Motion Verbs in English and Spanish: A Comparative Corpus Approach

Masterarbeit
zur Erlangung des akademischen Grades
Master of Arts (M.A.)
der Philologischen und der Philosophischen Fakultät
der Albert-Ludwigs-Universität Freiburg im Breisgau

vorgelegt von
Luz Andrea Alvariño Gabeiras
aus Ourense, Spanien

Wintersemester 2022

Englische Sprachwissenschaft

Erstgutachter: Prof. Dr. Dr. h.c. Christian Mair

Table of Contents

1. Introduction

As a conceptual domain, Motion permeates our everyday lives and language. Without locking ourselves in a giant plastic bubble, we cannot avoid experiencing events of motion; and even the giant bubble might roll or skip away at some point. As an universal experiential domain, Motion is an essential part of communication in most languages, if not all. Therefore, it serves as an excellent testing ground for theories deriving from Linguistic Relativity, which propose that we segment and conceptualize reality differently depending on the language we speak, leading to separate and relatively distinct view of the world. One such theory is Dan Slobin's Thinking for Speaking (1987, 1996a, 1996b, 2000, 2003, 2005), which proposes that speakers conceptualize reality differently while in the process of communication, leading to patterns that can be observed and studied through speech. Motion events make up a large part of Thinking for Speaking research, as they are frequent in almost all languages, and present enough typological variation to conceivably lead to differences in schematization. For instance, a simple clause such as 'the dog wiggled away,' in which the way that the dog moves is encapsulated in the main verb 'wiggle' and the direction in the attached particle 'away'—would be verbalized entirely differently in Spanish, as the most natural translation would be "el perro salió retorciendose" (*'the dog went-away wiggling'*), in which the direction can be found in the main verb, and the way the dog leaves is in the attached non-finite verb.

The present dissertation adapts Slobin's model, employing some of his more traditional methodologies like elicited narrations in combination with gesture analysis as a measure of conceptualization; alongside a relatively novel large corpus approach. This methodology serves the main goal to answer the over-arching research question of whether differences in motion event conceptualization present in short narrations and small corpora, such as a preference for manner-of-motion in English, can also be demonstrated in large corpora and nonlinguistic tasks. From this main exploratory point, a series of smaller research questions arise, which are presented in Chapter 3. The study compares English and Spanish as hallmark examples of the two main patterns outlined in Thinking for Speaking.

These patterns, alongside a detailed overview of the theoretical and typological background of Thinking for Speaking, are presented in Chapter 2; as well as a summary of the research framework applied in the gesture analysis, combining speech, gesture and cognition. Chapter 3 reviews the data collection and analysis methodologies employed in the study, the

results of which are then presented in Chapter 4. Chapter 5 attempts to contextualize the results within the existing research, and Chapter 6 concludes with a point-by-point summary of the study's findings.

All in all, the present dissertation aims to serve as a mainly corpus-based crosslinguistic examination of motion verbs and motion event description in English and Spanish, adding some supporting or contradicting data to the Thinking for Speaking framework within Linguistic Relativity.

2. Literature Review

As explained in the Introduction, Chapter 2 offers an in-depth introduction to the theories necessary to understand the research conducted later in the dissertation. The chapter begins by outlining Thinking for Speaking as a model, and later delves into its applications for gesture research.

2.1. Thinking for Speaking

2.1.1. Theoretical Background

Historical background and Theoretical Principles. Thinking for Speaking (hereafter named TFS) refers to Dan Slobin's model of linguistic relativity, which proposes a strong link between language and thought while in the process of communication (Slobin, 1987, 1991, 1996a, 1996b, 2000, 2005).

As a research area, linguistic relativity is centered on the idea that a speaker's native language determines or greatly influences their way of thinking about reality, their 'world view' (Kaye, 2009, p. 116; Pederson, 2009, p. 663). The 20th-century theory is generally attributed to scholars Edward Sapir and Benjamin Whorf and has in turn sometimes been named the 'Sapir-Whorf Hypothesis'. Sapir and Whorf, however, draw upon ideas from earlier thinkers such as Wilhelm von Humboldt or Franz Boas (Kaye, 2009, p. 116).

