (e.g., a figure walking toward the entrance to a building). Type B stimuli show a figure moving along an extended trajectory with a potential (but not clearly evident) goal point (e.g., a village or a house in the distance). In Type C stimuli the motion events show goal points that were reached and include a boundary crossing (a car entering a garage). The focus in the present study is placed on stimuli of the type B and C (see **Figure 2**). We expect that language-specific event schemata, which underlie information



FIGURE 2 | Screenshots illustrating video stimuli from the two critical conditions; **left**: Type B (extended trajectory); **right**: Type C (boundary crossing).

selection and linguistic encoding processes in the speaker's L1, would be likely to prevail in the spontaneous responses under these two conditions.

The remaining 40 videos show an agent performing an action on an object (e.g., a person knitting a scarf, building a model airplane), as well as activities (e.g., playing the flute or working in the garden) and states (e.g., a candle burning, ice cream melting on a plate). These videos were introduced as fillers in order to reduce priming effects, i.e., participants recycling sentence formats and lexical items.

Participants

The participants were Tunisian students ($N = 19$) of different technical subjects at the Heidelberg College of Technology (Fachhochschule) and were aged between 19 and 27. Their level of competence in German as a foreign language was C1 (based on the European Reference Frame). They have been living in Germany between 1 and 5 years. All participants studied German already in their home country. They were asked to fill out a questionnaire on their social and linguistic background. Three subjects were female, 16 male, and all with advanced knowledge of French and English. The L1 participants ($N = 19$) in the German group were 5 male and 14 female students at the University of Göttingen (aged between 20 and 32). All participants had a comparable socio-economic background with an advanced knowledge of English.

Procedure

The participants were tested individually. They were seated in front of a computer screen on which the instructions were displayed. This was followed by a training session with 3 items. They were then asked if the procedure was clear. The instructions in German were as follows:

Sie werden eine Reihe von kurzen Videos sehen, die alltägliche Situationen zeigen und nicht in Verbindung miteinander stehen. Ihre Aufgabe ist es zu sagen, was passiert. Sie können beginnen, sobald Sie erkennen, was in dem Video vor sich geht. Berücksichtigen Sie dabei keine Einzelheiten der Szene (z.B. der Himmel ist blau). Konzentrieren Sie sich bitte auf das, was passiert.

(You will see a set of short video clips showing everyday events which are not in any way connected to each other. Your task is

to tell “what is happening” and you may begin as soon as you recognize what is happening in the clip. It is not necessary to describe the video clips in detail (e.g., “the sky is blue”). Please focus on the event only). All verbal responses were audio-taped. Following the experiment, which lasted approximately 8 min, the participants were asked to complete the questionnaire on their educational and linguistic background.

Coding the Verbal Responses

The coding procedure, based on previous extensive studies in this domain, was carried out by two independent coders in order to assure consistency in the fine-grained semantic analysis of the utterances produced. Differences across the coders were limited and the few that occurred were easy to settle. The coding process was binary: a value of 1 was assigned if an utterance contained an element from one of the coding categories (e.g., if a manner verb occurred, 1 was filled in the column “manner”) and 0 was assigned if an utterance did not contain any element from the respective category (e.g., if a path verb occurred, the column “manner” was marked as 0, and the column “path” as 1).

Valid responses A response was coded as valid if it (1) referred to the video presented and (2) expressed a motion event. Responses that simply mentioned the figure in motion or other objects in the scene [e.g., “there is a woman (in the city)”, “I see a bus stop“], or an activity instead of a motion event (“Someone is going for a walk“), were classified as invalid.

Verb type All valid responses were coded for verb type - either manner or path. Neutral motion verbs (“to move”) were also coded (only three cases).

Satellites In German, as a satellite-framed language, the path component is expressed in the following linguistic forms: prepositional phrases predominate (e.g., *geht zu einer Kirche*, “walks to a church”), but there are also some cases in which spatial information is “verb-framed” as in (*betritt ein Gebäude*, “enters a building”) with the spatial argument encoded as a case marked nominal phrase. Prepositional phrases were coded under the label “adjunct”. Instead of in the main verb, information on the direction taken is frequently encoded in German in verb particles (e.g., *hineingehen* which roughly translates as *to-in(go)*). In many cases these forms do not occur adjacent to the verb, given the finite verb second (V2) pattern in word order, but at the end

TABLE 1 | Coding scheme for satellites.

Satellite semantic category	Example
No adjunct	Eine Frau läuft "A woman walks"
Location	Eine Frau läuft auf einer Straße "A woman walks on a street"
Path/object	Eine Frau läuft entlang einer Straße "A woman walks along a street"
Direction	Eine Frau geht in Richtung eines Gebäudes "A woman walks in the direction of a building"
Goal	Eine Frau geht in ein Gebäude "A woman goes in a building"
Goal reached	Eine Frau betritt ein Gebäude "A woman enters a building"
Path/object (verb particle)	Eine Frau geht eine Straße entlang "A woman walks a street along"
Direction (verb particle)	Eine Frau geht auf ein Gebäude zu "A woman walks a building toward"
Goal reached particle	Eine Frau geht in ein Gebäude hinein A woman walks in a building (seen from outside "hin" to inside "ein")

of the utterance (*sie geht in ein Haus hinein*, "she walks in a house to-in"). Note that particles in German remain attached to the verb in verb-final sentences (*Ich sehe, dass er in das Haus hineingeht*, "I see that he in the house intowalks"). In the present coding procedure a distinction was therefore drawn between adjuncts (case-marked NPs and prepositional phrases) and verb particles. Both adjuncts and particles were coded based on the spatial information they express (see **Table 1**). In order to capture the relevant language-specific patterns, the coding scheme also listed whether an adjunct occurred or not (see <https://osf.io/zs4ny/> for examples of L2 speakers' responses).

Motion event descriptions in German may include more than one path segment as in *Eine Frau geht über einen Platz in ein Gebäude*, "A woman walks across a plaza into a building". Since adjuncts and particles were coded per utterance, the final sum for all adjunct types may exceed 100%, as in **Figure 5**.

