A Cross-linguistic Study of Motion Event Acquisition in Japanese and Chinese: Through the Lens of Figure and Ground Components

ABSTRACT

Numerous cross-linguistic studies on motion events have been undertaken to test the thinking-for-speaking hypothesis, with a remarkable lead over the Path and Manner components. Yet, unexpectedly, the core entities in motion events such as Figure and Ground have garnered relatively limited focus. The current study shifts the focus to examine the way how subjects encode the Figure and Ground information through a narrative task, involving Chinese learners who learn Japanese as their foreign language, Chinese native speakers, and Japanese native speakers. The results revealed that the typological framework is too narrow to explain observed linguistic variations in the acquisition process. The striking differences followed from syntactic structure variations in the utterances related to how Japanese versus Chinese encode the Figure and Ground information, which suggest the joint impact of typological and syntactic factors on language acquisition. In addition, it was clearly shown that the JFL learners patterned to a large degree with the Japanese native speakers rather than the Chinese native speakers concerning the distribution of Figure and Ground information by motion event type. This means that to some extent the JFL learners have acquired the target-like way to report Figure and Ground information in the narrative task, showing a limited role of the L1 thinking for speaking patterns.

Keywords: motion events; Figure; Ground; typological; written narrative

INTRODUCTION

As a fundamental part of everyday human experiences, motion events have received increasing attention among researchers in the field of second language acquisition (Talmy, 1985; Slobin, 1996; Park et al., 2022; Wu et al., 2022). In Talmy's (1985) typological analysis, a motion event comprises four essential components: figure (the moving entity), ground (the reference point), motion (the presence of movement), and path (the trajectory that Figure follows with respect to Ground). To these, supplementary components (manner and cause) can be incorporated, as shown in the following example (adapted from Talmy, 1985):

(1) He swam across the pool.
Figure Motion+Manner Path Ground

Within the motion domain, numerous cross-linguistic research based on Talmy's (1985, 2000) binary typology of verb-framed (hence V-) and satellite-framed (hence S-) languages has provided the foundations in this field. Many of the empirical studies have been informed by Slobin's (1991) "thinking for speaking hypothesis" to testify the role of the learners' native languages (L1) in their acquisition of the second language (L2). While most of the findings support this hypothesis (Cadierno, 2004; Hasko, 2010; Sung & Kim, 2020), there are still a few counterexamples (Cadierno & Ruiz, 2006).

Recently, it has been shown that Path and Manner components have been underscored as typologically attractive and variable in this line of inquiry (Slobin, 2004; Akita, 2017; Matsumoto & Kawachi, 2020; Wang & Wei, 2023). Yet, unexpectedly, the core entities in motion events such as Figure and Ground (adapted from Gestalt Psychology) have garnered relatively limited focus in cross-linguistic research (Tajima & Duffield, 2012; Choi et al., 2018). This study seeks to delve into a detailed analysis of the components of Figure and Ground, which are a set of spatial concepts in linguistic expressions for motion events (Talmy, 1985;

Langacker, 1987). Grounded in Talmy's (2000) Figure-Ground theory, this study adopts a picture description task as the primary methodological approach, to compare the narratives of Chinese learners who learn Japanese as a foreign language (JFL learners) with native speakers of Japanese and native speakers of Chinese, regarding the presence and word order of Figure versus Ground information. Our purpose is to do clause-rank analysis to gain insights into the different types of Figure and Ground information encoded in a clause. Subsequently, a detailed linguistic analysis will be conducted to 1) identify similarities and differences in the depiction of Figure and Ground components across the three groups; and 2) examine the role of L1 transfer in shaping the descriptions of Figure and Ground information provided by JFL learners.

Given that Japanese and Chinese are two typologically overlapping languages to some extent (Chen & Guo, 2009; Matsumoto & Kawachi, 2020), their narratives offer a unique lens through which to explore the influence of L1 transfer on the acquisition of motion events in a second language context. Therefore, this study not only enhances our understanding of language typology and cognitive processes, but also informs pedagogical practices aimed at facilitating second language acquisition.

LITERATURE REVIEW

A central focus concerning motion events has been related to typological insights. Numerous studies have been carried out to test the thinking-for-speaking hypothesis based on Talmy's (1985, 2000) typological framework. However, given the inter- and intra-typological differences across languages, some studies indicate that the relevant components in motion events can be weighed differently (Li, 2017; Matsumoto, 2018), thus no exception for the Figure and Ground components. In this section we briefly discuss this typological line of research, with a particular emphasis on how Japanese and Chinese encode Figure and Ground information, and then set out to address the questions in the present study.

THE TYPOLOGICAL DIFFERENCES BETWEEN JAPANESE AND CHINESE

According to Talmy's classification (1985), Japanese belongs to the category of V-languages. However, there is currently no consensus regarding the classification of the Chinese language type. Most researchers believed that Chinese is an atypical S-language, especially regarding the serial verb constructions (Talmy, 1985, 2000; Shen, 2003). Talmy (1985) further proposed that there are three main typological patterns for the expression of motion events: 1) Motion + Manner / Cause, a language of this type can express motion occurring in various manners or by various causes, e.g., "The rock **sIid/rolled/bounced** down the hill", in this sentence the manners of the "rock" down the "hill" are indicated by the verb roots. While in another sentence, "I **pushed/threw/kicked** the keg into the storeroom", the cause notions are conflated in the verbs; 2) Motion + Path, as exemplified in the Spanish sentence "La botella **entró** a la cueva", the verb root "entró" expresses both the fact of Motion and the Path. Japanese is also an example of this type; 3) Manner + Figure, this conflation type, such as Atsugewi, expresses the verb roots of Motion with conflated Figure. This study refers to motion events that incorporate the components of cause, manner, and path as cause motion events, manner motion events, and path motion events, respectively.

