


The effects of bilingual proficiency on the acceptability of motion encoding strategies

Jean Costa-Silva , Shulin Zhang and Vera Lee-Schoenfeld

Department of Linguistics, University of Georgia, Athens, Georgia, USA

Research Article

Cite this article: Costa-Silva, J., Zhang, S. and Lee-Schoenfeld, V. (2025). The effects of bilingual proficiency on the acceptability of motion encoding strategies. *Bilingualism: Language and Cognition* 28, 444–456. <https://doi.org/10.1017/S136672892400052X>



Received: 15 June 2023
Revised: 02 June 2024
Accepted: 09 June 2024
First published online: 27 January 2025

Keywords:

motion events; proficiency; acceptability ratings; Brazilian Portuguese; English

Corresponding author:

Jean Costa-Silva;
Email: jeancosta@uga.edu

  This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

When describing motion events, English encodes Manner of motion in the verb and Path of motion in a satellite (s-framing). Brazilian Portuguese (BP), however, encodes Path in the verb and elaborates Manner adverbially (v-framing). This study investigates at what stages of L2 proficiency L2BP and English learners' acceptability ratings converge with those of L1 speakers when rating sentences with Manner elaborated in the verb (Manner-verb) or in an adverb: a prepositional phrase (Manner-prep) or subordinate clause (Manner-AdvClause). Participants ($n = 176$) consisted of L1/L2English and L1/L2BP speakers. L2ers were grouped according to language proficiency (Elementary, Intermediate and Advanced). Results of ordinal logistic regressions show that Intermediate proficiency is associated with Manner-verb (L2BP) and Manner-prep ratings (L2English), and that Advanced proficiency is associated with Manner-verb (L2English) and Manner-AdvClause judgments (L2English and L2BP). These findings contribute to the limited body of work on the acquisition of v-framed L2s and the development of low-proficiency learners.

Introduction

The relationship between language and thought has been one of the central pursuits of Cognitive Linguistics (Danhier & Mertins, 2016; Lucy, 1992, 1996; Slobin, 1996; Stam, 2010, among many others). Empirical research shows that thinking is completely possible without the aid of language; however, language assists in the habitual classification of the world into meaningful categories (Bylund & Athanasopoulos, 2014; Lucy, 1997; Slobin, 1996; Stam, 2010). Danhier and Mertins (2016), for instance, argue that language has an effect on how speakers of different languages perceive and describe the world. This effect can be traced back to the abstract planning stages of speech, which are influenced by the grammar and lexicon of the speakers' mother tongue (Danhier & Mertins, 2016). Their proposal is somewhat aligned with Slobin's (1996) seminal work, which posits that there is a specific type of thinking that is performed in the process of speaking and that this thinking is intimately tied to language. As Slobin (1996) puts it, the world itself does not present situations that need to be encoded in language – what happens is that one fits their thoughts into the linguistic frames of the language in their repertoire. Essentially, experiences are FILTERED through language. As we carry out this process of filtering, we perform the activity of thinking-for-speaking.

The THINKING-FOR-SPEAKING HYPOTHESIS (Slobin, 1996, 2000, 2004) proposes that language acts as a mechanism that drives speakers' attention to specific perceptual attributes of reality. The hypothesis has inspired a significant body of work on the perception and description of color (Athanasopoulos, 2009), objects (Pavlenko & Malt, 2011), time (Stam, 2010), as well as issues pertaining to lexical and grammaticalized concepts of space (Alloway & Corley, 2004; Cadierno & Ruiz, 2006; Filipović, 2011; Hasko, 2009; Pavlenko & Driagina, 2006; Pavlenko & Volynsky, 2015). Space is of particular interest because it consists of a 3D experience that needs to be fit into language, which is – with the exception of sign languages – one-dimensional and linear (Danhier & Mertins, 2016). When a speaker translates their experience with space into a language, they need to reduce the number of dimensions they describe (Danhier & Mertins, 2016). Therefore, language plays a role in directing speakers to decide what is left out (Slobin, 1996, 2004).

The Talmyan typology of motion events

Leonard Talmy's typology (Talmy, 1985, 1991, 2000) is one of the most prominent works to describe how languages structure the domain of Space and investigate the general nature of cognitive representation (Batoréo, 2014). Talmy classifies languages based on the relationship between surface structures and four main meaning components (Talmy, 1985, 1991; Slobin, 2004; Pavlenko & Volynsky, 2015). The first of these components, FIGURE, consists of the moving object, which is typically a nominal element (*The bus is approaching the school*). The second component, GROUND, is the reference point for the Figure and is also prototypically a

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

nominal element (*The bus is approaching the school*). The third element, PATH, refers to the course traveled by Figure in relation to the ground (*The child walks into the classroom*). The fourth component, MOTION, may be self-propelled or caused (*The child walks into the classroom*).

Although the expression of movement in space is a language universal, there is significant variation in how languages encode motion (Hasko, 2009). Based on their encoding strategy, Talmy (1985, 1991, 2000) proposes that languages can be categorized into three groups. The first group, named SATELLITE-FRAMED (S-FRAMED) LANGUAGES, consists of those that adopt a CONFLATED STRATEGY when describing motion:

1. The woman **ran down** the hill.

That is, the Manner of motion is expressed in the main verb of the clause, while the Path of motion is encoded in a satellite (a preposition, a particle or a prefix) (Hasko, 2009; Lewandowski, 2021). Because of the morphosyntactic structures, these languages allow the stacking of multiple satellites to a single verb, which allows for more detailed descriptions of Path within a single clause (Slobin, 2004). German and Slavic languages are prototypical examples of this group (Hasko, 2009; Filipović, 2011; Lewandowski, 2021; Ragnarsdóttir & Strömquist, 2004; Slobin, 2004).

The second group, named VERB-FRAMED (V-FRAMED) LANGUAGES, are those that employ a SEPARATED STRATEGY when encoding motion:

2. A mulher **desceu** o morro **correndo**. [Portuguese]
'The woman went down the hill running.'

In (2), the Path of motion is expressed in the main verb and Manner, if present, is encoded in an additional subordinate clause (Batoréo, 2014; Hasko, 2009; Lewandowski, 2021; Slobin, 2004). Languages in this group use few Motion verbs with generic meaning (e.g., *descer* 'descend/go down', *subir* 'ascend/go up', *entrar* 'enter/go in' *sair* 'exit/go out', in Portuguese). V-framed languages seem to treat the elaboration of Manner as narrative luxury: manner is encoded when speakers choose to do so (Slobin, 2004). Prototypical examples are Romance Languages, Hebrew and Turkish. The third group, named EQUIPOLLENTLY FRAMED LANGUAGES, is outside the scope of this study (see Beavers et al. (2010), Chen and Guo (2009), Slobin (2004) and Talmy (2012) for an in-depth discussion).

A significant subsequent body of work has shown that languages rarely rely on one lexicalization pattern exclusively: their affiliation to a typological group seems to be based on the most frequently encoding strategy employed (Lewandowski, 2021; Meirelles, 2019; Pavlenko & Volynsky, 2015). For instance, from a purely syntactic standpoint, both languages (English and Brazilian Portuguese) analyzed in this study accept structures such as (3):

3. I **go** to work **by car**.

In (3), Path is assigned to the main verb and Manner of motion is encoded in a prepositional phrase. This mixed strategy conflates Manner and Path information within the same clause (s-framed strategy) while still treating Manner as adverbial and optional (v-framed strategy). English, while a prototypical S-framed language, allows – at least to a certain degree of acceptability, the realization of Path verbs followed by adverbial prepositional phrases – as in (3) as well as in Manner constructions

(*The child enters the room walking*). However, the overall preference of English for Manner verbs and Path satellites places it within the S-framing range (Batoréo, 2014). Brazilian Portuguese, on the contrary, shows flexibility – presenting more restricted descriptions of motion (Batoréo & Ferrari, 2016) and disfavoring Manner verbs in the main clause especially in instances in which the main verb requires an inanimate direct complement (**Eu dirigi minha irmã para a escola*. Lit., 'I drove my sister to school').

There are also issues of intratypological variation (Hasko, 2009; Pavlenko & Volynsky, 2015; Slobin, 2004; Verhoeven & Stromqvist, 2004). Studies comparing Germanic and Slavic languages find that, despite being placed within the same group, these languages encode motion in significantly different ways (Gagarina, 2009; Hasko, 2009; Pavlenko & Volynsky, 2015; Lewandowski, 2021). In their work on the acquisition of Russian motion verbs by L1 English speakers, Hasko (2009) and Pavlenko and Volynsky (2015) find that the semantic repertoires of the languages are not parallel, and L1 English-L2 Russian learners will encode Manner in a less fine-grained way than L1 Russian speakers. Similar results are found by Lewandowski (2021) when comparing L1 German-L2 Polish and L1 Polish-L2 German descriptions of motion to that of monolinguals. Their conclusion is that German shows a higher tendency to encode Manner than Polish and that Russian shows a higher tendency than English. Degrees of variability are observed even between two variations of the same language. Batoréo (2014); Batoréo and Ferrari (2016) and (Meirelles, 2019) show that European Portuguese elaborates Manner in more fine-grained ways than the Brazilian variation. To address issues like this, Slobin (2004) proposes that languages should, instead, be ranked on a cline of Manner salience based on how salient the encoding of Manner is in the language.