RESULTS

Verb Type

The overview starts with the use of verb types (manner vs. path) in both conditions B and C. As shown in **Figure 3**, L1 and L2 speakers rarely show any differences with respect to the selection of a manner or path verb across the different conditions. Since there are very few neutral verbs (such as *sich bewegen* "to move"), they were not further analyzed. Given the obvious similarity in verb type selection they will not be compared statistically.

Verb Vocabulary Size

The analysis covers the proportion of unique (non-identical) verbs used in both conditions by dividing the number of unique verbs used by the number of available responses per participant and condition. A value of 0.2, for example, would indicate that a subject used only two different verbs across 10 items (see **Figure 4**). The analysis shows that there is almost no

difference regarding verb vocabulary size in the context of motion events, thus indicating that Tunisian L2 speakers of German have successfully acquired the lexical means that L1 speakers of German typically use in this context.

Satellites

Figure 5 depicts the frequency of occurrence of different satellite types in proportions per condition (Type B and Type C scenes). It shows that the two different stimulus types elicited different encoding patterns for both groups of speakers, as in previous studies (von Stutterheim et al., 2012; Flecken et al., 2015). Although various linguistic forms occur in Type B stimuli (endpoint not evident), Type C stimuli (endpoint reached) led almost exclusively to the use of forms referring to the endpoint, as expected. In both conditions, differences were observed between L1 and L2 speakers with respect to the path components encoded. **Figure 5** shows four clear differences for the Type B scenes: (1) L2 speakers produce more utterances with bare verb phrases relative to the L1 speakers. (2) L2 speakers produce fewer adjuncts expressing path information relative to a landmark (*entlang der Straße laufen* "walk along the road"). (3) L2 speakers produce fewer verb particles that express path information relative to a landmark (*die Strasse entlang laufen* "the street along-walk"). (4) L2 speakers produce fewer verb particles that express direction (*auf × zulaufen* "toward × walk"). **Figure 5** also shows a difference in Type C scenes: L2 speakers frequently use verb particles that express the state "goal reached" (*läuft in das Haus rein* ~ "walks in the house into").

Statistical Analysis of Satellite Type Occurrence

Separate Bayesian generalized mixed models, assuming a Bernoulli distribution with a logit-link function, were used to assess potential group differences for all satellite types in which visual inspection of the **Figure 5** suggested a clear effect and for which a reasonable number of occurrences was observed (occurrence in more than 10%). Specifically, we analyzed the satellite types "No adjunct", "Path/Object particle", and "Direction particle" in Type B scenes, and "No adjunct", and "Goal reached particle" in Type C scenes. In every model, occurrence (yes = 1/no = 0) was specified as the dependent variable. "Group" was included as the only predictor (contrast-coded: L1German = +1, L2German = -1). Random intercepts were specified for subjects and items, thereby accommodating random variance induced by inter-individual differences as well as variance introduced by the different video clips. The brms R package (Bürkner, 2017, 2018) was used to estimate parameters for the Bayesian models that were fitted with three Markov Chain Monte Carlo (MCMC) chains. Each chain contained 1,000 burn-in samples and 5,000 additional samples with the thinning parameter set to 1. This resulted in 4,000 posterior samples per chain. These were combined to one posterior sample consisting of 12,000 samples. In order to induce a level of conservatism in the models, given the relatively low number of observations, a normally distributed prior centered at zero was set with a standard deviation at 0.5 on the slope term (the group difference). This can be viewed as a weakly-informative prior. Intercepts had

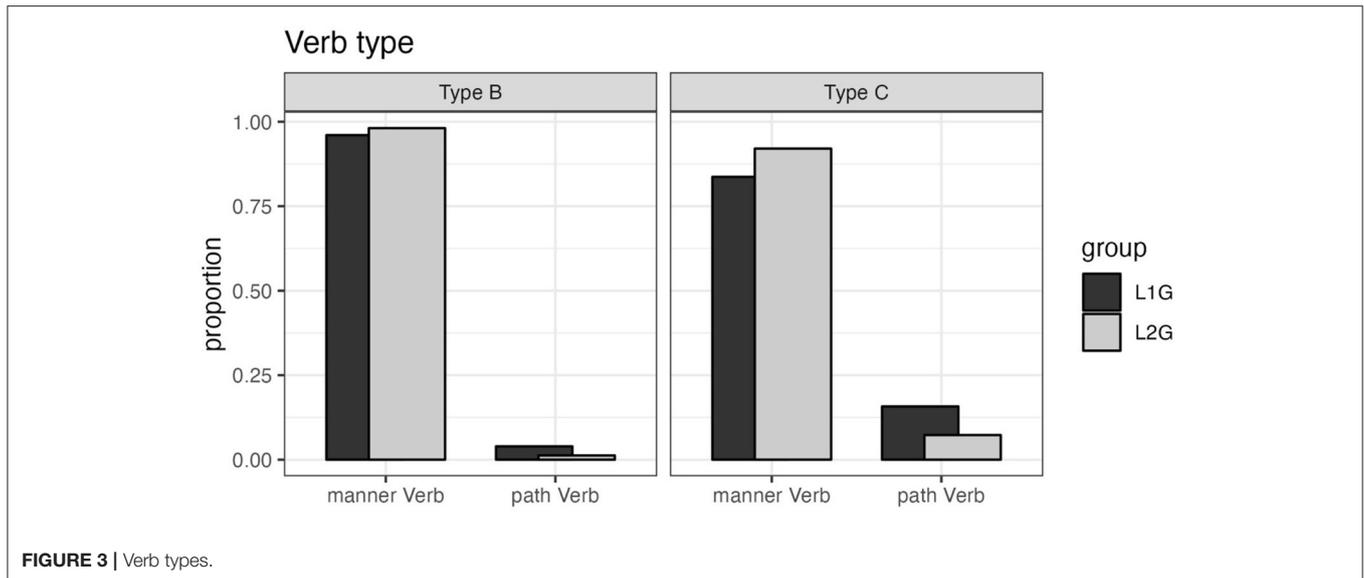


FIGURE 3 | Verb types.