As part of linguistic relativity studies, TFS works on three main central points:

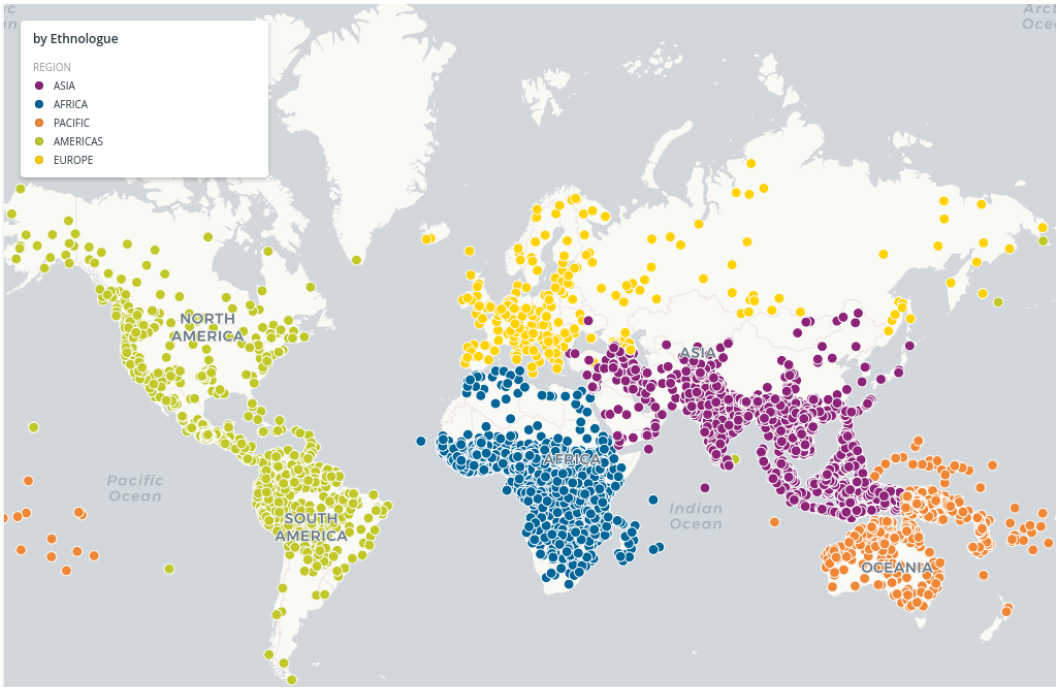
a. "each speech community embodies a distinct world view,"
b. "languages differ from one another", and
c. "thought and language are inseparable" (Slobin, 1996a, p. 70).

Much of linguistic relativity theory hinges on the notion of '*Weltansichten'* or world view. Humboldt famously stated that "the diversity of languages is not a diversity of sounds and signs but a diversity of the views of the world"[1] (1903, p. 27), giving way to the popular definition of world view as a unique way of observing and interacting with reality that arises from language.

[1] Translated. Original: "Ihre Verschiedenheit ist nicht eine von Schällen und Zeichen, sondern eine Verschiedenheit der Weltansichten selbst" (Trabant, 2000, p. 29).

Figure 1

Living languages of the world



Note: Taken from *Living Languages.* (2022). [Map]. SIL International, CARTO.

The linguistic diversity Humboldt is referring to hardly needs to be substantiated; as of its 25th edition (2022), the Ethnologue linguistic atlas compiles a total of 7,152 different 'living languages' (see Figure 1). The lexical and morphosyntactic differences between these languages have been at the forefront of linguistic typology research since the late twentieth century; the focus, which had once been on the 'hard-to-obtain' linguistic universals, has more recently shifted to formal diversity between languages (Song, 2011). For instance, example 0 illustrates lexical differences between languages in the description of 'body-part terms':

[0] West-Central Oromo (Cushitic; Ethiopia, Kenya); also English:

quba 'finger'

harka 'hand'

… Warlpiri (Pama-Nyungan, Australia):

rdaka 'hand and fingers'

(Moravcsik, 2013, p. 33)

If reality can only be perceived subjectively through language, as Humboldt suggests, it would be only logical that linguistic differences like the one above should result in different

world views for separate speech communities. Furthermore, linguistic relativity suggests that languages 'segment' objective reality in smaller sections or concepts for our subjective perception; in doing so, different languages prioritize different aspects of reality, which may manifest in their speakers' conceptualization of reality (Boas, 1911, p. 39, as cited in Lucy, 1992, p. 13; Humboldt, 1903, p. 29; Trabant, 2000, p. 34). As stated by Whorf (1956),