Thinking-for-speaking in the L2

As language provides at the very least some degree of input to cognition (Allen et al., 2007; Cadierno, 2017; Grigoroglou & Ganea, 2022; Hasko, 2009; Slobin, 1996; Stam, 2010), it is no wonder that thinking-for-speaking is of interest for the field of Second Language Acquisition (SLA). Referred to as ALTERNATIVE WAYS OF THINKING-FOR-SPEAKING, RETHINKING-FOR-SPEAKING OR THINKING-FOR-SPEAKING IN THE L2 (Cadierno, 2004; Cadierno & Ruiz, 2006; Cadierno, 2008; Ellis & Robinson, 2008), the hypothesis suggests that when the L2 patterns of thinking-for-speaking are different from the speaker's L1's, learners need to acquire new patterns to reach higher levels of L2 proficiency (Stam, 1998, 2010). As described by Choi and Bowerman (1991) and Slobin (2004), children acquire their L1 patterns of thinking-for-speaking at least as early as in their first year of life. Since these patterns are resistant to restructuring in late second language acquisition (Slobin, 1996), analyzing the challenging areas for mastering the additional language can help us define what the patterns are in the learner's L1 (Stam, 2010). Similarly, as Slobin (1996) argues, L2 acquisition of these patterns incurs learners determining what must be attended to and expressed in the L2.

Multiple authors have investigated how thinking-for-speaking develops in a second language (Cadierno, 2004; Cadierno & Ruiz, 2006; Elliott & Yountchi, 2009; Gagarina, 2009; Hasko, 2009; Lewandowski, 2021; Pavlenko & Volynsky, 2015; Stam, 2015). Motion is, as previously discussed, of great interest due to the

reduction of dimensions it undergoes in the process of encoding, which means language plays a role in determining what is left out. Several studies focused on speech (see Stam, 2015, for a summary) and encountered mixed results, showing that acquiring the L2 thinking-for-speaking patterns is a challenging task, especially when typologically distinct languages are considered (Hasko, 2009; Gagarina, 2009; Nogueira, 2009).

Several studies illustrate these challenges. Cadierno (2004) shows that the influence of the learner's L1 thinking-for-speaking patterns in their L2 occurs regardless of the typological directionality. In her study, speakers of S-framed languages used adverbs redundantly in their V-framed L2s and demonstrated difficulty in describing motion events that express crossing a physical boundary (Cadierno, 2004). When the opposite was observed (i.e., speakers of V-framed language acquiring an S-framed L2), students showed difficulty in verbalizing trajectory dynamically and using locative expressions (usually prepositional phrases), as well as in employing a wide range of Manner verbs. More advanced L2 learners were able to develop appropriate patterns of thinking for speaking in the L2, but this did not occur in the same way for all aspects of a motion event (Cadierno, 2004). Ibarretxe-Antuñano et al. (2016) found that both L2 Danish and Spanish intermediate learners had difficulties in reconstructing L2 verb meanings, and that, even though they were familiar with the basic L2 placement verbs, their choice and usage differ significantly from that of native speakers. Similarly, Ozyurek (2002) showed that Turkish learners of English as an additional language followed thinking-for-speaking patterns of their L1. The L1 influence was also observed in languages within the same typological group. In an investigation on the locus of difficulty L1 English L2 Russian learners, Hasko (2009) concluded that the choice of contextually appropriate encoding of Manner of motion was a problem that persisted through learners' narratives.

Cadierno (2017) points out that bidirectional transfer can also occur, as illustrated by multiple studies (see Athanasopoulos et al., 2015; Brown & Gullberg, 2010, 2011; Bylund & Athanasopoulos, 2014; Cadierno & Ruiz, 2006). Brown and Gullberg (2008, 2011, 2010, 2013) carried out extensive work on the acquisition of L2 Japanese and L2 English and found bidirectional influence for intermediate and advanced learners, showing that the patterns of L1 thinking-for-speaking patterns affect the L2 ones, and vice-versa. As Putnam (2019) puts it, this clash between L1 and L2 systems often leads to more gradient representations in the bilingual speaker who may opt for a linguistic form that is neither typical of their L1s nor their L2s. For late bilinguals, rethinking-for-speaking entails the restructuring of cognitive aspects of learning, constant monitoring and the competition between L1 and L2 systems during perception and production (Putnam, 2019; Selinker & Gass, 2008). Kellerman (1995) suggests that, unless the development is mediated, adult second language learners may not be completely aware of what these patterns look like and may learn L2 linguistic forms, but use them from an L1 perspective.

Some studies suggest that late bilinguals do acquire the thinking-for-speaking patterns of their L2. Stam (1998, 2006) and Stam and McCafferty (2009) showed that when L2 English learners narrated in English, they showed mixed L1 and L2 patterns of thinking-for-speaking. These results reflected their interlanguage development and indicated an overall improvement. Cadierno and Ruiz (2006) compared the expression of Path

and Manner of Motion by Danish learners of Spanish, Italian learners of Spanish and Spanish native speakers. Despite the expectation that Danish learners would show a higher degree of Path and Manner elaboration, the results displayed a limited influence of the L1 thinking-for-speaking patterns on the speech of Advanced learners. This is particularly encouraging, as Danish was only typologically different language in the group. Lewis's (2012) investigation of L1 English speakers learning Spanish while studying abroad showed that learners were able to demonstrate L2 patterns in their descriptions of path after only 6 months of immersion. Their findings suggest that learners can acquire their L2 patterns of thinking-for-speaking but do not specify to what extent (Lewis, 2012). As for the acquisition of English by Portuguese speakers. The few studies that have been carried out reported positive results (Nogueira, 2009; Chiappara, 2022). However, these focused on a significantly small number of participants (an average of 15 individuals) and only focused on a particular proficiency range.

Some questions about the acquisition of thinking-for-speaking patterns in the L2 remain unanswered. These include the effects of age of onset, the amount of exposure or the types of instruction (Cadierno, 2017; Stam, 2015). Cadierno (2017) points out that, while a significant body of research has focused on the acquisition of s-framed languages, work on the acquisition of v-framed ones remains scarce. Mengali (2020) underscores this scarcity and emphasizes that this is particularly true for Portuguese as a second language. Cadierno (2017) also points out that more work that includes data from early learners is needed: in her comprehensive review of the field, she showed that most of the work has focused on Intermediate and Advanced learners.

The present study fills some of these gaps in the literature by (i) analyzing the acquisition of Portuguese, a v-framed language, by L1 speakers of English; (ii) introducing results collected from a large sample of Elementary and Intermediate L2 speakers; (iii) providing data on the acquisition of English by L1 Portuguese speakers with a number of participants that much supersedes the average of 15 individuals.

The present study

This investigation focuses on the acquisition of English as an additional language by L1 Brazilian Portuguese speakers as well as Portuguese as an additional language by L1 English speakers. From a theoretical standpoint, English is an s-framed language that prototypically encodes Manner of motion in the main verb and Path of motion in a satellite. Brazilian Portuguese, on the other hand, is a canonical V-framed language that encodes Path of motion in the main verb, and Manner of motion is elaborated in an optional adverbial clause (Almeida, 2002; Batoréo, 2014; Batoréo & Ferrari, 2016).

Participants were asked to rate sets of sentences in which Manner of motion is encoded by employing the canonical s-framed structure: (i) Manner + Path (*The woman walks into the room*), as well as two prototypical v-framed construction types: (ii) Path Verb + Manner Adverbial Clause (*The woman enters the kitchen walking*) and (iii) Path Verb + Manner Adverbial Prepositional Phrase (*The woman enters the kitchen on foot*). The choice for the novel model of employing an acceptability judgment task followed Hwang (2023), who argues that it imposes less cognitive burden on participants.

This study aimed to answer three questions – one of a more typological focus and two centered on bilingual development:

RQ1. Considering that both English and Brazilian Portuguese allow descriptions of motion in with a Path verb and an optional Manner prepositional phrase, how does this structure rank compared to the canonical structures of each language?

RQ2. At what stage of second language development (Elementary, Intermediate, Advanced) do L1 English L2 learners of Portuguese (as v-framed language) begin to rate motion encoding structures as acceptable as monolingual L1 Portuguese speakers do?

RQ3. At what stage of second language development (Elementary, Intermediate, Advanced) do L1 Portuguese L2 learners of English (as V-framed language) begin to rate motion encoding structures as acceptable as monolingual L1 English speakers do?

For **RQ1**, we anticipate that Portuguese speakers will consider the Path verb + Manner prepositional phrase more acceptable than the prototypical Path verb + Manner adverbial clause. This will occur due to the lower complexity of a Manner prepositional phrase compared to an adverbial clause, which entails the assignment of an argument structure (Chomsky, 1957). We also anticipate that English speakers will consider Path verb + Manner prepositional phrase more acceptable than Path verb + Manner adverbial clause for a similar reason but will rank it as less acceptable than the canonical Manner verb + Path satellite constructions.

For **RQ2** and **RQ3**, we expect that learners at earlier stages will demonstrate higher acceptability of sentences that reflect the thinking-for-speaking strategies of their L1 (Cadierno, 2004, 2017). We also expect that this should change over time, and late bilingual speakers will begin to assign higher ratings to sentences that show the pattern of thinking-for-speaking of their L2 as early as at the Intermediate level (Cadierno & Ruiz, 2006). However, we expect that learners of Portuguese as an L2 will consider Path verb + Manner prepositional phrase less acceptable than monolingual speakers of Portuguese do due to the permissibility of the structure in both languages, but the preference for v-framing in Portuguese. Learners of English as an L2 will consider Path verb + Manner adverbial clause more acceptable than monolingual speakers of English do due to the grammaticality of the structure in English, despite v-framing being less preferable.