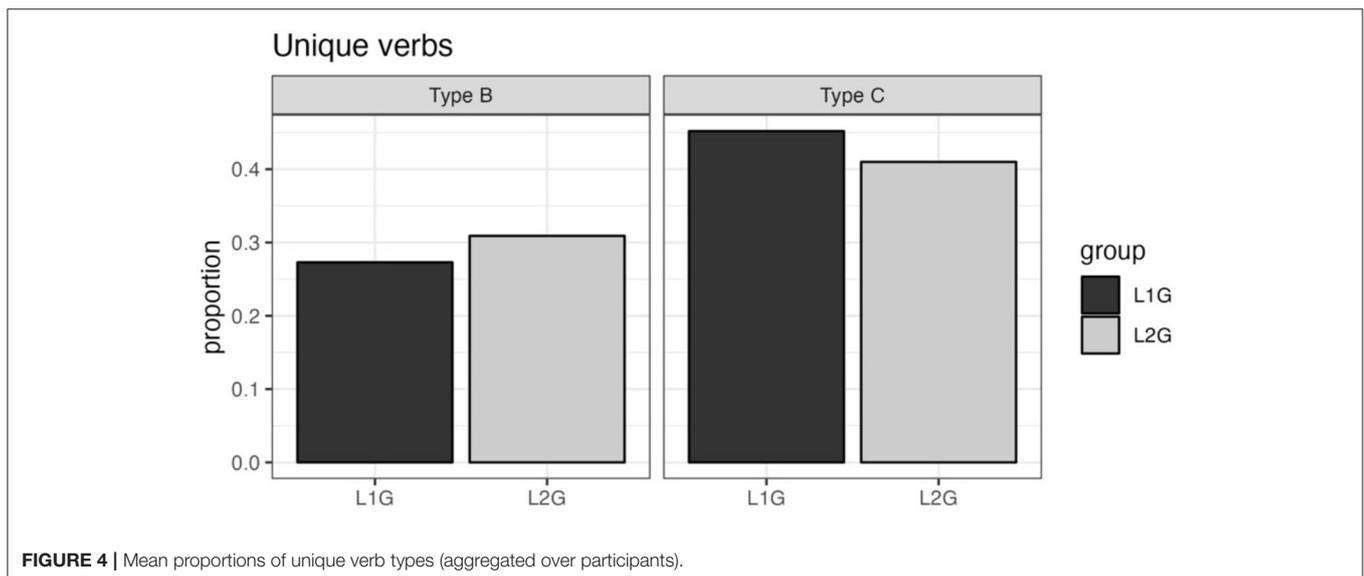


FIGURE 4 | Mean proportions of unique verb types (aggregated over participants).

Student's t distributed priors, $b0 \sim t_{(5,0,10)}$, standard deviations of random effects had Cauchy distributed priors, $\sigma \sim \text{Cauchy}(0, 2)$. Model convergence was evaluated based on the Gelman–Rubin convergence statistic \hat{R} (Gelman and Rubin, 1992), with \hat{R} -values close to 1 indicating negligible differences between within- and between- chain variances. Table 2 reports the mean group differences on the logit scale and the corresponding credible intervals (CI) to describe the posterior distributions of sampled regression weights (see <https://osf.io/zs4ny/> for the data and the R script).

In sum, and given the data and the respective models, the predictor language group showed an effect on the occurrence of no adjunct (L2German > L1German) and path/object particle (L2German < L1German) in Type B stimuli, as well as on the occurrence of no adjunct (L2German > L1German) and a “goal reached” particle (L2German > L1German) in Type C stimuli.

DISCUSSION

The present study investigated potential L1 effects in information selection and preverbal planning in the speaker's use of the L2. Verbalizations by German L1 speakers were compared with L2 speakers of German with Tunisian Arabic as their L1. The selection of these two groups was motivated by previous studies on motion events showing how speakers of Tunisian Arabic direct attention to the temporal/aspectual characteristics of a situation when speaking in their L1 (von Stutterheim et al., 2017). This presents a clear contrast with L1 German speakers who focus on spatial features of the places traversed (von Stutterheim et al., 2012). We argue that this clustering of temporal and spatial concepts can only be adequately captured by adopting a more complex and abstract format of representation: an *event schema*. Before