> the users of markedly different grammars are pointed by the grammars towards different types of observations and different evaluations of externally similar acts of observation, and hence are not equivalent as observers but must arrive at somewhat different views of the world. (p. 221, as cited in Lucy, 1992, p. 38)

TFS, accordingly, suggests that speakers conceptualize reality through the 'grammatical system' of their language, which determines how the "content of any particular sentence" is understood (Slobin, 1991, p. 7). Certain aspects of reality are "enshrined" in the formal characteristics of every language (Slobin, 1996a, p. 71); if these characteristics vary, so will the segments of objective reality that speakers are pointed towards. In example (2),

(2)[2] English: 'Me and the dog walked'
German: *Ich* *und* *der* *Hund* *gingen* *spazieren*
me and the.NOM dog us.went walk

English: 'I fed the dog'
German: *Ich* *fütterte* *den* *Hund*
I fed the.ACC dog

Therefore, the typological differences between languages might lead to subsequent differences in schematization.

Finally, whether language merely influences cognition or rather completely determines it is the single most conflictive point within linguistic relativity research. Boas (1911), for instance, recognizes a two-way interdependence between language and thought, stating that cognition and culture may also influence language; as humans sharing the same key "fundamental psychological processes" will lead to the same core "grammatical concepts" in all languages (p. 67, as cited in Lucy, 1992, p. 14). In contrast, Sapir and Whorf believe that

[2] Example format loosely inspired by Moravcsik (2013).

linguistic differences are not a reflection of cognition but rather, that thought is only possible when channeled through language, making the two inseparable (Lucy, 1992, pp. 17–22),

> humans … are very much at the mercy of the particular language which has become the medium of expression for their society … the 'real world' is to a large extent unconsciously built up on the habits of the group. No two languages are ever sufficiently similar to be considered as representing the same social reality. The worlds in which different societies live are distinct worlds, not merely the same world with different labels attached. (Sapir, 1929, p. 209, as cited in Lucy, p. 22)

Nonetheless, Sapir's 'weaker' linguistic relativism follows Boas in stating that although language and thought are inseparable, language does not inherently determine or restrict cognition, and linguistic differences alone are not enough to account for differences in culture and behavior (Kaye, 2009, p. 118; Lucy, 1992, pp. 17–22; Pederson, 2010, p. 1013; Whorf, 1956, p. 221),

> culture may be defined as what a society does and thinks. Language is a particular *how* of thought. It is difficult to see what particular causal relations may be expected to subsist between a selected inventory of experience (culture, a significant selection made by society) and the particular manner in which the society expresses all experience [language]. (Sapir, 1921, p. 2018, as cited in Lucy, 1992, p. 22)

In contrast, Whorf's 'stronger' linguistic determinism (Pederson, 2009, pp. 664–665) posits that thought shapes thought and ultimately behavior, as reality must be filtered through language in order to be cognitively processed, and it is the resulting conceptualization that determines a speaker's actions.

Following the Sapir-Whorf hypothesis, TFS pushes back against the idea of thought independent from language. Where Boas states that in speaking, we "express only a part of the complete concept that we have in mind" (Boas, 1911, pp. 38–39, as cited in Slobin, 1991, p. 8), Slobin argues that without language, it is near impossible to determine what aspects of reality would constitute a 'complete' mental image:

> a particular utterance is never a direct reflection of "objective" or perceived reality or of an inevitable and universal mental representation of a situation. This is evident within any given language, because the same situation can be described in different ways; and it is evident across languages, because each language provides a limited set of options for the grammatical encoding of characteristics of objects and events. (Slobin, 1987, p. 435)

As children acquire a particular language, they learn to focus on the domains that are expected in its grammatical characteristics. For instance, Turkish has a past tense inflection for 'witnessed events' and one for 'non-witnessed events,' so speakers must necessarily pay attention to that aspect of reality. Taking the example, Slobin questions whether even speakers of a language without such inflections would invariably notice whether someone has or has not witnessed the event and incorporate that supposedly complete mental image (Slobin, 1991, pp. 7–11).