Because Manner encoding in the verb results in a larger verb lexicon in s-framed languages (Slobin, 2004), learners of L2 English will consider S-framing structures acceptable relatively early due to the salience of elaborate description of Manner in the L2. Contrastingly, learners of L2 Portuguese will consider v-framing structures less acceptable and also relatively early because of the low occurrence of Manner verbs in the language.

Participants

Two hundred and eight participants were recruited via social media announcements as well as collaborations with higher education institutions in Brazil, Canada and the United States. To participate, individuals needed to be 18 years old. Participants consisted of monolingual speakers of Brazilian Portuguese ($n = 20$), monolingual speakers of English ($n = 31$), L2 English learners whose first language was Brazilian Portuguese ($n = 88$), and L2 Portuguese

learners whose first language was English ($n = 69$). The study collected informed consent from participants prior to its start.

To isolate the effects of L3 transfer on the learners' English and Portuguese, we excluded from the data L2 English learners who reported an Intermediate level in another s-framed language as well as L2 Portuguese learners who reported the same level in another v-framed language. L2 learners also rated ungrammatical sentences (i.e., sentences that violated the target language syntax) on a scale from 1 to 6. Participants who assigned 4 or higher to any of the sentences were also removed from the study ($n = 26$). No L1 speaker rated any of the ungrammatical sentences as acceptable. Finally, we also removed incomplete responses from the data ($n = 6$).

This procedure left us with one hundred and seventy-six participants. The first group ($n = 27$) consisted of monolingual L1 speakers of English ('L1 English') aged between 18 and 63. The second group ($n = 19$) was made up of monolingual L1 speakers of Portuguese ('L1 Portuguese') aged between 22 and 58. The third group ($n = 73$) consisted of L2 English speakers whose L1 is Brazilian Portuguese ('L2 English'). These participants' age ranged between 18 and 39. They had an average of 13.07 years learning the L2 and an average age of onset of 19.87. The fourth group ($n = 57$) consisted of L2 Brazilian Portuguese speakers whose L1 is English ('L2 Portuguese'). Their age ranged between 18 and 55. They had an average of 1.8 years learning their L2 and an average age of onset of 1.80. While both groups' average age of onset was early adulthood, the amount of time spent learning their L2 varied significantly.

See **Table 1** for background information on participants.

Proficiency measurements

L2 English and L2 Portuguese learners took a self-reported proficiency test designed at the Michigan State University. The administration of a self-reported exam follows recent trends in the field: they are adequate for low-stakes L2-proficiency measurements because of their low cost, and they make intuitive sense to learners (Winke et al., 2023). The exam is aligned with the American Council on the Teaching of Foreign Languages (ACTFL) standards for language learning and places test-takers in one of five bands: (i) Level 1: Novice Low to Novice High, (ii) Level 2: Novice High to Intermediate Mid, (iii) Level 3: Intermediate Mid to Advanced Low, (iv) Level 4: Intermediate High to Advanced Mid and (v) Level 5: Advanced Mid to Superior. The self-reported proficiency test is divided into five sections. In each section, participants rated ten language-related Can-Do statements according to their perceived language skills (e.g., *I cannot do this yet, with much help, with little*

Table 1. Background information of participants

Group	Age at testing	Age of onset	Years learning L2
L1 English ($n = 27$)	34.48 (SD = 12.92; range = 18–63)	NA	NA
L1 Portuguese ($n = 19$)	36.11 (SD = 10.76; range = 22–58)	NA	NA
L2 English ($n = 73$)	32.95 (SD = 11.39; range = 18–39)	19.87 (SD = 12.22; range = 4–60)	13.07 (SD = 11.22; range = 0.08–45)
L2 Portuguese ($n = 57$)	20.81 (SD = 5.27; range = 18–55)	19.00 (SD = 2.61; range = 13–32)	1.80 (SD = 4.30; range = 0.08–30)

help, I can do it well). Participants were also asked to rate whether the skills described were important to them. The test performs conservative scoring: each set of 10 statements has 10 possible points, with one point awarded if the person selected the highest ability level (mastery, a '4') on the Likert-scale (Tigchelaar et al., 2017). A participant who scored 8 out of 10 on a set was able to advance to the next set. There was a total of 5 sets: one for each of the test levels. The conservative approach was adopted to prevent over-assignment in the higher levels (Tigchelaar et al., 2017).

To add another layer of validity to their self-reported results, we analyzed whether there was a correlation between the self-reported proficiency exam and participants' years of learning the L2. We expected a positive correlation between the factors. The Kendall rank correlation test showed a strong positive correlation ($t = 0.64$, $Z = 7.65$, $p < 0.001$) for L2 English learners and a moderate positive correlation for L2 Portuguese learners ($t = 0.52$, $Z = 5.03$, $p < .001$). This means that their proficiency assessment is coherent with the length of exposure.

Participants' proficiency levels can be seen in Table 2.

As the self-reported proficiency exam places learners in one of five levels – which would result in many small subgroups – we decided to conflate the five test levels into three bands: Elementary, Intermediate and Advanced. Our aim was to find an adequate scope that was neither too broad to generalize or too narrow for patterns to be identified. First, we correlated the ACTFL levels with the Common European Framework of Reference (CEFR) bands. We noted that the one CEFR band ranged over one or two ACTFL levels, i.e., both test levels 1 and 2 fit within the CEFR's 'Basic User' (Elementary); similarly, levels 3 and 4 were equivalent to the CEFR's 'Independent User' (Intermediate). Level 5 corresponded to the CEFR's 'Proficient User' (Advanced). This system allowed us to reduce the number of subgroups from 5 to 3.

Methods

The decision to use an acceptability judgment task was motivated by multiple factors. First, as seen in Hwang (2023), they are more sensitive to nuances of grammar and impose less cognitive burden on learners. Second, they have a high potential to assess specific

target structures (Grey & Tagarelli, 2018). Third, they are easy and efficient to administer, as they can be applied via the internet. Acceptability judgments are also helpful for late bilingual learners who may demonstrate understanding of an L2 semantic feature, yet not have undergone the process of automatization (i.e., the conscious, controlled processing of declarative knowledge in natural speech) (Hasko, 2009; Selinker & Gass, 2008).

Participants were asked to rate 75 sentences on a 1–6 Likert scale. They were instructed to select 1 for sentences they deemed 'not acceptable' and 6 for those they considered 'completely acceptable'. A six-point Likert scale was selected to avoid the selection of a middle point and invite participants to consider the items of measurement (as argued for by Chomeya (2010)). The 75 sentences were distributed as: 24 target structures and 48 distractors – a format similar to the one used by Hwang (2023), as well as 3 extra ungrammatical sentences (sentences that violated subject–verb agreement and word order) to add another layer of validity to their proficiency measurement results. L2 participants who rated ungrammatical sentences with a 4 or higher were excluded from the study. No L1 participant rated an ungrammatical sentence with a score higher than a 1.

The target structures were elaborated following Gagarina (2009)'s list of common Manner verbs. We analyzed their absolute frequency in the *Contemporary Corpus of American English* (COCA) and selected the top 4 most frequent verbs: 'walk', 'swim', 'drive' and 'fly'. For each of these structures, we collected two sentences from COCA, which we manipulated to display the Path verb and Manner prepositional phrase as well as Path verb and Manner adverbial clause structures:

4. (a) I walk into the kitchen and ask him for the phone.
[Manner verb + Path satellite]
- (b) I enter the kitchen **walking** and ask him for the phone.
[Path verb + Manner adverbial clause]
- (c) I enter the kitchen **on foot** and ask him for the phone.
[Path verb + Manner prepositional phrase]

Brazilian Portuguese sentences underwent the opposite process. To determine the corpus for the BP analysis, we searched for constructions that were structurally opposite from the English patterns (Path+Manner vs Manner+Path). We ran these in the *NOW: Corpus do Português* and located samples that also displayed high absolute frequency in Brazilian Portuguese. This allowed us to find direct correspondents in both languages. The BP constructions consisted of *entrar* 'to enter/walk in(to)', *atravessar* 'to cross', *levar* 'to take/drive' and *viajar* 'to travel/fly'. These sentences were also manipulated from their canonical Path verb+ Manner adverbial clause or Path verb + Manner prepositional phrase to the Manner verb + Path satellite construction:

5. (a) *Ronaldo entrou na cozinha andando*. [Path verb + Manner adverbial clause]
'Ronaldo entered the kitchen walking.'
- (b) *Ronaldo **entrou** na cozinha a pé*. [Path verb + Manner prepositional phrase]
'Ronaldo entered the kitchen on foot.'
- (c) *Ronaldo **andou** para dentro da cozinha*. [Manner verb + Path satellite]
'Ronaldo walked into the kitchen.'