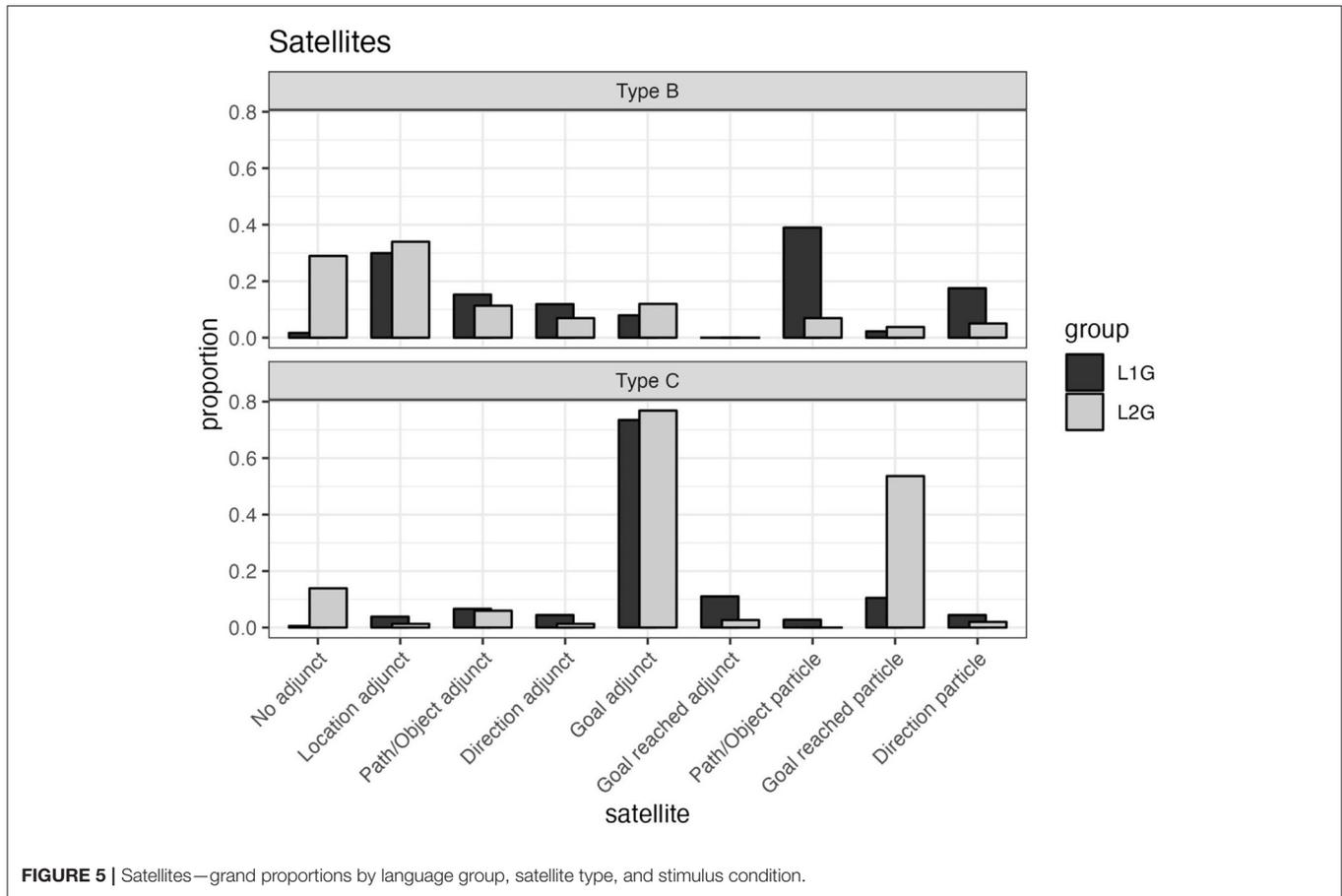


FIGURE 5 | Satellites—grand proportions by language group, satellite type, and stimulus condition.

TABLE 2 | Modeling results.

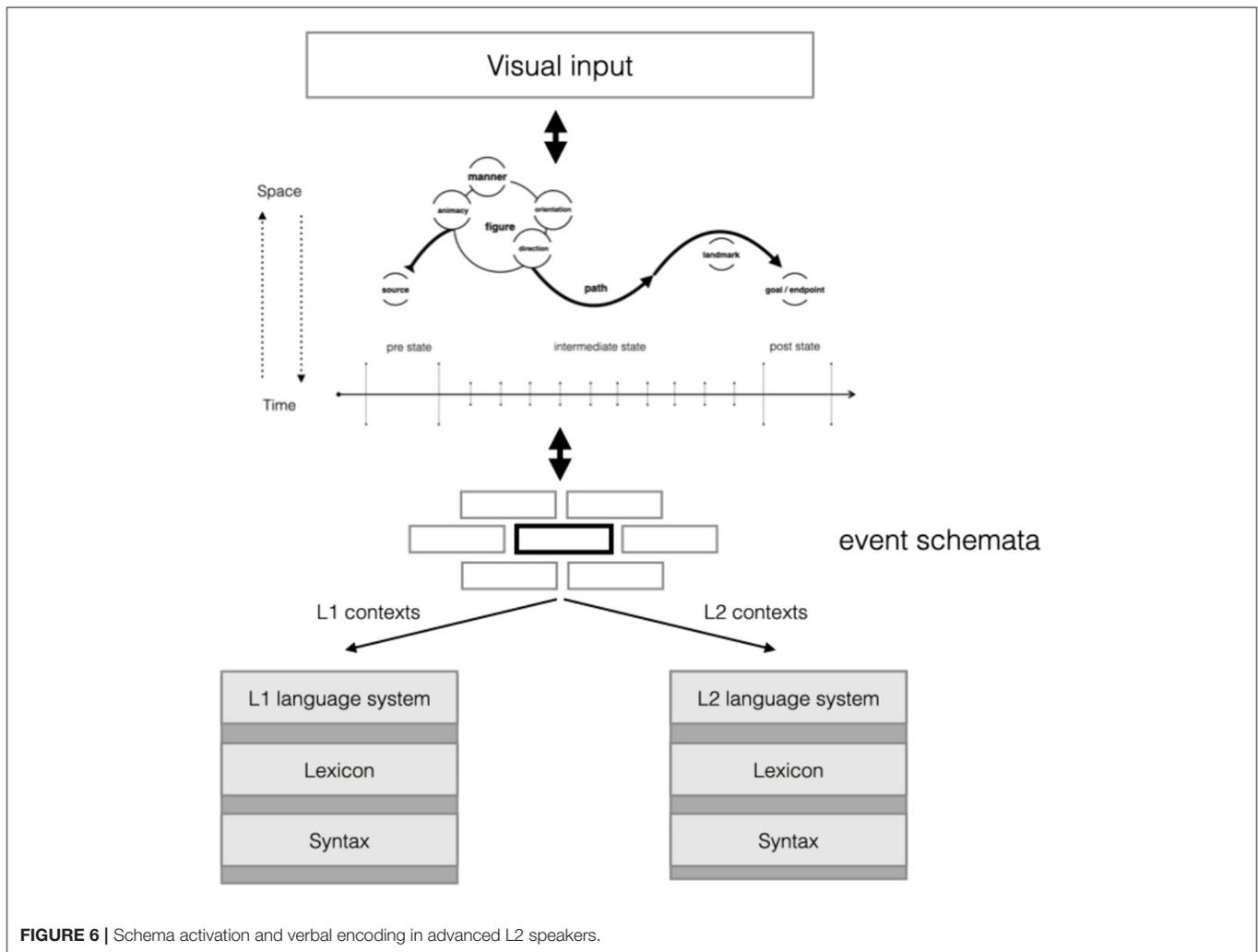
Scene type	Satellite	Group difference estimate (se)	95% Credible interval (CI)
B	No adjunct	-2.56 (0.46)	[-3.50, -1.71]*
B	Path/object particle	2.04 (0.56)	[0.91, 3.14]*
B	Direction particle	1.18 (0.61)	[-0.02, 2.35]*
C	No adjunct	-2.21 (0.56)	[-3.37, -1.15]*
C	Goal reached particle	-2.31 (0.45)	[-3.21, -1.44]*

Estimates are on the logit scale; the range of the 95% CI indicates whether the mean difference between groups included zero or not (a star indicates that it did not).

proceeding however, it is necessary to summarize and discuss the relevant findings.