MOTION EVENTS IN SECOND LANGUAGE ACQUISITION

Based on Talmy's (1985, 2000) typological framework, a line of research has been conducted to investigate the expressions of motion events by second language learners with different language backgrounds (e.g., Slobin, 2004; Sung & Kim, 2020; Wang & Wei, 2023). Generally, these studies carried out written or spoken narrative tasks, sometimes alongside native speakers performing the same task in their native language. Then

a comparison was conducted between participants with different language backgrounds to test Slobin's (1991) thinking-for-speaking hypothesis. Accordingly, while the L1 and L2 pertain to the same typological patterns, the L1 typically facilitates the acquisition of motion events in the L2 (Hasko, 2010; Gu & Lin, 2017). Conversely, while the L1 and L2 adapt to different typological patterns, the L1 exerts a negative influence on the L2 acquisition of motion events (Sung & Kim, 2020).

However, it is worth noting that the findings of the aforementioned studies are not consistently aligned, with most studies supporting the thinking-for-speaking hypothesis while a few present challenges to it (Cadierno & Ruiz, 2006; Schmiedtová, 2013). There are some factors contributing to the inconsistency, and the different components that being studied should also be an important factor. Yet, there have been too many studies focusing on the construal related to the direction of motion or manner of motion, the core entities of a motion event, that is Figure and Ground, have largely been ignored (Tajima & Duffield, 2012; Choi et al., 2018).

HOW JAPANESE AND CHINESE ENCODE FIGURE AND GROUND RELATION

According to Talmy (1978, p. 627), Figure is "a moving or conceptually movable point whose path or site is conceived as a variable", while Ground is "a reference point, having a stationary setting within a reference-frame". Thus, Figure holds perceptual salience due to its ease of mobility, whereas Ground, in contrast, is comparatively less salient (Talmy, 1985). Furthermore, Talmy (2000: 334) proposed the precedence principle, emphasizing that "the Figure has syntactic precedence over the Ground". Some studies have found that Talmy's canonical rules are uniform across languages (Muehleisen & Imai, 1997; Li, 2011). That's to say, almost all the languages align Figure and Ground in a typical way. But things are not that systematic: many cross-linguistic analyses reveal significant variations either in inter- or intra-language analysis, since the relevant components can be weighted differently (Li, 2011). In the following section, the encoding of Figure and Ground in Japanese and Chinese will be discussed systematically.

Firstly, the variations in clausal word order may be an important factor that affects the encoding of the two core entities. According to Tajima & Duffield (2012), languages are divided into two types: head-initial languages (such as English) and head-final languages (such as Japanese and Korean), due to the position of head element in a sentence. As a head-final language, the typical word order in Japanese is subject-object-verb (SOV): the verb succeeds its object; tense and mood affixes are both manifested as verbal suffixes; postpositions are used to indicate spatial relations (as shown in example 2).

(2) ジョンが 花瓶を 割った。
S O V
John-ga kabin-o wat-ta.
'John broke the vase.' (adapted from Tajima & Duffield, 2012)

However, as for Mandarin Chinese, it is hard to simply classify it into head-initial or head-final. Basically, Chinese is considered head-initial concerning its subject-verb-object (SVO) word order, the precedence of temporal and modal auxiliaries before the verb phrase. However, when comes to the lexical noun phrases (see example 4), it is head-final.

(3) Zhangsan meiyou kanjian Lisi.
S V O
'Zhangsan did not see Lisi.'

(4) {ni zui xihuan de} nei-ben shu mai-wan le.

S V

'The book that you like most has been sold out.' (cited from Tajima & Duffield, 2012)

Secondly, the variations in the grammatical status of clausal subjects between Japanese and Chinese seem to be another important factor leading to different encoding for Figure and Ground relations (Ikegami, 1981; Tajima & Duffield, 2012). Broadly, it is nearly universal to put clausal subjects initially in the sentences, the same as for Chinese. However, according to Ikegami (1981: 200), Japanese sentences typically begin with a topic argument (abbreviated as T), as opposed to subjects. That is to say, Japanese subjects marked by the particle "ga" may not invariably appear at the beginning of clauses (as shown in example 5).

(5) 日本は、山が多い。
T S
Nihon-wa yama-ga oo-i.
'There are many mountains in Japan.'

(6) 家のそば（G）に、バイク（F）が ある。
uti-no soba-ni baiku-ga aru.
'The bike (F) is near the house (G).' (adapted from Ikegami, 1981)

Further, Ikegami (1981) stated that clause-initial topicalization is often employed to present Ground information, thereby establishing the context, as exemplified in (6). Note that in Japanese the Figure-Ground word order is also allowed as an option, yet the Ground-Figure order seems to be predominance due to the syntactic Japanese topic argument.

Thirdly, the encoding of Figure and Ground varies for the types of motion events involved. In a nonagentive clause expressing spontaneous voluntary motion, as illustrated in (7a), the Figure is typically the subject and the Ground is the (oblique) object. In an agentive clause expressing caused-motion as in (7b), where the agent serves as the subject, the Figure assumes the role of the direct object and the Ground is the oblique object describing the goal or location, typically indicated by spatial prepositions such as "on" in 7b. This effect of motion event type on the encoding of core entities is also applicable to both Japanese and Chinese.