L1 speakers filled out a background questionnaire which collected information about their age, experiences abroad (visits and

Table 2. Participants' Distribution Based on Proficiency Levels

Group	Age at testing	Age of onset	Years learning L2
L2 English Elementary (n = 33)	31.27 (SD = 11.82; range = 18–70)	26.87 (SD = 12.17; range = 13–60)	4.40 (SD = 3.72; range = 0.08–15)
L2 English Intermediate (n = 15)	35.47 (SD = 13.12; range = 18–57)	16.34 (SD = 16.34; range = 4–44)	19.5 (SD = 19.5; range = 4–43.7)
L2 English Advanced (n = 25)	33.64 (SD = 9.12; range = 23–58)	12.96 (SD = 7.16; range = 4–43)	20.67 (SD = 8.24; range = 10–45)
L2 Portuguese Elementary (n = 43)	19.86 (SD = 2.58; range = 18–33)	19.29 (SD = 2.56; range = 17.5–32)	0.59 (SD = 0.42; range = 0.08–1.5)
L2 Portuguese Intermediate (n = 10)	24.20 (SD = 10.47; range = 19–55)	19.07 (SD = 2.11; range = 17.5–25)	5.12 (SD = 8.53; range = 0.91–30)
L2 Portuguese Advanced (n = 4)	22.25 (SD = 3.36; range = 18–26)	15.73 (SD = 2.09; range = 13.5–18)	6.52 (SD = 3.89; range = 1.6–12)

stays), proficiency in additional languages and academic and professional background. The aim was to prevent their knowledge of a foreign language of the opposite typological group of the tested language from having an effect on the results. It also served to determine whether they had spent a significant amount of time in a target-language-speaking country. A period of more than 30 days was considered significant. L2 speakers completed a similar background questionnaire that also included questions about the amount of time spent learning the additional language, method of instruction and age of onset. None of their answers to these questions excluded them from the study.

L1 participants completed an acceptability judgment task of sentences in their native language. The aim of this type of assessment was two-fold: (i) it provided us with a baseline against which bilingual ratings could be compared and (ii) it allowed us to assess the preference for Path verb and optional Manner prepositional phrases in relation to Path verb and optional Manner adverbial clause (structures that are licit in both languages). As this study focuses on unidirectional effects (i.e., the effects of the L1 in the L2 only), L2 participants completed the preference task in their additional language. To reduce the impact of low language proficiency in Elementary L2 speakers' results, glosses of less frequent words were provided to all L2 speakers.

Data analysis

We ran two main types of analysis of the data: within-group and between-group comparisons. Within-group comparisons helped us establish a baseline in the languages that informed us (i) what acceptability ratings should be expected from bilingual speakers in relation to monolingual speakers; (ii) whether speakers of each language displayed any preference between the two verb-framing structures (Path verb + Manner prepositional phrase or Path verb + Manner adverbial clause), which are acceptable in both BP and English. To perform this analysis, we created a model for ordinal logistic regression (Cumulative Link Mixed Model fitted with the Laplace approximation) using the `clmm()` function from the ordinal package in the R software (R Core Team, 2021). Our motivation for employing this type of regression was twofold. From a theoretical standpoint, as seen in Verissimo (2021), it is inappropriate to assume that ordinal data shows equidistance between points – a requirement of metric methods such as ANOVA or linear regressions. From a practical perspective, ordinal logistic regressions demonstrate higher predictive power for models with ordinal data and multiple independent variables, as seen in Kissling (2018) and Tare et al. (2018).

We created a model with one dependent variable (Ratings), one fixed effect (Pattern) and two random effects (Participant and Events).¹ The patterns were coded as Manner+Path (sentences with a Manner verb followed by a Path satellite within the same clause), Path+Verb (sentences with a Path verb followed by a Manner subordinate adverbial clause) and Path+Prep (sentences with a Path verb followed by a Manner prepositional phrase, also within the same clause). Our formula consisted of: Ratings ~ Pattern + (1| Participant) + (1| Events). We also ran the `emmeans()` function to contrast the variables.

¹We included *Events* as a random variable after one of the reviewers suggested we accounted for the difference in our data: some sentences consisted of one clause with one motion event while others consisted of two clauses with two motion events. No effects were found.

Between-group analysis helps us determine whether acceptability ratings change across proficiency levels, in particular, in relation to those provided by monolingual speakers. For this analysis, we created a second model for ordinal logistic regression (Cumulative Link Mixed Model fitted with the Laplace approximation) again using the `clmm()` function from the ordinal package in the R software (R Core Team, 2021). Our second model included one dependent variable (Ratings), one fixed effect (Proficiency) and two random effects (Participant and Events).² Proficiency included in the model consisted of Monolingual (L1 speakers), Elementary, Intermediate and Advanced. Our formula consisted of: Ratings ~ Proficiency + (1| Participant) + (1| Events). As in the first model, we ran the `emmeans()` function for contrast.

These approaches allowed us to test (i) whether L1 speakers demonstrate differences in their acceptability ratings of Path Verb + Manner prepositional phrase and Path Verb + Manner adverbial clause (RQ1) and (ii) whether L2 learners' acceptability ratings begin to converge with those of L1 speakers as they become more proficient in their L2 (RQ2 and RQ3).

Results

The graphed results (Figures 1 and 2) give us an overview of the tendencies in both L1 and L2 speakers' ratings across proficiency levels and pattern types. For the sake of convenience, we refer to structures with Manner encoding in the verb and Path encoding in a satellite as MANNER-VERB, Path encoding in the verb and Manner encoding in a prepositional phrase as MANNER-PREP and Path encoding in the verb and Manner encoding in an adverbial phrase as MANNER-ADVCLAUSE. In this section, we present the results of both within- and between-group comparisons.

Within-group comparisons: monolinguals

The results of the ordinal logistic regression (Table 3) show that monolingual speakers of English ('L1 English') rated Manner-verb structures as more acceptable than both Manner-prep (OR = -0.44, $Z = -2.56$, $p = 0.02$) and Manner-AdvClause constructions. However, the difference in the acceptability ratings of Manner-verb and Manner-AdvClause structures was not statistically significant (OR = -0.18, $Z = -1.03$, $p = 0.55$). Similarly, there was no particular difference between their ratings for Manner-prep and Manner-AdvClause constructions (OR = -0.18, $Z = -1.03$, $p = 0.55$).

The results for monolingual speakers of Portuguese ('L1 Portuguese') show a lower acceptability of the Manner-verb structure compared to both Manner-prep (OR = -1.73, $Z = -7.57$, $p < 0.01$) and Manner-AdvClause (OR = -0.84, $Z = -3.98$, $p < 0.01$). There was also a higher acceptability of Manner-prep structure in relation to the Manner-AdvClause (OR = -0.89, $Z = 3.91$, $p < 0.01$). For both groups, the regression showed no effects of the Participant variable in the result (L1 English: [$s^2 = 1.741$, $SD = 1.32$], L1 Portuguese: [$s^2 = 0.33$, $SD = 0.58$]).

Overall, L1 English speakers displayed higher acceptability of structures in which Manner is encoded in the verb (s-framing strategy) compared to the ones in which Manner is encoded adverbially (v-framing strategy). However, there was no significant difference in their assessment of Manner encoding in adverbial clauses in relation to the other two strategies. L1 Portuguese speakers

²As in the previous model, we included *Events* as a random variable, as suggested by one of the reviewers. No effects were found.

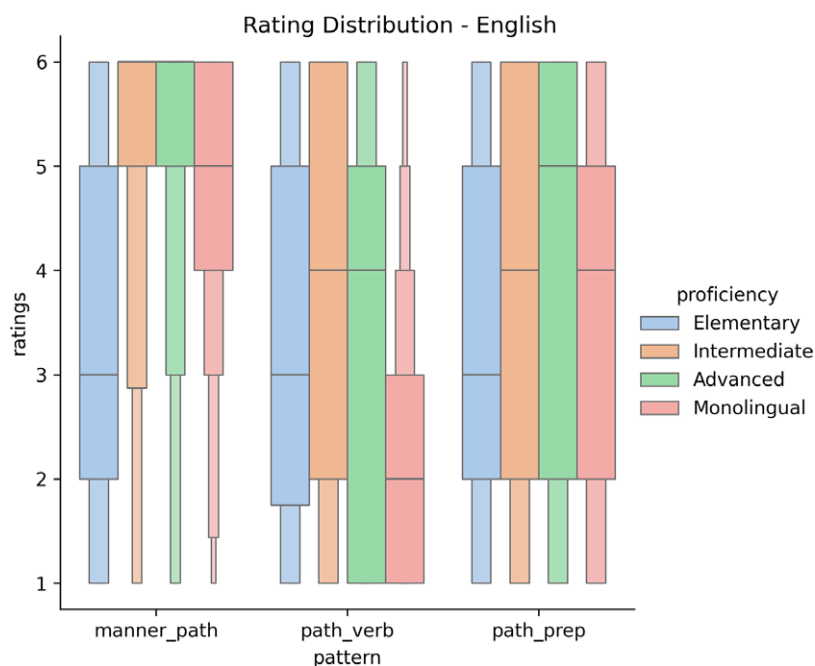


Figure 1. Sentence ratings by English speakers.