Starting with situations in which no evident end point or landmark was shown in the video (Type B), there was no difference between the groups in the selection of verb types. The verbs used by the Tunisian L2 speakers of German showed a similar range of types as in the L1 group. In one sense this is not surprising since German has a very low number of path verbs. Significantly, however, the implications based on the use of manner verbs, clearly observed in the data of the German L1 speakers, were not observed in the L2 data. The implications in question concern the spatial information encoded *via* satellites, i.e., adjuncts and particles: 60% of the L2 speakers’ verbalizations did not provide any information on the path of motion. 30% of these verbalizations did not refer to

any spatial features of the situation at all (no adjuncts), while 30% referred to the location where the motion takes place by means of a locational adjunct (preposition + dative case, e.g., *X läuft auf einer Straße* “X is walking on a street”). By contrast, almost 100% of the L1 German speakers’ descriptions provided information on the path: References were made either to the trajectory (*die Straße entlang* “along the road”) or to some place at goal (*auf ein Dorf zu* “toward a village”). Note that particles are used in 57% of the utterances in the L1 descriptions, whereas this holds for only 15% of the L2 speakers. Thus, in the Type B situations, spatial information is low in the data of L2 speakers, in particular with respect to information on the path of motion of any kind. In the data of the German speakers, spatial information on the changes in place is expressed in almost 100% of the utterances.



The second data set analyzed shows situations in which boundary crossings occur in the stimulus videos (Type C). Again, similarities occur across the two groups with respect to verb type. Manner verbs occur in 92% of the L2 speakers' responses and in 83% of those of the L1 speakers. Consistency across the two groups can also be observed with respect to the spatial information provided. 80% of the utterances in both groups refer to the object that marks the goal point of the motion event by use of a satellite. However, there is a relevant difference in the use of verb particles. 53% of the L2 speakers' utterances explicitly encode the feature "goal reached" by means of the verb particles *rein/hinein* ("into/hither-in"). A sentence such as *die Frau läuft in den Supermarkt hinein* "the woman is walking into the supermarket hither-in" states that the person has actually crossed the borderline which is viewed as separating the outside and the inside of supermarket. In the data of the L1 speakers, particles which refer to the boundary crossing occur in only 10% of the utterances in contrast to the L2s. The L1 and L2 speakers of German differ with respect to the selection of a specific phase of the motion event, i.e., the post state.

We will now link the empirical data to our theoretical conception. Two of the processes that precede the articulation of an utterance when referring to a motion event are as follows: (1) the activation of event schemata and (2) the process of verbal encoding (Gerwien and von Stutterheim, in press). The event schemata provide the abstract conceptual gist that profiles the relevant spatial and temporal features of the representation that is mapped onto the external situation, which in this case concerns the visual input. In other words, an event schema pre-specifies conceptual relations between objects (figure, ground, goal, etc.). Importantly, event schemata are holistic representations. Similar to scripts, they form through experience across the life-span, and an essential assumption in our approach is that language is an important factor in this. In the initial stages of perception a schema is automatically activated (Zacks and Swallow, 2007; Zacks et al., 2007) in a bottom-up manner. Depending on the specific schema that has been activated, attention is directed to details of the external situation in a top-down fashion. Event schemata thus form the basis for information selection as well as for the interpretation of external situations. There are specific mechanisms that control how an event schema is held active

and under what circumstances the current schema is deactivated (Zacks and Swallow, 2007; Zacks et al., 2007). In the process of verbal encoding, the components of the active event schema are further specified by linking them to the respective linguistic means. Processes at the level of verbalization carve out further details that the activated event schema does not provide. On the other hand, if relevant components are not represented in the schema, they will not form part of the linguistic representation. In the present theoretical approach, the findings on differences present at the linguistic surface between L1 and L2 speakers point to the specific event schemata that speakers activate when exposed to different types of situations (see **Figure 6**).

The theoretical approach we propose for the domain of motion events links general theories of event cognition (Radvansky and Zacks, 2014) and research on typological differences in motion event encoding. It is specifically supported by findings from a recent study on event-unit formation in French and German that showed that speakers of the two languages form differently sized units when exposed to identical visual input showing motion events both in a verbal and in a non-verbal task (Gerwien and von Stutterheim, 2018).

In the present section, the empirical findings are presented from the perspective of event schema theory. The L1 German data represent what can be viewed as the standard use of event schemata in the German language community. Almost 100% of all utterances include reference to spatial information which is typically information on the path of motion. In situations with no evident endpoint (Type B), the activated event schema includes path information that can be derived from contours of ground (e.g., the trajectory). In situations with boundary crossing at a goal (Type C), the event schema includes path information that can be derived from relevant features of the goal. Since a specification of an aspectual perspective is not required in German, we assume that temporal features of this kind would not have relevant weightings in the event schema. The extent to which temporal aspectual features are relevant at all would require further research.

Since the L2 data deviate from the L1 data with respect to the occurrence of critical conceptual features in the utterances, this supports our theoretical assumption of selective schema activation. In situations with no evident endpoint (Type B), Tunisian L2 speakers frequently activate an event schema that focuses on manner of motion only and does not include information on the path. They therefore do not extract path information from contours of the ground to the same extent as speakers of their target language given the same situation type. In TA the expression of spatial information is constrained by the aspectual concepts attributed to a situation (imperfective/progressive vs. perfective, see templates in Section Results). The choice of a specific temporal aspect depends on the critical feature “boundary”, either the source or the goal of the path taken by the figure. The event schema that is typically activated in Type B situations (no evident boundary) provides information on temporal features—ongoing motion—and the downgrading of spatial information. The observation that Tunisian L2 speakers of German display a preference for this event schema can be directly explained by findings from a

previous study which show that the same event schema is activated in Tunisian L1 speakers who are exposed to the same stimuli as those in the current study (von Stutterheim et al., 2017).