(7) a. Sarah runs into the room. (spontaneous motion)
b. Sarah puts the cup on the table. (caused motion)
(adapted from Choi et al., 2018)

Finally, the specific types of Ground within a given language play a critical role in the encoding. Talmy (2000) presented that different path terms are used depending on the three basic types of path phases: leading from a Source (FROM), leading to a Goal (TO), or one in between (VIA) (p. 53). Tajima & Duffield (2012) conducted a series of experiments involving Japanese, English, and Chinese native speakers, and found that Japanese speakers are unable to omit Ground information, as it is obligatory to mention it before reporting Figure information. Consequently, they are expected to include a disproportionate number of Ground elements relative to Figure elements. Further, as compared to English and Chinese speakers, Japanese speakers are inclined to devote greater attention to the Ground element. However, there remains a scarcity of detailed quantitative research concerning the encoding of the three specific types of Ground elements.

RESEARCH QUESTIONS

Previous studies have shown that the line of inquiry is fruitful for Second Language Acquisition research by examining the process of adaption to L2 learners' thinking in a new language. However, this type of research mainly focused on the expressions of the semantic components of Path and Manner, leaving other components like Figure and Ground neglected. Thus an analysis of the encoding of Figure and Ground components could make significant contributions to this line of inquiry. Further, previous studies provide clear evidence to suggest that relevant syntactic variations in Japanese and Chinese are sensitive to the encoding of Figure and Ground relations. Therefore, we set out to examine how JFL learners express Figure and Ground components in describing pictures, by comparing them with Japanese native speakers and Chinese native speakers. The goal is to see whether the narratives by JFL learners will be more aligned with Japanese native speakers or Chinese native speakers, addressing the role of L1 transfer in shaping the Figure and Ground descriptions. Specifically, this study addressed the following three research questions:

1) What are the differences in the number of elicited motion events across the three groups?
2) What is the encoding of Figure and Ground information across the three language groups (including an overview of frequency and percentage of Figure and Ground components, description of various Ground components, and the encoding of F-G word order)?

MATERIALS AND METHODS

PARTICIPANTS

The current study involved 22 subjects per language group to do the narrative task: JFL learners and Chinese native speakers were recruited at a University in China, while Japanese native speakers at a University in Japan. All the subjects were matched with ages ranging from 18-25 years old. In order to control for the language background, while JFL learners are subjects in China who learn Japanese as their major, Japanese native speakers and Chinese native speakers are those who have no knowledge of Chinese and Japanese respectively. All the subjects volunteered to participate in the task and received monetary compensation for participation. This study was conducted by a teacher at the respective university. To guarantee the participants' full cooperation, the instructions about the purpose and requirements of collecting their data were first provided before the task.

MATERIALS

This study employed a written picture elicitation task utilizing giving subjects a serial of wordless cartoon pictures in a comic book <Father and Son> (see Figure 1). It was about an incident concerning a father and his son titled "The Stubborn Apple", and it was intentionally selected because each of the pictures contained certain motion events. The JFL learner group was required to write the story down first in Japanese and then in Chinese for about 20 minutes respectively. Such an order can probably minimize the potential language transfer from Chinese. The Japanese and Chinese native speakers were also instructed to describe what happened in their respective languages by referring to the pictures. Our purpose was to compare the descriptions of three language groups with respect to the presentation of different types of spatial information encoded in their narratives.

Figure 1 Test Materials for the Written Narratives

DESIGN AND PROCEDURE

To ask the questions outlined in 2.4, we constructed a descriptive and comparative cross-linguistic quantitative design. The procedure was the same for all the three groups. The subjects were tested in their respective university, being required to sit in front of a laptop computer on which the instructions were presented. They were given three minutes to prepare with the series of pictures on the screen, then the teacher asked if they had any questions. Following this, the actual experiment started and continued for approximately 30 minutes.

CODING

Ultimately, we built four corpora based on the samples of the written narratives from the three groups of subjects (JFL learners produce two corpora presenting as Japanese and Chinese narrative data respectively). The written data in the four corpora were further divided into clauses, which were defined by Berman & Slobin (2013) as a linguistic unit that contains a unified predicate (as cited in Chen & Guo, 2009; Rungrojsuwan, 2023). They further explained "a unified predict" as one expressing a single situation. In this study, a predicate is a single motion verb or motion verb construction. Based on the motion verbs employed, the raters first identified the "utterances" expressed by motion verbs, which were coded as [CME], [MME], and [PME] for cause-of-motion-event, manner-of-motion-event, and path-of-motion-event respectively (as compared to previous studies, this study did not involve fictive motion). In this way, all of the narrative data in the four corpus were coded into these three motion event categories. Following that the Figure and Ground information in the narratives were coded within each clause, as indicated by parenthetical function markings denoted by the abbreviations "F" and "G". Further the coding for the Ground information was also aligned with previous studies such as Chen & Guo (2009), which means that the information representing Source, Via, and Goal was evidenced by parenthesized function markings abbreviatedly symbolized as "G1", "G2", and "G3". Note that as compared to Chen & Guo (2009), this study also involved agentive motion, and the agent in the utterances was coded as "AG". An example from a Chinese JFL learner is as follows:

(8) 父[AG]は木[G1]にかかった靴[F]をかさで取った。[CME] (S0420)
Chichi-wa ki-ni kakat-ta kutsu-wo kasa-de tot-ta.
My father took the shoe off the tree with an umbrella.
(9) 父与子[F]来到了公园[G3]。[PME](S0432)

Fu yu zi lai dao le gong yuan.
The father and son came to the park.