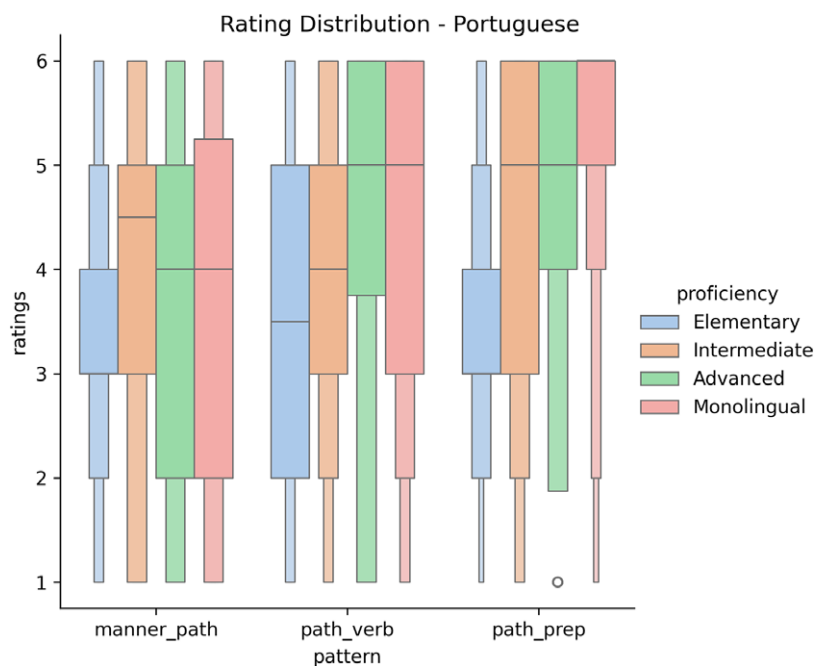


Figure 2. Sentence ratings by Portuguese speakers.

demonstrated a clear preference for the adverbial encodings of Manner (v-framing strategies) over the s-framing strategy and rated the encoding via adverbial clause lower than the prepositional counterpart.

Within-group comparisons: bilinguals

Acceptability ratings by L2 English speakers

The results of the ordinal logistic regression for L2 English Elementary speakers (Table 3) showed that they consider Manner-verb

constructions slightly more acceptable than Manner-preposition constructions ($OR = 0.46$, $Z = 2.58$, $p = 0.01$). Similar ratings are observed for Manner-AdvClause over Manner-verb ($OR = 0.63$, $Z = 3.92$, $p = 0.00$). They did not, however, show any particular preference for Manner-prep over Manner-AdvClause constructions in their L2 ($OR = 0.17$, $Z = 1.07$, $p = 0.53$).

For L2 English Intermediate speakers, Manner-verb structures received a significantly higher acceptability rate than both Manner-prep ($OR = 1.80$, $Z = 6.78$, $p < 0.01$) and Manner-AdvClause ($OR = 3.19$, $Z = 11.00$, $p < 0.01$). They also demonstrated higher

Table 3. Within-group comparison

Language	Proficiency	Contrast (Manner encoding)	Estimate	SE	Z-ratio	p-value
Portuguese	Elementary	Main.verb – Preposition	0.46	0.16	2.85	0.01
		Main.verb – Adv.clause	0.63	0.16	3.92	0.00
		Preposition – Adv.clause	0.17	0.16	1.07	0.53
	Intermediate	Main.verb – Preposition	1.80	0.26	6.78	<.01
		Main.verb – Adv.clause	3.19	0.29	11.00	<.01
		Preposition – Adv.clause	1.39	0.24	5.72	<.01
	Advanced	Main.verb – Preposition	0.21	0.17	1.20	0.45
		Main.verb – Adv.clause	−0.00	0.18	−0.04	0.99
		Preposition – Adv.clause	−0.22	0.18	−1.23	0.43
	Monolingual	Main.verb – Preposition	−0.44	0.17	−2.56	0.02
		Main.verb – Adv.clause	−0.18	0.17	−1.03	0.55
		Preposition – Adv.clause	0.26	0.17	1.50	0.28
English	Elementary	Main.verb – Preposition	1.15	0.14	7.90	<.01
		Main.verb – Adv.clause	1.31	0.14	8.91	<.01
		Preposition – Adv.clause	0.15	0.13	1.13	0.49
	Intermediate	Main.verb – Preposition	1.18	0.29	3.95	<.01
		Main.verb – Adv.clause	2.61	0.33	7.87	<.01
		Preposition – Adv.clause	1.43	0.29	4.79	<.01
	Advanced	Main.verb – Preposition	−1.05	0.47	−2.23	0.06
		Main.verb – Adv.clause	−0.43	0.46	−0.94	0.61
		Preposition – Adv.clause	0.62	0.45	1.35	0.36
	Monolingual	Main.verb – Preposition	−1.73	0.22	−7.57	<.01
		Main.verb – Adv.clause	−0.84	0.21	−3.98	<.01
		Preposition – Adv.clause	0.89	0.22	3.91	<.01

acceptability of Manner-prep over Manner-AdvClause (OR = 1.39, $Z = 5.27$, $p = <0.01$).

Acceptability ratings provided by L2 English advanced speakers, however, showed no particular preference for any of the structures, with Manner-verb and Manner-prep (OR = 0.21, $Z = 1.20$, $p = 0.45$), Manner-verb and Manner-AdvClause (OR = 0.00, $Z = -0.04$, $p = 0.99$) and Manner-prep and Manner-AdvClause (OR = −0.22, $Z = -1.23$, $p = 0.43$) yielding no significant difference. Participant variability had no effect for any of the groups (L2 English Elementary: [$s^2 = 1.16$, $SD = 1.07$], L2 English Intermediate: [$s^2 = 1.93$, $SD = 1.39$], L2 English Advanced: [$s^2 = 0.44$, $SD = 0.66$]).

Acceptability ratings by L2 BP speakers

The results of the ordinal logistic regression for L2 BP Elementary speakers (Table 3) showed a slightly lower acceptability of Manner-verb constructions over Manner-preposition constructions (OR = 1.15, $Z = 7.90$, $p = <0.01$). Higher ratings for Manner-verb over Manner-AdvClause (OR = 1.31, $Z = 8.91$, $p = <0.01$) are also observed. L2 BP Elementary speakers did not show, however, any significant difference in their ratings for Manner-prep and Manner-AdvClause constructions in Portuguese (OR = 0.15, $Z = 1.13$, $p = 0.49$).

L2 BP Intermediate speakers assigned slightly higher acceptability ratings to Manner-prep structures in comparison to Manner-AdvClause (OR = 1.43, $Z = 4.79$, $p = <0.01$) and Manner-verb

(OR = 1.18, $Z = 3.95$, $p = <0.01$). Manner-verb constructions received higher ratings than Manner-AdvClause sentences (OR = 2.61, $Z = 7.87$, $p = <0.01$).

Similarly to the ratings by L2 English Advanced speakers, those provided by L2 BP Advanced speakers showed no particular preference for any of the structures, with Manner-verb and Manner-prep (OR = −1.05, $Z = -2.23$, $p = 0.06$), Manner-verb and Manner-AdvClause (OR = −0.43, $Z = -0.94$, $p = 0.61$) and Manner-prep and Manner-AdvClause (OR = 0.62, $Z = 1.35$, $p = 0.36$) yielding no significant difference. Participant variability also had no effect for any of the groups (L2 English Elementary: [$s^2 = 0.53$, $SD = 0.73$], L2 English Intermediate: [$s^2 = 1.37$, $SD = 1.17$], L2 English Elementary: [$s^2 = 0.59$, $SD = 0.76$]).

Between-group comparisons

Manner-encoded-in-the-verb constructions

The distribution of L2 English speakers' ratings (Table 4) showed differences among the levels tested, in particular, in relation to the ratings by L1 English speakers. The ordinal logistic regression (Table 4) showed that Elementary and Advanced speakers assigned similar acceptability ratings to Manner-verb constructions as those provided by L1 speakers (L2 English Elementary: [OR = −0.34, $Z = -0.91$, $p = 0.79$]; L2 English Advanced: [OR = 0.04, $Z = 0.11$,

Table 4. Between-Group Comparison

Language	Pattern	Proficiency	Estimate	SE	Z-value	p-value
Portuguese	Manner verb + Path satellite	Monolingual	-	-	-	-
		Elementary	1.43	0.29	4.86	1.17e-6
		Intermediate	0.24	0.40	0.61	0.53
		Advanced	-0.33	0.54	-0.61	0.53
	Path verb + Manner clause	Monolingual	-	-	-	-
		Elementary	-1.00	0.33	-2.99	0.002
		Intermediate	-2.85	0.47	-5.95	2e-9
		Advanced	-1.10	0.63	-1.73	0.08
	Path verb + Manner prep	Monolingual	-	-	-	-
		Elementary	-1.95	0.30	-6.30	2.82e-10
		Intermediate	-2.76	0.41	-6.60	4.08e-11
		Advanced	-1.97	0.55	-3.55	0.00
English	Manner verb + Path satellite	Monolingual	-	-	-	-
		Elementary	0.13	0.16	0.80	0.42
		Intermediate	1.53	0.21	7.19	6.5e-13
		Advanced	-0.60	0.16	-0.41	0.67
	Path verb + Manner clause	Monolingual	-	-	-	-
		Elementary	-0.68	0.16	-4.08	4.42e-05
		Intermediate	-1.07	0.20	-5.37	7.76e-08
		Advanced	-0.21	0.16	-1.29	0.19
	Path verb + Manner prep	Monolingual	-	-	-	-
		Elementary	-0.73	0.17	-4.31	1.62e-05
		Intermediate	-0.21	0.19	-1.90	0.27
		Advanced	-0.56	0.16	-3.38	0.0007

$p = 0.99$)). There was also no significant difference between the ratings of Elementary and Advanced speakers ($OR = 0.39$, $Z = 1.01$, $p = 0.73$). Interestingly, there were differences between the ratings of Intermediate and L1 speakers ($OR = -1.88$, $Z = 3.94$, $p < 0.01$) as well as Intermediate and Advanced speakers ($OR = 1.93$, $Z = 4.00$, $p < 0.01$).