In Type C situations the activated event schema of the L2 speakers appears, at first sight, highly similar to that of L1 speakers of German. It includes motion, as well as path features which can be derived from the presence of a goal point. However, the fact that the L2 speakers, in contrast to the L1 speakers of German, refer to the post state of the motion event by use of the particles *rein/hinein* (“into/hither-in”) in 50% of all utterances—that is, they mention that the figure has crossed the boundary at the goal—points to the activation of a different event schema. As we interpret this effect to indicate a means to compensate for the absence of aspectual markers in the German language (event ongoing, event completed), this schema specifies the temporal dimension in addition to the spatial dimension. The event schema that L1 German speakers activate, similar to what we concluded with respect to Type B situations, does not have relevant weightings for the temporal aspectual dimension.

Speakers of German with Tunisian as their L1 have to deal with the challenges posed by the fundamentally different criteria in the activation of event schemata. Different conceptual domains contribute to different degrees, and associated information has to be extracted from the input in order to comply with the relevant schema. The study shows how Tunisian L2 speakers of German differ from L1 speakers of German in what they encode. Their preferences clearly show similarities with the data of L1 speakers of Tunisian. This supports our claim that the differences arise at a pre-verbal stage, i.e., the stage at which information is selected for expression. This can be explained by the specific content of the event schemata which can be traced back to the syntactic and lexical features of the given language, e.g., the presence and role of grammatical aspect or that of spatial particles, as in the case of Tunisian and German. As language use can be understood as a cognitive experience, we assume that these language-specific event schemata, and their use, develop during language acquisition and are reinforced *via* day to day communication to form part of the pragmatic knowledge of a language community. The present study shows how this group of L2 speakers activates event schemata based on the conceptual templates described for their L1, as illustrated schematically in **Figure 5**.

Are there other factors which might explain the present results? There could be questions concerning the group of participants and the experimental design. Are the findings simply driven by a lack of proficiency of the L2 participants? There are substantial indications that this is not the case. The L2 speakers use manner verbs in both situation types, as is the case for the L1 speakers. There is also no further difference concerning the range of lexical items. In addition to the use of individual lexical items, the comparison of the use of more complex constructions shows that the L2 speakers have acquired constructions in which particles are combined with manner verbs, as evident in the Type C verbalizations (verb+particle_{boundarycrossing}). Although the L2 speakers have acquired these constructions (verb+particle_{path-object}), they are not used in Type B situations. Furthermore, their linguistic

competence is well advanced in that their verbalizations show that they have overcome the boundary crossing constraint which is a core feature of their L1 (manner verbs cannot combine with boundary crossing satellites in most verb-framed languages, as in Tunisian Arabic). Taken together, the present results do not indicate a lack of linguistic competence, at either the lexical or structural level, which would prevent this group of L2 speakers from using prepositional phrases as adjuncts or particles when describing information on the path of motion.

Are the differences driven by individual participants or items? The statistical models account for the variance that stems from random variables (participants and items) and the fixed factor (language) on a separate basis. Despite some degree of variance due to individual participants and items, the language group effect can be shown to prevail.

One could also ask how a different theoretical framework would explain the findings. Obviously, the concept of “conceptual transfer” could be an alternative in describing the L2 processes involved (Jarvis, 1998, 2011; Muñoz and Cadierno, 2019). Jarvis (2016) describes conceptual transfer as one type of cross-linguistic factor. This position states that if speakers do “not form the same mental representations of the same scene [...] then this would be a case of conceptual transfer assuming that the two groups differing patterns of conceptualization are directly tied to their language backgrounds.” (p. 610). From our point of view the theoretical position put forward by Jarvis is very similar to the present explanation. However, the term “transfer” would imply that “something” (in this case a conceptual representation) is “shifted” from one place to another. This would be at odds with our view on what information representation entails in speakers of more than one language, because it is unclear what should be located where. It seems that the original theoretical concept that gave birth to the term “transfer” implies that there are two (or more) conceptual modules (De Bot, 1992), or systems, between which some form of transfer can take place. However, we are not aware of reliable empirical evidence that would support a multi-component representational system of this nature, whereby one component “contains” one language and the other the other language. Seen from our theoretical standpoint, multilingual speakers may have one abstract repertoire for all event schemata. Features of the external world would lead to the automatic activation of the relevant event schema given one’s life-long experience. The experience given with the use of a specific L1 (language A) within a given language community of native speakers leads to a bias in the activation of certain event schemata in relation to a specific situation. Speakers with a different L1 (language B) may develop a different bias. If speakers with language A as their L1 set out to acquire language B as an L2, ultimate attainment would require them to detect the basis for activating a specific schema given a specific situation. This would mean that they can therefore ultimately “adjust” the bias stemming from the use of the L1. In our conception, the process would run from perception to schema activation to information selection to formulation. The network metaphor would therefore seem to be more adequate in describing the relevant sequential steps—compared to the metaphors of two

“modules” between which transfer processes can occur. Given an overall store of abstract conceptual templates (= event schemata), their automatic activation in a specific context is subject to implicit learning from experience.

The present study has crossed boundaries into new territory at two levels. It shows how temporal concepts are core constitutive factors in the expression of motion events in languages such as Tunisian Arabic in which aspectual categories are fully grammaticalized. As with L1 speakers of Tunisian when verbalizing a motion event, L2 speakers of German with Tunisian as their L1 show a high level of sensitivity to the temporal structure of scenes showing motion events. However, this sensitivity to the temporal dimension does not make motion event verbalizations more specific compared to those by German L1 speakers. The findings for the L2 speakers show how temporal information pushes spatial information into the background. In other words, the temporal and spatial conceptual domain can be shown to interact in the context of information selection. The necessity of going beyond spatial-cognitive categories in the analysis of motion events has called for a different theoretical approach which treats (motion) events as conceptual units, rather than a composition of different components (in the sense of Talmy). This has led us to link motion event cognition and schema theory (Bartlett, 1932; Brewer and Nakamura, 1984; Hintzman, 1986; Zacks, 2004; Gerwien and von Stutterheim, 2018) whereby event schemata form part of a generic knowledge system when interpreting the world and talking about it. As the present study on advanced second language speakers indicates, these schemata can be viewed as attentional filters at the interface between perception and selective representation, as well as at the interface between information selection from a knowledge base and verbal encoding. In fact, the L2 data presented in this article lend further support to the deeply entrenched language-specific nature of event schemata. The automatic and spontaneous selection of information for verbal expression in a specific linguistic environment requires an efficient cognitive mechanism, which, as we argue, event schemata can be seen to provide. Given the fact that the type of information that L2 speakers select and encode is identical to those that they select and encode in their L1, and since it is possible to explain why some information appears on the linguistic surface while other features do not, depending on the nature of the speakers’ language system, the current findings clearly underline the role of language in the formation of event schemata.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: osf.io/zs4ny/.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and

institutional requirements. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

JG and CS contributed equally to the manuscript. The statistical analysis was done by JG. Both authors contributed to the article and approved the submitted version.