For the coding process, two advanced Japanese learners first coded five samples together. While meeting discrepancies, they referred to dictionaries or examples available in previous studies and discussed the doubtful cases until a consensus was reached. The inter-rater reliability of coding that of the written elicitation task was 0.94, showing a good agreement between the two raters (Cohen, 1968). After the data coding of written performance, all the data were entered into AntConc 3.5.8 for further analysis.

RESULTS AND DISCUSSION

AN OVERVIEW OF THE EXPRESSION OF MOTION EVENTS IN THE FOUR CORPORA

The first step in the data analysis set out to examine the overall number of elicited utterances (segmented by the occurrence of a motion verb or a motion construction). Figure 2 shows the differences in the number of elicited motion events across the three groups.

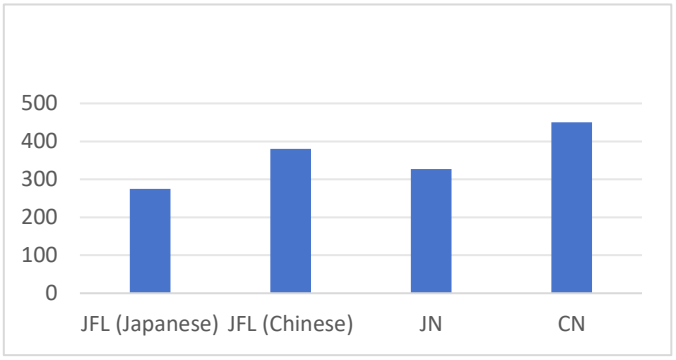

Figure 2. Numbers of Utterances Descripted by JFL Learners, Japanese Native Speakers, and Chinese Native Speakers

As was previously shown, speakers of Chinese native speakers produced the highest number of utterances, compared to the other two language groups. JFL learners produced more utterances when they conducted the narrative task in Chinese in contrast to in Japanese. This result converges with Slobin's (1996) as well as Chen & Guo's ((2009)) findings that S-languages are more likely to break an event into several segments and describe what happens segment-by-segment in the scene, thus producing more utterances than V-languages.

However, when identifying possible differences with respect to the types of motion events expressed in each group, we found that the performance exhibited greater complexity. To assess the numerical differences in the utterances expressed across three groups, the number of tokens based on motion event types for each language group was countered and percentage calculations were conducted. Table 1 shows that the rates of occurrence of cause motion events were very similar in the four coupora of slightly more than 50%. This indicated that the factor of motion event type does not significantly affect the proportions for cause motion events across the three groups. However, when identifying the specifics, we found that the Japanese cause motion events expressed in JFL learners took a slight lead over Japanese native speakers. We speculated that this difference between the two groups may be characteristic of languages that prefer to use causative constructions with "shi, rang, ling" versus languages that prefer to avoid causative constructions. As shown in example (10), the subject S0425 employed the causative verb "つけさせます (tsuke sase masu)" to

express the event of marking on the tree, which was not typically observed in Japanese sentences.

(10) 男の子[AG]は木の枝[G3]にしるし[F]をつけさせます。[CME](S0425)
Otoko-no-ko-wa ki-no-eda-ni sirusi-wo tsuke-sasemasu.
The boy makes a mark on the tree branch.

Table 1 The Tokens and Percentages of Expressed Utterances by Motion Event Type in the Four Corpora

	CME	MME	PME	Total
JFL (Japanese)	149 (54.2%)	30 (10.9%)	96 (34.9%)	275
JFL (Chinese)	201 (52.9%)	113 (29.7%)	66 (17.4%)	380
JN	168 (51.4%)	47 (14.4%)	112 (34.3%)	327
CN	233 (51.8%)	143 (31.8%)	74 (16.4%)	450

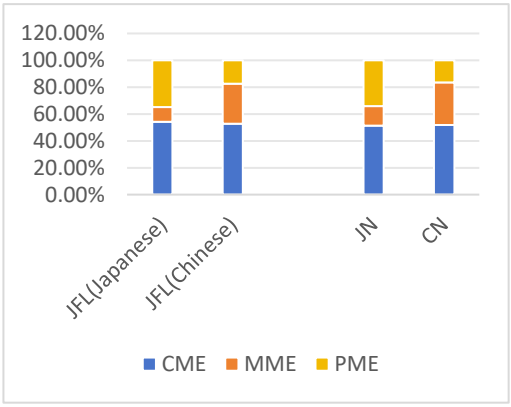

Figure 3 Visual Representation of Occurrences of Three Types of Motion Events in the Four Corpora

Notably, as shown in Table 1, the performance of JFL learners in Chinese was likely to be that of Chinese native speakers, while the performance of JFL learners in Japanese seemed to resemble that of Japanese native speakers. This becomes more evident upon visual representation of the percentages of occurrences of three types of motion events in the four corpora in Figure 3. We argued that this resonates with the finding reported by Slobin (2004) that S-languages tend to employ more manner verbs while V-languages tend to use more path verbs in describing the same motion event.

THE ENCODING OF FIGURE AND GROUND INFORMATION IN THE FOUR CORPORA

AN OVERVIEW OF FIGURE AND GROUND INFORMATION

The second step in the analysis focused on whether the expression of Figure and Ground components deviated across the three language groups. For this analysis, we quantified the occurrences of the two components observed within each group. Table 2 presents an overview of the entity information that subjects

provided in their narratives.