The ordinal logistic regression for the L2 BP speakers' data showed a difference in the acceptability ratings by L1 and Elementary speakers ($OR = -1.43$, $Z = -4.86$, $p < 0.01$) as well as Elementary and the other two L2 levels (L2 BP Intermediate: [$OR = 1.18$, $Z = 3.25$, $p < 0.01$]; L2 BP Advanced: [$OR = 1.77$, $Z = 3.39$, $p < 0.01$]). A positive effect of proficiency can be observed in the comparison between the ratings by Intermediate and Advanced learners in relation to L1 speakers. There was no statistical difference between the ratings of L2 BP Intermediate and L1 speakers ($OR = -0.24$, $Z = -0.61$, $p < 0.92$) or between the acceptability judgments by L2 Advanced and L1 speakers ($OR = 0.33$, $Z = 0.61$, $p < 0.92$). The effects of L2 proficiency were stable from the Intermediate level onwards, as there were no differences between the judgments by Intermediate and Advanced speakers either ($OR = 0.58$, $Z = 0.99$, $p < 0.75$).

As in the previous tests, participant variability had no effect for any of the groups (L2 English: [$s^2 = 1.67$, $SD = 1.29$], L2 BP: [$s^2 = 0.65$, $SD = 0.80$]). The significance of these findings will be addressed in the Discussion section.

Manner-in-the-prepositional-phrase constructions

For Manner-prep constructions, English speakers' ratings showed a positive effect of L2 proficiency. The ordinal logistic regression showed that there is a significant difference between the acceptability judgments of L1 and Elementary speakers ($OR = 0.94$, $Z = 2.71$, $p = 0.03$). There were, however, no differences between the ratings by L1 and Intermediate ($OR = 0.39$, $Z = 0.93$, $p = 0.78$) or L1 and Advanced speakers ($OR = 0.75$, $Z = 2.06$, $p = 0.16$). There were also no effects of proficiency between Intermediate and Advanced learners ($OR = 0.35$, $Z = 0.83$, $p = 0.84$).

L2 BP data showed differences in the acceptability ratings compared to those by L1 and all L2 speakers (Elementary: [$OR = 1.95$, $Z = 6.30$, $p < 0.01$]; Intermediate: [$OR = 2.76$, $Z = 6.60$, $p < 0.01$]; Advanced: [$OR = 1.97$, $Z = 3.55$, $p < 0.01$]). It also showed no differences in the judgments from Elementary to other L2 proficiency levels (Intermediate: [$OR = 0.81$, $Z = 2.28$, $p = 0.10$]; Advanced: [$OR = 0.02$, $Z = 0.05$, $p = 0.99$]) or between Intermediate and Advanced ($OR = -0.78$, $Z = -1.34$, $p = 0.53$). Participant variability had no effect for any of the L2 groups (L2 English: [$s^2 = 1.33$, $SD = 1.15$], L2 BP: [$s^2 = 0.65$, $SD = 0.80$]).

Manner-in-the-adverbial-clause constructions

For L2 English speakers' ratings, the ordinal logistic regression showed that there were significant differences in the acceptability judgments by Elementary and Intermediate speakers compared to

L1 speakers (Elementary: [OR = 0.95, $Z = 2.70$, $p = 0.03$]; Intermediate: [OR = 1.48, $Z = 3.40$, $p < 0.01$]). There were no differences between the ratings by Elementary and Intermediate speakers (OR = 0.53, $Z = 1.25$, $p = 0.59$). Higher L2 proficiency had an effect as Advanced and L1 speakers' ratings converged (OR = 0.30, $Z = 0.81$, $p = 0.84$). This observation was also supported by the fact that there was a significant difference in the judgments by Intermediate and Advanced speakers (OR = -1.18 , $Z = -2.70$, $p < 0.03$).

L2 BP speakers' data showed a similar effect of high proficiency in participants' results. There were differences in the ratings provided by Elementary and Intermediate speakers compared to those by L1 speakers (Elementary: [OR = 1.00, $Z = 2.99$, $p = 0.01$]; Intermediate: [OR = 2.85, $Z = 5.95$, $p < 0.01$]). There were, however, differences between the ratings by Elementary and Intermediate speakers (OR = 1.84, $Z = 4.37$, $p < 0.01$). Highly proficient learners showed no difference in their ratings in relation to L1 speakers (OR = 1.10, $Z = 1.73$, $p = 0.30$). This effect was corroborated by the difference in acceptability judgment by their less proficient counterparts (Intermediate: OR = -1.75 , $Z = 2.56$, $p = 0.05$).

While participant variability had no effect for the English group ($s^2 = 1.38$, $SD = 1.17$), our model showed that it played a role in the results by the BP group ($s^2 = 0.98$, $SD = 0.99$). This fact, which helps clarify some points in the data, will be addressed in the following section.

Discussion

The present study set out to answer three research questions. First, considering that both English and Brazilian Portuguese allow descriptions of motion with a Path verb and an optional Manner prepositional phrase, we investigated where the Manner-prep structure ranks compared to the canonical structures of each language. Second, our goal was to identify at what stage of L2 development (i.e., the effects of L2 proficiency) L2 Portuguese learners' acceptability ratings of Manner-verb, Manner-AdvClause and Manner-prep converged with those of L1 Portuguese speakers. Third, our aim was to identify at what stage of L2 development (i.e., the effects of L2 proficiency) L2 English learners' acceptability ratings of the same three structures approximated to those of L1 English speakers. In this section, we attempt to answer these three questions based on the results we encountered.

The elaboration of Manner in a prepositional phrase

L1 English speakers demonstrate a higher acceptability for Manner-verb constructions over Manner-Path, which is expected considering the theoretical work in the field of semantic typology (Slobin, 2004; Talmy, 1985, 1991, 2000). Manner-verb is, after all, the most frequent encoding strategy employed by s-framed languages, which is the canonical classification English has received in the literature (Lewandowski, 2021; Pavlenko & Volynsky, 2015; Slobin, 2004). Interestingly, however, L1 English did not rate Manner-prep and Manner-AdvClause significantly differently from one another. As previously discussed, the acceptability of Manner-AdvClause constructions comes to no surprise as English allows encoding of the type (*Mary crossed the river swimming*). What is remarkable in participants' judgments is that the structure received ratings that are not distinguishable from those for Manner-prep constructions. According to Slobin (2004), what plays a more important role is when Manner *must* be encoded as opposed to when it *may* be encoded. In this study, our focus was on instances in which Manner

was encoded, and no Mannerless clauses were provided. Due to the Manner elaboration being presented and licit, speakers may assume that it has done so for a reason (i.e., a pragmatic motivation). The investigation of speakers' assumptions about the obligatoriness of Manner encoding is outside the scope of this work.

As for the question at hand, although L1 English speakers rated Manner-prep structures as less acceptable than the canonical Manner-verb one, they did not show any particular higher or lower judgment of the structure compared to Manner-AdvClause – the other canonical v-framed structure. This seems to indicate that, while the Manner-verb structure is preferred, the v-framed constructions may be used in lieu of one another.

For L1 BP speakers, ratings show acceptability that matches what one would expect based on the literature. Manner-prep and Manner-AdvClause constructions – the prototypical v-framing structures – were preferred over Manner-verb – the common s-framing strategy. What the data contributes to is showing how Manner-prep fares in relation to Manner-AdvClause. Judgments show that speakers consider Manner-prep encoding more acceptable than Manner-AdvClause, which is most often used as an example of canonical v-framing strategy. Overall, this corroborates the argument that BP favors verb-framing (Batoréo & Ferrari, 2016; Nogueira, 2009).

L2 English development

For Manner-verb structures, Advanced L2 English speakers' acceptability ratings converged with those by L1 English speakers – showing that high proficiency does have an effect on their judgments of the canonical s-framed structure. For Manner-prep constructions, the effects were observed earlier, at the Intermediate level. This convergence also holds for L2 Advanced learners. Therefore, bilingual learners' judgments were sensitive to lower acceptability of Manner encoding in the preposition at a somewhat early stage of L2 development. For Manner-AdvClause constructions, the convergence occurred somewhat later, at the Advanced level – as observed for Manner-verb structures. Speakers at the Intermediate level showed significant difference from L1ers, a difference also observed between this group and Advanced learners. These results are particularly promising if we consider the sample size ($n = 73$) and the years learning the L2 ($M = 13.07$, $SD = 11.22$).

The data partially supported our hypothesis that learners at earlier stages would demonstrate higher acceptability ratings for sentences that reflect the thinking-for-speaking strategies of their L1. For instance, L2 English speakers demonstrated late convergence for Manner-verb and Manner-AdvClause constructions, despite the high ratings they received from L1 English speakers. It did not, however, support our prediction for Manner-prep, where convergence began to occur significantly early, at the Intermediate level. The fact that Manner-AdvClause sentences received different ratings from Elementary and Intermediate learners when compared to L1 speakers also supports our hypothesis that there would be differences in their ratings despite the grammaticality of the structure in English. The data also partially supported our prediction that late bilingual speakers will begin to assign higher ratings to sentences that show the pattern of thinking-for-speaking of their L2 as early as at the Intermediate level. While this was true for Manner-prep constructions, the other two types – which were equally considered significantly acceptable by L1 speakers, received lower ratings by L2ers, which suggests transfer from the L1. Our final hypothesis, that the large verb lexicon of the S-framed languages would indicate an early

convergence of ratings was not entirely supported. L2 Elementary and Intermediate learners did not rate Manner-path constructions as highly as their L1 counterparts.