REFERENCES

- Alderson-Day, B., and Fernyhough, C. (2015). Inner speech: development, cognitive functions, phenomenology, and neurobiology. *Psychol. Bull.* 141, 931–965. doi: 10.1037/bul0000021
- Arik, E. (2010). Describing motion events in sign languages. *Poznań Stud. Contemp. Linguist.* 46, 367–390. doi: 10.2478/v10010-010-0019-1
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., et al. (2015). Two languages, two minds: flexible cognitive processing driven by language of operation. *Psychol. Sci.* 26, 518–526. doi: 10.1177/0956797614567509
- Aurnague, M. (2011). How motion verbs are spatial: the spatial foundations of intransitive motion verbs in French. *Lingvist. Investig.* 34, 1–34. doi: 10.1075/li.34.1.01aur
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. London: Cambridge University Press.
- Bassetti, B., and Cook, V. (2011). “Relating language and cognition: the second language user,” in *Language and Bilingual Cognition*, eds V. Cook and B. Bassetti (Hove: Psychology Press), 143–190.
- Beavers, J., Levin, B., and Tham, S. W. (2010). The typology of motion expressions revisited. *J. Linguist.* 46, 331–377. doi: 10.1017/S002226709990272
- Brewer, W. F., and Nakamura, G. V. (1984). “The nature and functions of schemas,” in *Center for the Study of Reading Technical Report* (Champaign, IL: University of Illinois at Urbana-Champaign).
- Bürkner, P.-C. (2017). brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* 80, 1–28. doi: 10.18637/JSS.V080.101
- Bürkner, P.-C. (2018). Advanced bayesian multilevel modeling with the R package brms. *R J.* 10, 395–411. doi: 10.32614
- Bylund, E., and Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingual. Lang. Cogn.* 14, 47–59. doi: 10.1017/S1366728910000180
- Cadierno, T., and Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annu. Rev. Cogn. Linguist.* 4, 183–216. doi: 10.1075/arcl.4.08cad
- Carroll, M. (2000). Representing path in language production in English and German: Alternative perspectives on figure and ground. In *Räumliche Konzepte und sprachliche Strukturen*, eds C. Habel and C. von Stutterheim (Tübingen: Niemeyer), 97–118. doi: 10.1515/9783110952162.97/HTML
- Carroll, M., and von Stutterheim, C. (2011). “Event representation, time event relations, and clause structure: a crosslinguistic study of English and German,” in *Event Representation in Language and Cognition*, eds J. Bohnemeyer and E. Pederson (Cambridge: Cambridge University Press), 68–83.
- Carroll, M., Weimar, K., Flecken, M., Lambert, M., and Stutterheim, C., von. (2012). Tracing trajectories: motion event construal by advanced L2 French-English and L2 French-German speakers. *Lang. Interact. Acquisit.* 3, 202–230. doi: 10.1075/lia.3.2.03car
- Croft, W. A., Barðdal, J., Hollmann, W., Sotirova, V., and Taoka, C. (2010). “Revising Talmy’s typological classification of complex event constructions,” in *Contrastive Studies in Construction Grammar*, ed H. C. Boas (Amsterdam: Benjamins), 201–236.
- De Bot, K. (1992). A bilingual production model: levelt’s. *Applied Linguistics* 13, 1–24.
- De Knop, S. (2020). *Expressions of Motion Events in German: An Integrative Constructionist Approach for FLT*. Available online at: <http://journals.openedition.org/cognitextes>.
- Evans, V., and Chilton, P. (2007). *Language, Cognition and Space*. London: Equinox.

ACKNOWLEDGMENTS

For the publication fee we acknowledge financial support by Deutsche Forschungsgemeinschaft within the funding programme Open Access Publikationskosten as well as by Heidelberg University. We also thank two reviewers for their helpful comments on a first version of the manuscript.