Table 2 The Tokens and Percentages of Figure and Ground Information in the Four Corpora

	F	G
JFL (Japanese)	234 (85.1%)	88 (32%)
JFL (Chinese)	314 (82.6%)	122 (32.1%)
JN	268 (82.0%)	113 (34.6%)
CN	360 (80%)	154 (34.2%)

As discussed in previous studies (see section 2.3), it was predicted that the performance of the Japanese native speaker group should diverge significantly from that of the other two groups, which means that the Japanese tend to attend more ground elements than other language groups do. However, this anticipated outcome was not observed in Table 2. We further counted the occurrences of Figure and Ground information expressed for each motion event type, detecting whether there was a tendency in a given language to provide more ground information.

Table 3 The Tokens and Percentages of Figure and Ground Information by Motion Event Type in the Four Corpora

	CME		MME		PME	
	F	G	F	G	F	G
JFL (Japanese)	126 (45.8%)	40 (14.5%)	23 (8.4%)	16 (5.8%)	85 (30.9%)	32 (11.6%)
JFL (Chinese)	175 (46.1%)	42 (11.1%)	77 (20.3%)	70 (18.4%)	62 (16.3%)	10 (2.6%)
JN	151 (46.2%)	33 (10.1%)	34 (10.4%)	30 (9.2%)	83 (25.4%)	50 (15.3%)
CN	205 (45.6%)	46 (10.2%)	95 (21.1%)	86 (19.1%)	60 (13.3%)	22 (4.9%)

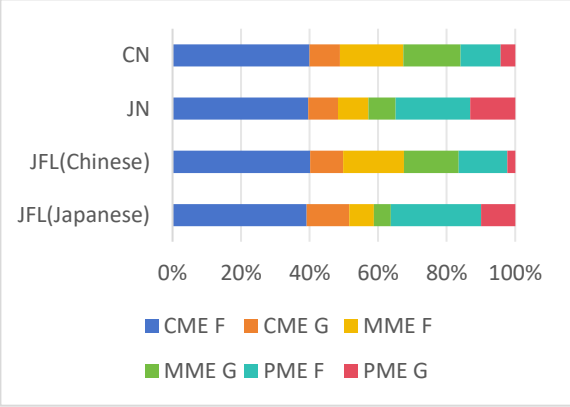

Figure 4 Visual Representation of Occurrences of Figure and Ground Elements by Motion Event Type in the Four Corpora

Unexpectedly, based on Table 3 and Figure 4, there was a general tendency to produce more Figure information than Ground information by each group. However, given the factor of motion event type, potential differences arose due to the complexity of the expressed motion events as seen in Figure 4.

Specifically, we firstly compared the Chinese narratives of JFL learners to those of Chinese native speakers, and found that the cause and manner motion events had some similarity in terms of the statistical distribution of Figure and Ground components, except for path motion events, in which Chinese native speakers referred to more ground information. Then when comparing the Japanese narratives of JFL learners with those of Japanese native speakers, we found that the types of motion events affected the F-G encoding differentially, particularly in their use of the manner and path motion events. In light of these events, Japanese native speakers produced more ground information than JFL learners, revealing the preference for Japanese to attend more ground elements to some extent. Further through the comparison across the three groups, it was apparent that the Japanese narratives by JFL learners perform similarly to Japanese native speakers rather than Chinese native speakers, indicating less language transfer from Chinese.

DESCRIPTION OF GROUND COMPONENTS

The next step in the analysis concerned a closer look at the differences in the specific types of Ground information encoded which led to contrasts expressed behind the numbers. The statistical data across the four corpora were presented in Table 4, with a graphical depiction provided in Figure 5. Roughly, the Ground information across the four corpora exhibited a similar tendency highlighted in the preceding section: the Japanese narratives by JFL learners demonstrated parity with Japanese native speakers, while the Chinese narratives by JFL learners displayed a similar pattern to Chinese native speakers. Further, a comparison of the results indicated a need to explore the variations in the preferred type of Ground information across the four corpora.

Table 4 The Tokens and Percentages of the Specific Types of Ground Information in the Four Corpora

	CME			MME			PME		
	G1	G2	G3	G1	G2	G3	G1	G2	G3
JFL (Japanese)	5 (1.8%)	0 (0%)	35 (12.7%)	1 (0.4%)	3 (1.1%)	12 (4.4%)	7 (2.5%)	1 (0.4%)	24 (8.7%)
JFL (Chinese)	15 (3.9%)	0 (0%)	27 (7.1%)	0 (0%)	1 (0.3%)	69 (18.2%)	3 (0.8%)	1 (0.3%)	6 (1.6%)

JN	4 (1.2%)	0 (0%)	29 (8.9%)	1 (0.3%)	1 (0.3%)	28 (8.6%)	11 (3.4%)	1 (0.3%)	38 (11.6%)
CN	11 (2.4%)	0 (0%)	35 (7.8%)	0 (0%)	4 (0.9%)	82 (18.2%)	3 (0.7%)	1 (0.2%)	18 (4.0%)

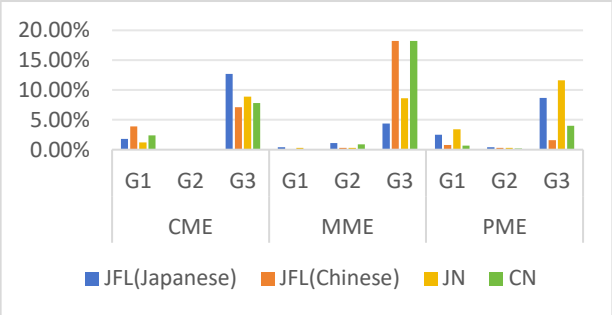

Figure 5 Visual Representation of Occurrences of Specific Types of Ground Components in the Four Corpora