L2 Portuguese development

For Manner-verb structures, Intermediate proficiency had an effect on the judgments of L2 Portuguese learners, who provided acceptability ratings that converge with those of L1 Portuguese speakers. As the results suggested that this convergence also holds for L2 Advanced learners, it seems that as bilingual learners of BP (a v-framed language) judgments are sensitive to lower acceptability of Manner encoding in the main verb in BP at a somewhat early stage of L2 development. For Manner-prep constructions, the data showed no converge of acceptability among any of the L2 groups and the L1 participants. This indicates a locus of difficulty in the development of judgment of learners of a v-framed language, which is of particular relevance since this is the most acceptable structure according to L1 BP speakers. It is important to note, however, that the Advanced sample is very small ($n = 4$), so these results cannot be generalized. For Manner-AdvClause constructions, the converge also does not occur during Elementary or Intermediate stages. We do observe, however, that Advanced speakers assign acceptability ratings that match those of L1 BP speakers. This result is particularly interesting because L1 English participants rated Manner-AdvClause constructions as acceptable as Manner-verb (the canonical s-framed structure). It would not be a far-reaching assumption that if L1 transfer were to occur, L2 speakers' ratings would converge earlier. Naturally, there are several other variables that need to be accounted for in this case, so this issue remains for future studies.

These results support our prediction that L2 BP learners would consider Manner-prep less acceptable than L1 speakers regardless of it being the highest-rated structure by L1 speakers. This indicates L1 influence, especially since this structure was rated significantly low by L1 English speakers in their language. Similarly, the data confirmed our hypothesis that L2 BP speakers would consider v-framed structures less acceptable in general despite the low frequency of Manner verbs in BP. Learners were able to determine that Manner-verb constructions are disfavored compared to other types, but not that v-framed structures are preferred overall. Our assumption that late bilingual speakers would begin to assign higher ratings to sentences that show the pattern of thinking-for-speaking of their L2 as early as at the Intermediate level was only met for Manner-prep constructions. However, as we argued for the Manner-prep constructions, we note the limited Advanced sample size: the convergence might not properly illustrate the acquisition of a v-framed L2. Another point that needs to be acknowledged for the acquisition of BP as a second language is that BP learners, despite their L2 proficiency, have a significantly lower average of years spent learning a language compared to those learning English (see Table 2). Considering that the teaching of motion encoding is often not emphasized in the language classroom (Mengali, 2020), overall exposure may have a significant effect as it supports the development of vocabulary and structure (Gass & Mackey, 2006).

Conclusion

This study investigated at what stages of L2 proficiency (Elementary, Intermediate and Advanced) L2 Brazilian Portuguese and L2 English learners' acceptability ratings converge with those of L1 speakers when rating sentences using the three structures. Our

goal was to fill gaps in the literature by (i) providing more data on the acquisition of v-framed L2s – in this case, Brazilian Portuguese – by speakers of s-framed languages; (ii) introducing results collected from a significantly large sample of beginners and intermediate L2 speakers – both acquiring an s-framed and a v-framed language; and by (iii) offering more data on the acquisition of L2 English by L1 Portuguese speakers – in contrast with most studies that have hardly superseded an average of 15 participants.

Our first question was how Manner-prep structures rank compared to the canonical lexicalization pattern of English (Manner-verb) and Portuguese (Manner-AdvClause). We found that L1 English speakers rated Manner-prep as less acceptable than the prototypical Manner-verb structure, but did not show any particular preference for the structure compared to Manner-AdvClause. Therefore, it seems that, while the Manner-verb structure is preferred, the v-framed constructions may be used in lieu of one another. L1 BP speakers rated Manner-prep and Manner-AdvClause – the prototypical v-framing structures – higher than Manner-verb constructions. These findings corroborate the argument that BP favors verb-framing (Batoréo & Ferrari, 2016; Nogueira, 2009).

Our second question was at what stage of L2 development L2 Portuguese learners' acceptability ratings of Manner-verb, Manner-AdvClause and Manner-prep converged with those of L1 Portuguese speakers. The assumption that late bilinguals would assign higher ratings to sentences that show the pattern of thinking-for-speaking of their L2 starting at the Intermediate level was only met for Manner-prep structures. Overall, while learners were able to determine that Manner-verb constructions are less preferred compared to other types in their L2, they did not rate v-framed structures as highly as L1 speakers did. Although this suggests crosslinguistic influence from the L1 (English speakers rated s-framed higher than v-framed structure), it also shows signs of a change toward the structure of their L2 (in many cases, Manner-verb constructions are not grammatical in Portuguese).

Our third and final question was at what stage of L2 development L2 English learners' acceptability ratings of the same three structures approximated to those of L1 English speakers. Results showed that, at least to an extent, learners at earlier stages demonstrated higher acceptability ratings for sentences that reflect the thinking-for-speaking strategies of their L1. L2 English speakers demonstrated late convergence for Manner-verb and Manner-AdvClause constructions, despite the high ratings they received from L1 English speakers. However, the convergence occurred as early as at the Intermediate level for Manner-prep structures. For both languages, results show that convergence starts at the Intermediate level for Manner-prep constructions and at a later proficiency stage for Manner-verb and Manner-AdvClause structures.

Moving forward, we recommend carrying out work with a larger sample size of advanced learners to provide a clearer picture of v-framed structure acquisition at higher proficiency levels. We also recommend more work on the acquisition of the English-Portuguese pair with a focus on production tasks, which is currently an understudied area. More specifically, as proposed by Lewandowski (2021), we suggest more elicitation of motion encoding in speech and writing via video clips that display nonstatic portrayals of motion.

Data availability statement. Data availability: The data that support the findings of this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.10822888>.

Acknowledgments. The authors would like to thank Dr. Melissa de Sá from the Federal Institute of Minas Gerais for assisting in the data collection process.

Competing interest. The authors declare none.