- Flecken, M., Carroll, M., Weimar, K., and von Stutterheim, C. (2015). Driving along the road or heading for the village? Conceptual differences underlying motion event encoding in French, German, and French-German L2 users. *Modern Lang. J.* 99, 100–122. doi: 10.1111/j.1540-4781.2015.12181.x
- Flecken, M., Stutterheim, C., von, and Carroll, M. (2014). Grammatical aspect influences motion event perception: findings from a cross-linguistic non-verbal recognition task. *Lang. Cogn.* 6, 45–78. doi: 10.1017/langcog.2013.2
- Gelman, A., and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* 7, 457–472. doi: 10.1214/SS/1177011136
- Gerwien, J., and von Stutterheim, C. (2018). Event segmentation: cross-linguistic differences in verbal and non-verbal tasks. *Cognition* 180, 225–237. doi: 10.1016/j.cognition.2018.07.008
- Gerwien, J., and von Stutterheim, C. (in press). “Describing motion events,” in *Pragmatics of Space (HoPs 14)*, eds A. H. Jucker and H. Hausendorf (Berlin: De Gruyter Mouton).
- Gleitman, L., and Papafragou, A. (2005). “Language and thought,” in *Cambridge Handbook of Thinking and Reasoning*, eds K. J. Holyoak and B. Morrison (Cambridge: Cambridge University Press), 633–661.
- Haun, D. B. M., Rapold, C. J., Janzen, G., and Levinson, S. C. (2011). Plasticity of human spatial cognition: spatial language and cognition covary across cultures. *Cognition* 119, 70–80. doi: 10.1016/j.cognition.2010.12.009
- Hijazo-Gascón, A. (2021). *Moving Across Languages: Motion Events in Spanish as a Second Language*. Berlin: Mouton de Gruyter.
- Hintzman, D. L. (1986). “Schema abstraction” in a multiple-trace memory model. *Psychological Review* 93, 411–428. doi: 10.1037/0033-295X.93.4.411
- Jarvis, S. (1998). *Conceptual Transfer in the Interlingual Lexicon*. Bloomington, IN: Indiana University Linguistics Club Publications.
- Jarvis, S. (2011). Conceptual transfer: crosslinguistic effects in categorization and construal. *Bilingualism* 14, 1–8. doi: 10.1017/S1366728910000155
- Jarvis, S. (2016). Clarifying the scope of conceptual transfer. *Lang. Learn.* 66, 608–635. doi: 10.1111/LANG.12154
- Lambert, M., Stutterheim, C., and von Gerwien, J. (2022). Under the Surface. A survey on principles of language use in advanced L2 speakers. *Lang. Interact. Acquisit.* 13, 1–29. doi: 10.1075/lia
- Landau, B., and Jackendoff, R. (1993). Whence and whither in spatial language and spatial cognition? *Behav. Brain Sci.* 16, 255–265. doi: 10.1017/S0140525X00029927
- Levinson, S. C. (1996). Language and space. *Annu. Rev. Anthropol.* 25, 353–382. doi: 10.1146/annurev.anthro.25.1.353
- Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity*, Vol. 5. Cambridge: Cambridge University Press.
- Louhichi, I. (2018). The description of motion in Tunisian Arabic: a thinking-for-speaking approach. *Sino-US English Teach* 15. doi: 10.17265/1539-8072/2018.07.004
- Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., and Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends Cogn. Sci.* 8, 108–114. doi: 10.1016/j.tics.2004.01.003
- Miller, G., and Johnson-Laird, P. (1976). *Language and Perception*. Cambridge University Press.
- Muñoz, M., and Cadierno, T. (2019). Mr Bean exits the garage driving or does he drive out of the garage? Bidirectional transfer in the expression of Path. *Int. Rev. Appl. Linguist. Lang. Teach.* 57, 45–69. doi: 10.1515/iral-2018-2006
- Radvansky, G. A., and Zacks, J. M. (2014). *Event Cognition*. Oxford: Oxford University Press.

- Saidi, D. (2006). "Étude typologique de l'expression de déplacement en arabe tunésien" in *Autour des langues et du langage: perspective pluridisciplinaire*, eds M. Loiseau, M. Abouzaid, L. Buson, V. Goossens, C. Surcouf, et al. (Grenoble: Presses Universitaires de Grenoble), 119–126.
- Sarda, L. (2019). French motion verbs. In M. Aurnague and D. Stosic (Eds.), *The Semantics of Dynamic Space in French: Descriptive, Experimental and Formal Studies on Motion Expression*, Vol. 66 (John Benjamins Publishing Company), 67–107.
- Slobin, D. I. (1996). "From 'thought and language' to 'thinking for speaking.'" in *Rethinking Linguistic Relativity*, eds J. J. Gumperz and S. C. Levinson (Cambridge: Cambridge University Press), 70–96.
- Slobin, D. I. (2004). "The many ways to search for a frog: linguistic typology and the expression of motion events," in *Relating events in narrative: Typological and contextual perspectives*, eds S. Strömquist and L. Verhoeven (Mahwah, NJ: Lawrence Erlbaum Associates), 219–257.
- Slobin, D. I., and Hoiting, N. (1994). "Reference to movement in spoken and signed languages: typological considerations," in *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society: General Session Dedicated to the Contributions of Charles J. Fillmore*, 487–505. doi: 10.3765/BLS.V2011.1466
- Stutterheim, C., von, Lambert, M., and Gerwien, J. (2021). Limitations on the role of frequency in L2 acquisition. *Lang. Cogn.* 13, 291–321. doi: 10.1017/LANGCOG.2021.5
- Talmy, L. (2000). *Towards a Cognitive Semantics Vol.2: Typology and Process in Concept Structuring*. Cambridge: MIT Press.
- von Stutterheim, C., Andermann, M., Carroll, M., Flecken, M., and Schmiedtová, B. (2012). How grammaticized concepts shape event conceptualization in language production: insights from linguistic analysis, eye tracking data, and memory performance. *Linguistics* 50, 833–867. doi: 10.1515/ling-2012-0026
- von Stutterheim, C., Bouhaous, A., and Carroll, M. (2017). From time to space: the impact of aspectual categories on the construal of motion events: the case of Tunisian Arabic and Modern Standard Arabic. *Linguistics* 55, 207–249. doi: 10.1515/ling-2016-0038
- von Stutterheim, C., and Gerwien, J. (in Press). "Die Bedeutung sprachspezifischer Ereignisschemata für die Argumentstruktur. Ein Vergleich zwischen dem Ausdruck von Bewegungsereignissen im Deutschen und im Französischen," in *Propositionale Argumente: Theorie und Empirie (Studien zur deutschen Sprache: Forschungen des Instituts für deutsche Sprache)*, eds J. Hartmann and A. Wöllstein, A (Tübingen: Narr Francke Attempto).
- von Stutterheim, C., Gerwien, J., Bouhaous, A., Carroll, M., and Lambert, M. (2020). What makes up a reportable event in a language? Motion events as an important test domain in linguistic typology. *Linguistics* 58, 1659–1700. doi: 10.1515/ling-2020-0212
- Vulchanova, M., and van der Zee, E. (eds.) (2013). *Motion Encoding in Language and Space*. Oxford: Oxford University Press.
- Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cogn. Sci.* 28, 979–1008. doi: 10.1016/j.cogsci.2004.06.003
- Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., and Reynolds, J. R. (2007). "Event perception: a mind-brain perspective. *Psychological bulletin*, 133, 273.
- Zacks, J. M., and Swallow, K. M. (2007). Event segmentation. *Curr. Direct. Psychol. Sci.* 16, 80–84. doi: 10.1111/j.1467-8721.2007.00480.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Gerwien and von Stutterheim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.