Specifically, as noted in Table 4 and Figure 5, the occurrence of G3 was the highest in all the Ground information, and it had nearly identical rates across the four corpus: 25.8% in the Japanese narratives of JFL learners, 26.9% in the Chinese narratives of JFL learners, 29.1% in the narratives produced by Japanese native speakers, and 30% in that of Chinese native speakers. However, when considering the impact of motion event type, the performance was notably intricate, especially the G3 information expressed for manner and path motion events by Japanese native speakers was evidently higher (as shown in example 11), aligned with Tajima & Duffield's (2012) finding that Japanese relatively mentioned more Ground information. In example 11, subject SJ17 placed a destination ground before the motion verb "ochi-ru", whereas Chinese native speakers typically prefer to add a deixis verb "lai" before the serial verb construction "diao xia".

(11) あの頑固なリンゴ[F]が自ら地面[G3]に落ちたのです。(SJ17)
Ano ganko-na ringo-ga mizukara jimen-ni ochi-ta no desu.
That stubborn apple fell to the ground by itself.

When concerned with the G1 information, the Japanese narratives by JFL learners seemed remarkably different from their corresponding Chinese narratives, yet behaved more similarly to Japanese native speakers. This observation suggests that the JFL learners had acquired the expression for GI components to some extent. Further, while instances of G1 information represented were few, even fewer were realized by G2 components, which can be explained with reference to the nature of the experience of construing motion events.

THE ENCODING OF F-G WORD ORDER

The final section in the analysis provided an overview of the differences in the encoding of Figure trajectory to the Ground in the four corpora. Firstly, proportions of different F-G relations were calculated for each group by motion event type. Our results suggested that the F-G relation had a striking lead over the G-F

relation, which was aligned with Talmy's (2000) proposal that the Figure holds syntactic precedence over the Ground.

Table 5 The Tokens and Percentages of the F-G Relations in the Four Corpora

	CME		MME		PME	
	F-G	G-F	F-G	G-F	F-G	G-F
JFL	21	9	14	0	27	0
(Japanese)	(29.6%)	(12.7%)	(19.7%)	(0%)	(38.0%)	(0%)
JFL	23	13	43	0	9	0
(Chinese)	(26.1%)	(14.8%)	(48.9%)	(0%)	(10.2%)	(0%)
JN	23	7	15	10	35	3
	(24.7%)	(7.5%)	(16.1%)	(10.8%)	(37.6%)	(3.2%)
CN	21	11	51	3	16	0
	(20.6%)	(10.8%)	(50%)	(2.9%)	(15.7%)	(0%)

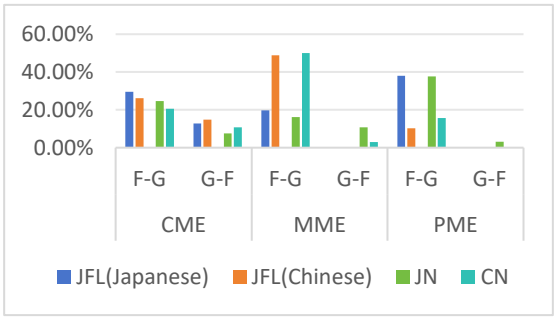

Figure 6 Visual Representation of Occurrences of F-G Relations in the Four Corpora

However, given the syntactic structure contrasts described in section 2.3, we expect differences concerning the word order of the Figure and Ground components across the language groups. Regarding the F-G relations, there was a tendency for the Japanese narratives produced by JFL learners to exhibit alignment with those of Japanese native speakers, whereas the Chinese narratives demonstrated alignment with those of Chinese native speakers. In terms of the G-F relations, it was evident that Japanese native speakers showed a remarkable distribution that some kind of Ground information was referenced before the Figure information, especially for manner motion events and path motion events (as shown in example 12).

(12) 父親の頭[G3]に子供の投げたもの[F]が当たった。(SJ20)

Chichioya-no atama-ni kodomo-no nage-ta mono-ga atat-ta.

The father's head was hit by something thrown by the child.

(13) 两人[AG]想尽办法将树上[G1]的苹果[F]摘下来。(s0432)

Liang ren xiang jin ban fa jiang shu shang de ping guo zhai xia lai.

They tried their best to pick the apple from the tree.

In example 12, the subject SJ20 was likely to organize the sentence by first placing peripheral elements, thus constructing the Ground to Figure order in reporting the scene. This was rarely observed in their counterparts by Chinese expressions. However, the G-F order is also applicable in Chinese narratives when the Ground occurs in the attributive clause that modifies the Figure as shown in example 13.

GENERAL DISCUSSION

The current study examined the expression of Figure and Ground information by the three language groups through a written narrative task. Our results partly approved the applicability of the typological framework proposed by Talmy (1985, 2000). However, when regarding the motion event type, the results were not exactly like what would be expected based on previous typological findings. Further, in the encoding of Figure and Ground components, this study found that the Japanese narratives by JFL learners performed similarly to Japanese native speakers rather than Chinese native speakers, showing a limited role of the L1 thinking for speaking patterns.