Ethical standard. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

- Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition*, 102(1), 16–48. <https://doi.org/10.1016/j.cognition.2005.12.006>
- Alloway, T. P., & Corley, M. (2004). Speak before you think: The role of language in verb concepts. *Journal of Cognition and Culture*, 4(2), 319–345. <https://doi.org/10.1163/1568537041725105>
- Almeida, M. (2002). Verbs of motion: The implications of cognitive semantics in teaching grammar. *Hispania*, 85(3), 609–617. <https://doi.org/10.2307/4141150>
- Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, 12(1), 83–95. <https://doi.org/10.1017/S136672890800388X>
- Athanasopoulos, P., Bylund, E., Montero-Melis, G., Damjanovic, L., Schartner, A., Kibbe, A., & Thierry, G. (2015). Two languages, two minds: Flexible cognitive processing driven by language of operation. *Psychological Science*, 26(4), 518–526. <https://doi.org/10.1177/0956797614567509>
- Batoréo, H. (2014). Perspective point (viewpointing) and events of motion in European Portuguese. *International Journal of Cognitive Linguistics*, 5(1), 55–74.
- Batoréo, H. J., & Ferrari, L. (2016). Events of motion and Talmyan typology: Verb-framed and satellite-framed patterns in Portuguese. *Cognitive Semantics*, 2(1), 59–79. <https://doi.org/10.1163/23526416-00201004>
- Beavers, J., Levin, B., & Tham, S. W. (2010). The typology of motion expressions revisited. *Journal of linguistics*, 46(2), 331–377. <https://doi.org/10.1017/S002226709990272>
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1–L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, 30(2), 225–251. <https://doi.org/10.1017/S0272263108080327>
- Brown, A., & Gullberg, M. (2010). Changes in encoding of path of motion in a first language during acquisition of a second language. *Cognitive Linguistics*, 21(2), 263–286. <https://doi.org/10.1515/COGL.2010.010>
- Brown, A., & Gullberg, M. (2011). Bidirectional cross-linguistic influence in event conceptualization? Expressions of path among Japanese learners of English. *Bilingualism: Language and Cognition*, 14(1), 79–94. <https://doi.org/10.1017/S1366728910000064>
- Brown, A., & Gullberg, M. (2013). L1–L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition*, 16(3), 477–494. <https://doi.org/10.1017/S1366728912000491>
- Bylund, E., & Athanasopoulos, P. (2014). Linguistic relativity in SLA: Toward a new research program. *Language Learning*, 64(4), 952–985. <https://doi.org/10.1111/lang.12080>
- Cadierno, T. (2004). Expressing motion events in a second language: A cognitive typological perspective. In M. Achard, & S. Neimeier (Eds.), *Cognitive linguistics, second language acquisition and foreign language pedagogy* (pp. 13–49). Mouton de Gruyter.
- Cadierno, T. (2008). Learning to talk about motion in a foreign language. In P. Robinson, & N. C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 249–285). Routledge.
- Cadierno, T. (2017). Thinking-for-speaking about motion in a second language. In I. Ibarretxe-Antunano (Ed.), *Motion and Space across Languages: Theory and Applications* (pp. 279–300). John Benjamins Publishing Company.
- Cadierno, T., & Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annual Review of Cognitive Linguistics*, 4(1), 183–216. <https://doi.org/10.1075/arcl.4.08cad>
- Chen, L., & Guo, J. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of Pragmatics*, 41(9), 1749–1766. <https://doi.org/10.1016/j.pragma.2008.10.015>
- Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, 41(1–3), 83–121. [https://doi.org/10.1016/0010-0277\(91\)90033-Z](https://doi.org/10.1016/0010-0277(91)90033-Z)
- Chomeya, R. (2010). Quality of psychology test between Likert scale 5 and 6 points. *Journal of Social Sciences*, 6(3), 399–403. <https://doi.org/10.3844/jssp.2010.399.403>
- Chomsky, N. (1957). *Syntactic Structures*. De Gruyter Mouton.
- Danhier, R. D., & Mertins, B. (2016). Language-specific information structure in German and Spanish route directions. *Vigo International Journal of Applied Linguistics*, 13, 55–92.
- Elliott, E., & Yountchi, L. (2009). Total physical response and Russian multi- and unidirectional verbs of motion: A case study in acquisition. *The Slavic and East European Journal*, 53(3), 428–450.
- Ellis, N. C., & Robinson, P. (2008). An introduction to cognitive linguistics, second language acquisition, and language instruction. In P. Robinson, & N. C. Ellis (eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition* (pp. 13–34). Routledge.
- Filipović, L. (2011). Speaking and remembering in one or two languages: Bilingual vs. monolingual lexicalization and memory for motion events. *International Journal of Bilingualism*, 15(4), 466–485. <https://doi.org/10.1177/136700691140306>
- Gagarina, N. (2009). Verbs of motion in Russian: An acquisitional perspective. *The Slavic and East European Journal*, 50(3), 451–470.
- Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, 19(1), 3–17. <https://doi.org/10.1075/aila.19.03gas>
- Grey, S., & Tagarelli, K. M. (2018). Psycholinguistic methods. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The Palgrave Handbook of Applied Linguistics Research Methodology* (pp. 287–312). Palgrave Macmillan.
- Grigoroglou, M., & Ganea, P. A. (2022). Language as a mechanism for reasoning about possibilities. *Philosophical Transactions of the Royal Society B*. <https://doi.org/10.1098/rstb.2021.0334>
- Hasko, V. (2009). The locus of difficulties in the acquisition of Russian verbs of motion by highly proficient learners. *The Slavic and East European Journal* 50(3), 360–385.
- Hwang, H. (2023). Wanna contraction in first language acquisition, child second language acquisition, and adult second language acquisition. *Bilingualism: Language and Cognition*, 27(3), 322–333. <https://doi.org/10.1017/S1366728923000640>
- Ibarretxe-Antuñano, I., Cadierno, T., & Hijazo-Gascón, A. (2016). The role of force dynamics and intentionality in the reconstruction of L2 verb meanings: A Danish-Spanish bidirectional study. *Review of Cognitive Linguistics. Published under the auspices of the Spanish Cognitive Linguistics Association*, 14(1), 136–160. <https://doi.org/10.1075/rcl.14.1.06iba>
- Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual Review of Applied Linguistics*, 15, 125–150. <https://doi.org/10.1017/S0267190500002658>
- Kissling, E. M. (2018). Pronunciation instruction can improve L2 learners' bottom-up processing for listening. *The Modern Language Journal*, 102(4), 653–675. <https://doi.org/10.1111/modl.12512>
- Lewandowski, W. (2021). Variable motion event encoding within languages and language types: a usage-based perspective. *Language and Cognition*, 13(1), 34–65. <https://doi.org/10.1017/langcog.2020.25>
- Lewis, T. N. (2012). The effect of context on the L2 thinking for speaking development of path gestures. *L2 Journal*, 4(2). <https://doi.org/10.5070/L24211612>
- Lucy, J. A. (1992). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge University Press.
- Lucy, J. A. (1996). *Grammatical Categories and Cognition: A Case Study of the Linguistic Relativity Hypothesis*. Cambridge University Press.
- Lucy, J. A. (1997). Linguistic relativity. *Annual review of Anthropology*, 26(1), 291–312. <https://doi.org/10.1146/annurev.anthro.26.1.291>
- Meirelles, L. L. (2019). Verbos de movimento do português brasileiro: evidências contra uma tipologia binária/Brazilian Portuguese motion verbs: evidence against a two-way typology. *Revista de Estudos da Linguagem*, 27(2), 1101–1124. <https://doi.org/10.17851/2237-2083.27.2.1101-1124>

- Mengali, R. (2020). (Re)pensando o estilo retórico para expressar o movimento: Noções aplicáveis ao ensino do português como língua estrangeira. *Cadernos de Pós-Graduação em Letras*, 20(3), 64–75.
- Nogueira, A. R. (2009). Processos de lexicalização na interlíngua de alunos aprendizes de inglês: uma análise dos verbos de movimento. [Master's thesis]. Fortaleza, Ceará. Available at <https://repositorio.ufc.br/handle/riufc/2821>
- Chiappara, J. P., Siqueira, J. S., Oliveira, A. A., & Gediel, A. L. (2022). *Estudos de linguística, ensino, e literatura em múltiplas perspectivas*. Divisão Gráfica Universitária.
- Ozyurek, A. (2002). Speech-gesture relationship across languages and in second language learners: Implications for spatial thinking and speaking. *BUCLD 26 Proceedings*, 500–509.
- Pavlenko, A., & Driagina, V. (2006). *Advanced-level narrative skills in Russian: A workbook for students and teachers*. Unpublished manuscript, Center for Advanced Language Proficiency Education and Research, State College, PA.
- Pavlenko, A., & Malt, B. C. (2011). Kitchen Russian: Cross-linguistic differences and first-language object naming by Russian–English bilinguals. *Bilingualism: Language and Cognition*, 14(1), 19–45. <https://doi.org/10.1017/S136672891000026X>
- Pavlenko, A., & Volynsky, M. (2015). Motion encoding in Russian and English: Moving beyond Talmy's typology. *The Modern Language Journal*, 99(S1), 32–48. <https://doi.org/10.1111/modl.12177>
- Putnam, M. T. (2019). The (in)stability of grammars. *Studies in Second Language Acquisition*, 41(2), 275–278. <https://doi.org/10.1017/S0272263119000299>
- R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. <https://www.R-project.org/>
- Ragnarsdóttir, H., & Strömqvist, S. (2004). Time, space, and manner in Swedish and Icelandic: Narrative construction in two closely related languages. In H. Ragnarsdóttir & S. Strömqvist (Eds.), *Relating events in narrative, vol. 2* (pp. 113–141). Psychology Press.
- Selinker, L., & Gass, S. M. (2008). *Second Language Acquisition*. Lawrence Erlbaum Ass.
- Slobin, D. I. (1996). From “thought and language” to “thinking for speaking”. In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge University Press.
- Slobin, D. I. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Neimeier, & R. Dirven (Eds.), *Evidence for Linguistic Relativity* (pp. 107–138). John Benjamins Publishing Company. <https://doi.org/10.1075/cilt.198.10slo>
- Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In H. Ragnarsdóttir & S. Strömqvist (Eds.), *Relating events in narrative, vol. 2* (pp. 219–257). Psychology Press.
- Stam, G. (1998). Changes in patterns of thinking about motion with L2 acquisition. *Faculty Publications*, 34.
- Stam, G. (2006). Thinking-for-speaking about motion: L1 and L2 speech and gesture. *International Review of Applied Linguistics in Language Teaching*, 44(2), 145–171. <https://doi.org/10.1515/IRAL.2006.006>
- Stam, G. (2010). Chapter 3: Can an L2 Speaker's Patterns of Thinking for Speaking Change?. In Z. Han & T. Cadierno (Eds.), *Linguistic Relativity in SLA: Thinking for Speaking* (pp. 59–83). Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781847692788-005>
- Stam, G. (2015). Changes in thinking for speaking: A longitudinal case study. *The Modern Language Journal*, 99(S1), 83–99. <https://doi.org/10.1111/j.1540-4781.2015.12180.x>
- Stam, G., & McCafferty, S. G. (2009). *Gesture: Second Language Acquisition and Classroom Research*. Routledge. <https://doi.org/10.4324/9780203866993>
- Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language Typology and Syntactic Description* (pp. 36–149). Cambridge University Press.
- Talmy, L. (1991). Path to realization: A typology of event conflation. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society*, 17, 480–519. <https://doi.org/10.3765/bls.v17i0.1620>
- Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems* (Vol. 1). MIT press.
- Talmy, L. (2012). Main verb properties. *International Journal of Cognitive Linguistics*, 3(1), 1–23.
- Tare, M., Golonka, E., Lancaster, A. K., Bonilla, C., Doughty, C. J., Belnap, R. K., & Jackson, S. R. (2018). The role of cognitive aptitudes in a study abroad language-learning environment. In C. Sanz & A. Morales-Front (Eds.), *The Routledge Handbook of Study Abroad Research and Practice* (pp. 406–420). Routledge.
- Tigheelaar, M., Bowles, R. P., Winke, P., & Gass, S. (2017). Assessing the validity of ACTFL can-do statements for spoken proficiency: A Rasch analysis. *Foreign Language Annals*, 50(3), 584–600. <https://doi.org/10.1111/flan.12286>
- Verhoeven, L., & Strömqvist, S. (2004). *Relating events in narrative, volume 2: Typological and contextual perspectives*. Taylor & Francis.
- Veríssimo, J. (2021). Analysis of rating scales: A pervasive problem in bilingualism research and a solution with Bayesian ordinal models. *Bilingualism: Language and Cognition*, 24(5), 842–848. <https://doi.org/10.1017/S1366728921000316>
- Winke, P., Zhang, X., & Pierce, S. J. (2023). A closer look at a marginalized test method: Self-assessment as a measure of speaking proficiency. *Studies in Second Language Acquisition*, 45(2), 416–441. <https://doi.org/10.1017/S0272263122000079>