By assuming that the formal characteristics of a language intimately influence the resulting mental representations, Slobin adopts a tight view of the relationship between word forms and their conceptual meanings, in which "words map closely onto coherent packets of nonlinguistic knowledge constituting concepts" and "cross-linguistically variable word meanings lead to culturally variable concepts" (Malt et al., 2010, pp. 29, 31), that is,

> by acquiring and using the naming patterns of their native language, speakers of different languages come to have concepts shaped by those patterns. If different languages have different meanings associated with words, then speakers of those languages will have correspondingly different concepts, producing the tight linkage. (Malt et al., 2010, p. 31)

From this point of view, our 'mental image' of reality (or world view) will be determined to a great extent by whatever languages we speak, taking that culture and behavior depend on the mental images resulting from lexicalization. In this sense, TFS follows the stronger current of linguistic relativity.

Why 'Speaking'? TFS focuses on the thinking that is directly involved in speech, occurring directly before or while producing an utterance (Slobin, 1987, p. 435, 1991, p. 23, 1996a, pp. 70–75, 2000, p. 108). Theories within linguistic relativity have been criticized for not allowing empirical testing and relying too strongly on faith-based beliefs to affirm that

cognition varies in different languages (Chomsky et al., 2005; Langacker, 1976; Lenneberg, 1953; McWhorter, 2014; Pinker, 2007). Linguistic determinism, in particular, has been largely rejected; the possibility of translating from one language to another, or the ability to acquire new languages throughout our lifetimes—among other aspects—have been taken to prove that cognition cannot be completely determined by language (Pederson, 2009, p. 665). However, instead of attempting to concretely define concepts like 'world view' and finding a way to prove their relation to language, Slobin narrows down TFS theory to one over-arching hypothesis: that regardless of whether an individual's native language *always* determines the way they think about the world, it will constrict the way they can verbally describe it and impact the mental processes involved in active communication (Slobin, 1996a, pp. 70–75). For Slobin (1987, p. 435),

> the activity of thinking takes on a particular quality when it is employed in the activity of speaking. In the evanescent time frame of constructing utterances in discourse, one fits one's thoughts into available linguistic forms.

That is, for reality to be verbalized it must be "filtered through language" using the formal characteristics of a speaker's language (Slobin, 1996a, pp. 75–76); accordingly, speakers must think about whatever semantic domains are nested in said formal aspects while they are in use (p.71). Slobin replaces the relationship between thought and language with the mirroring notions of 'thinking' and 'speaking.' By using the gerund form, Slobin re-defines the concepts as dynamic entities, directly involved in the ongoing act of producing speech (Slobin, 1987, 1991, 1996a). Since the model focuses on the aspects of cognition involved with communication, it offers the opportunity of circumventing issues of traditional linguistic relativity, such as the issue of measuring responses to 'non-linguistic tasks,' necessary for the study of 'non-linguistic cognition (Pederson, 2009, p. 671).

Pederson (2009) criticizes Slobin for proposing a "model of cognition … built with separate processes of Thinking for Speaking and thinking for not speaking," as even when they are not actively engaging in speech, speakers would need to 'encode' the necessary elements for later 'retelling' (p. 672). However, Slobin's argument is not so much that cognition is radically different when speaking and *not* speaking; instead, TFS proposes that even if one cannot guarantee that one's language *always* determines thought, it is likely that if some aspects of reality are always expected in the grammatical categories of a language, those aspects will be more salient in the speakers' mental imagery. Consequently, since grammatical

categories vary across languages, the aspects of conceptual domains that are highlighted might also vary for speakers of different languages.

Another significant challenge to linguistic relativity has been proving a 'causal relation'[3] between one specific behavior or cultural element and its linguistic representation. Even if one has established that "cognitive systems vary across two different cultures," a feat that is already hard in itself, it could be possible that "the difference in language use and cognitive patterns across the two cultures" are both driven by pre-existing cultural differences (Pederson, 2009, pp. 670–671). In this sense, cultural variation has been the suggested reason for differences in lexicon across cultures; for example, a fishing community may have needed more words for the different states of the ocean than a group living in-land and relying on hunting, leading to varying salience in those semantic domains in each of their languages' lexicon. Consequently,

> elaborated vocabulary sets in expert domains and impoverished sets where there is little experience and impoverished sets where there is little experience … are also not taken as particularly revealing. (Pederson, 2010, p. 1016)

If, on the other hand, the linguistic feature observed is "a long fossilized grammatical pattern," it would be more plausible that the linguistic variation predates the cultural one (Pederson, 2009, p. 671).