Specifically, this study showed that Chinese narratives produced more utterances, and they tend to employ more manner verbs while Japanese narratives tend to use more path verbs in describing the same motion event. In addition, the F-G relation in the narratives across the three groups had a striking lead over the G-F relation. Such findings followed from language-specific properties related to how S-framed (Chinese) versus V-framed (Japanese) languages grammaticalize the Figure and Ground information. They therefore indicated the impact of typological constraints on second language acquisition. On the other hand, branded as a so-called S-language in Talmy's (1985) framework, it was suggested that Chinese speakers tend to attach more ground elements than Japanese (Slobin, 1996, 2004; Chen & Guo, 2009). However, the Japanese do not necessarily produce fewer ground elements, conversely, as Tajima & Duffield (2012) mentioned, Japanese are required to attend more to the Ground elements because they cannot skip them. In our dataset, we found variations for different rates of occurrence of Ground (G1, G2, G3) and their prominence in each corpus by motion event type (Figure 5). These differences in typological studies allow a detailed cross-linguistic analysis of the linguistic representation of the motion events in the future.

Secondly, regarding the impact of motion event type, this study clearly showed that Japanese native speakers produced evidently more ground information than JFL learners, particularly in their use of the manner and path motion events; Japanese native speakers expressed more G3 information for manner and path motion events; Japanese native speakers showed a remarkable distribution that some kind of Ground information was referenced before the Figure information, especially for manner and path motion events. The results reconfirmed that Talmy's (1985, 2000) typology framework was proposed in a broad sense, and it could not reflect the conceptualizations of motion events that occur around us. Therefore, diverse types of motion event descriptions should be examined to test the validity of previous proposals in a narrow sense.

Thirdly, the results also suggest that subjects in the three groups differed remarkably in the way they linguistically encoded the syntactic F-G relations. We cannot simply address that a language is likely to report more Ground information, either before or after the Figure information, than another language. It is worth considering that several marked syntactic variations (as discussed in 2.3) had to do with the expression of the two core entities. However, it still puts into question to what extent the F-G relation is affected by syntactic complexity (Lozinska & Hendriks, 2024).

Finally, the findings of the token and type analysis generally showed that the JFL learners patterned to a large degree with the Japanese native speakers, separately from the Chinese native speakers, which means that JFL learners have acquired the target-like way to report more Ground information and more Ground-Figure order sentences, especially in spontaneous motion events (as seen in Figure 6). These results seemed to support the view of a rather limited role of the L1 thinking for speaking patterns (Cadierno & Ruiz, 2006). This leads us to hypothesize that in this study the influence of the L1 thinking for speaking patterns might be applicable for cause motion events, but that it is limited for path and manner motion events in the

acquisition process. This explanation agrees with previous studies that L1 transfer is a complex phenomenon constrained by factors such as markedness of the linguistic feature (Fagard et al., 2017; Liao et al., 2021).

Regarding that there are some factors contributing to the inconsistency of the language transfer studies, such as typological differences, modality (written versus spoken, picture-based versus movies-based narration), task differences (e.g., elicitation task, picture-matching task, memory task, or translation task), and the complexity of structures being investigated, future studies should conduct detailed analysis about how and to what extent do L2 learners acquire the motion events, especially for language combinations that are typologically overlapping, such as Japanese and Chinese.

CONCLUSION

This research showed striking cross-linguistic differences in how subjects in the three groups expressed Figure and Ground information. Such differences demonstrated that the notion of a two-fold typology system is too broad to explain observed linguistic representation in the acquisition process. We partly converge with the typological proposals for that by rough contrast Japanese and Chinese native speakers in our dataset exhibited a respective pattern aligned with the proposal. However, the striking differences followed from syntactic structure variations in the utterances related to how Japanese versus Chinese encode the Figure and Ground information. These results suggest the joint impact of typological and syntactic factors on language acquisition. Furthermore, there is a misalignment with the role of L1-specific patterns on L2 learners' acquisition of Figure and Ground information. Such misalignment calls for further studies on the influencing factors.

We hope to have shown that the cross-linguistic research on the acquisition of Figure and Ground information can contribute to the SLA research. This research not only constitutes the empirical basis for testing the typological framework, but also sheds light on the question of the extent to which JFL learners can adapt to thinking for speaking in a new language. Furthermore, this kind of research bears significant pedagogical implications for enhancing language learners' production, which means that the syntactic variations observed in the encoding of motion events can be incorporated into teaching practice (Xu et al., 2024). Future research is necessary to address further questions raised by these results concerning the relationship between language acquisition and teaching.

Finally, there may be some uncontrollable factors affecting the acquisition process, such as the Japanese and Chinese native speakers in the current study may not be purely monolinguals, which is difficult to fulfill in reality. We also concede that the number of elicited pictures presented to subjects was smaller than standard for such studies, and also a limitation arising from the content constraints in certain picture books which results in a restricted corpus quality about identical thematic content. Future research can consider further enriching data collection materials, but not being constrained by particular storylines.

REFERENCES

Akita, K. (2017). *The typology of manner expressions: A preliminary look*. InI. Ibarretxe-Antuñano (Ed.), Motion and space across languages: Theory and applications (pp. 39–60).

Anastasio, S. (2023). Motion event construal in L2 French and Italian: from acquisitional perspectives to pedagogical implications. *International Review of Applied Linguistics in Language Teaching*, 61(1), 37-60.

Berman, R. A., & Slobin, D. I. (Eds.). (2013). *Relating events in narrative: A crosslinguistic*

developmental study. Psychology Press.

Brown, A., & Gullberg, M. (2013). L1–L2 convergence in clausal packaging in Japanese and English. *Bilingualism: Language and Cognition*, *16*(3), 477-494.

Cadierno, T. (2004). Expressing motion events in a second language: A cognitive typological perspective. *Cognitive linguistics, second language acquisition, and foreign language teaching*, 13-49.

Cadierno, T., & Ruiz, L. (2006). Motion events in Spanish L2 acquisition. *Annual review of cognitive linguistics*, *4*(1), 183-216.

Chen, L., & Guo, J. (2009). Motion events in Chinese novels: Evidence for an equipollently-framed language. *Journal of pragmatics*, *41*(9), 1749-1766.

Choi, S., Goller, F., Hong, U., Ansorge, U., & Yun, H. (2018). Figure and Ground in spatial language: evidence from German and Korean. *Language and Cognition*, *10*(4), 665-700.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, *70*(4), 213.

Fagard, B., Stosic, D., & Cerruti, M. (2017). Within-type variation in Satellite-framed languages: The case of Serbian. *STUF-Language Typology and Universals*, *70*(4), 637-660.

Gu, Y., & Lin, Z. J. (2017). An empirical study of Chinese EFL learners motion event encoding in the way-construction. Modern Foreign Languages, (1), 336–347.

Hasko, V. (2010). The role of thinking for speaking in adult L2 speech: The case of (non) unidirectionality encoding by American learners of Russian. *Linguistic relativity in SLA: Thinking for speaking*, 34-58.

Ikegami, Y. (1981). *Suru to Naru no Gengogaku [Linguistics for "doing" and "occurring"].* Tokyo: Taishukan Shoten.

Kashyap, A. K., & Matthiessen, C. M. (2017). Figure and ground in the construal of motion: a registerial perspective. *Word*, *63*(1), 62-91.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Volume I: Theoretical prerequisites (Vol. 1)*. Stanford university press.

Liao, Y., Dijkstra, K., & Zwaan, R. A. (2021). Directional prepositions and event endpoint conceptualization: A study of naar and richting in Dutch. *Language and Cognition*, *13*(2), 161-190.

Li, F.Y. (2017). The representation of Path components in the typical translocational Motion Events. *Foreign Language Education*, 38 (4), 1-6.

Li, X. (2011). A contrastive study of the grammatical realization of figure and ground in motion events between Chinese and English. *Foreign Language Teaching and Research (bimonthly)*. 43(5), 689-701.

Lozinska, J., & Hendriks, H. (2024). Lexicalization of Pulling and Pushing Events in Polish and English. *GEMA Online® Journal of Language Studies*, 24(1), 1-17.

Matsumoto, Y. (2018). Motion event descriptions in Japanese: Typological perspectives. *Handbook of Japanese contrastive linguistics*, 273-289.

Matsumoto, Y., & Kawachi, K. (Eds.). (2020). *Broader perspectives on motion event descriptions (Vol. 69)*. John Benjamins Publishing Company.

Muehleisen, V., & Imai, M. (1997). Transitivity and the incorporation of ground information in Japanese path verbs. *Lexical and syntactic constructions and the construction of meaning*, 329-346.

Özçalışkan, Ş. (2004). Typological variation in encoding the manner, path, and ground components of a metaphorical motion event. *Annual Review of Cognitive Linguistics*, *2*(1), 73-102.

Park, H. I., Jarvis, S., & Kim, J. E. (2022). Exploring motion event construal: How much attention do speakers of different languages and cultures pay to context?. *Lingua*, *265*, 103164.

Rungrojsuwan, S. (2023). The Development of Narrative Skills of Thai Children. *GEMA Online Journal of Language Studies*, *23*(3).

Schmiedtová, B. (2013). Traces of L1 patterns in the event construal of Czech advanced speakers of L2 English and L2 German. *International Review of Applied Linguistics in Language Teaching*, *51*(2), 87-116.

Shen, J. X. (2003). The resultative construction in Chinese: A typological perspective. *Chinese Teaching in the World*, (3), 17–23.

Slobin, D. I. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics. Quarterly publication of the international pragmatics association (ipra)*, 1(1), 7-25.

Slobin, D. I. (1996). Two ways to travel: Verbs of motion in English and Spanish. *Grammatical constructions: Their form and meaning*, *195217*.

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, Volume 2. Psychology Press.

Sung, M. & Kim, K. (2020). Spontaneous motion in L1- and L2-English speech: A corpus-based study. *English Teaching*, 75(1), 49–66.

Tajima, Y., & Duffield, N. (2012). Linguistic versus cultural relativity: On Japanese-Chinese differences in picture description and recall. *Cognitive Linguistics*, *23*(4), 675-709.

Talmy, L. (1978). Figure and ground in complex sentences. *Universals of human language*, (4), 631-649.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description*, *3*(99), 36-149.

Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems* (Vol. 1). MIT press.

Wang, Y., & Wei, L. (2023). Multilingual learning and cognitive restructuring: The role of audiovisual media exposure in Cantonese–English–Japanese multilinguals' motion event cognition. *International Journal of Bilingualism*, *27*(3), 331-348.

Wu, S. L., Nunome, T., & Wang, J. (2022). Crosslinguistic influence in the conceptualization of motion events: A synthesis study on L2 acquisition of Chinese motion expressions. *Second Language Research*, 02676583221132205.

Xu, Z. R, Roslina, M., & Zaid, Z. (2024). Conceptual transfer studies in the context of Second Language Acquisition: A scoping review. *Akademika*, 94(2), 110-128.