Attending to these aspects, conducting viable linguistic relativity research requires looking into aspects of language other than lexicon, such as syntax or morphology, which are more likely to predate cultural and behavioral differences. In addition, one must look into relatively universal domains of human life which do not rely on "environmentally or culturally variable experience" (Pederson, 2010, p. 1016). In response, TFS research looks at

1. a selection of languages and a semantic domain that is encoded with some frequency in all of the languages;
2. the semantic domain is encoded by special grammatical constructions or obligatory lexical selections in at least some of the languages under comparison;
3. the domain is relatively more codable in some of the languages to be compared;
4. a selection of discourse situations in which the semantic domain is regularly accessed. (Slobin, 2003, p. 3; emphasis removed)

[3] As described in Sapir (1921: 2018).

One of the key factors being the selection of a domain for observation that is both relatively universal and enshrined in the formal features of a language. To this avail, TFS has largely focused on the domain of Motion.

Why Motion? Motion is one of the 'culture-specific' "experiential domains" in which our knowledge of the world is organized, from the point of view of cognitive grammar and semantics (Croft, 2009, p. 164; Lakoff, 1987, p. 95). Within these, Langacker (1986) proposes the existence of 'basic domains' which are inherent to cognition, either due to being common to "all human environments" or belonging to the innate physical aspects of human experience (Pederson, 2010, p. 1018). For instance, Langacker names as basic domains "the experience of time and our capacity for dealing with two- and three-dimensional spatial configurations" as well as the ones "associated with the various senses: color space … the pitch scale … [and] a range of possible temperature sensations" (Langacker, 1986, p. 5).

Studies in linguistic relativity have traditionally focused on a small set of said "cognitively fundamental domains" (Pederson, 2009, p. 673), namely: *Color, Space, Time, Grammatical gender and number, Logic and arithmetic number, Emotion and personality,* and *Motion*. As a semantic domain, 'Motion' is concerned with 'motion events,'

> events of animate motion … [that] involve an animate being moving along a directed path, expressed linguistically using a verb of motion, a directional element, and possible mention of one or more associated ground elements (source medium, way-station, goal). (Slobin, 2005, p. 1)

Motion events can be defined as "perceptually universal," involving essential aspects of human experience like space or time; in addition, "they are complex enough to suggest a range of conceptualizations should be possible" in different languages (Pederson, 2009, p. 675), making the domain of Motion an ideal candidate for linguistic relativity research.

2.1.2. Typological Background

In order to study the ways in which different formal realizations of Motion may impact conceptualization, however, one must first demonstrate that motion events are verbalized differently from language to language, and classify said grammatical distinctions. For this, TFS largely draws from Leonard Talmy's typologies of motion events (1985, 1991,

2000a, 2000b, 2009) and their later variations, all based on the differences in lexicalization patterns and verbal framing between languages.

Lexicalization and Verb Framing. Semantic domains such as Motion are often based on the description of real-world events, meaning things that have occurred or are in the process of occurring (Rosen, 1999, p. 1). As real-world events are cognitively processed and verbalized, a process known as 'lexicalization' occurs, wherein the conceptual components that make up our perception of reality are mapped onto the 'lexical units' of a clause. When this process is undergone differently from language to language, it leads to diverse 'lexicalization patterns' (Levin & Rappaport Hovav, 2019, p. 395).

In order to be able to compare the mapping of concepts onto words and classify lexicalization patterns, one must look into events that contain identifiable conceptual components used across different languages (Levin & Rappaport Hovav, 2019, p. 396). For Talmy (1991), events can be classified into three types:

a. 'simplex events' are those that can be "expressed by a single clause" that cannot "be further partitioned with the resulting sub-portions also able to be cognized as events" (Talmy, 1991, p. 481), for instance: *The candle went out.*
b. 'complex events' are those that need to be expressed by "a complex sentence consisting of a main verb and a subordinate clause," which can be further subdivided into separate events. Such as: *The candle went out because something blew on it.* (Talmy, 1991, pp. 481–482).
c. 'macro events' are 'conceptually conflated' events. They are "understood as complex" and could be expressed through a "multi-clause syntactic structure," but are represented by a single clause, and "conceptualized as simplex." For example: *The candle blew out"* (Talmy, 1991, p. 481).

Talmy focuses on the latter, stating that within macro-events, one can find 'framing' or 'domain-schematizing events', simplex events that give shape to the "schematic structure in any of a particular set of organized conceptual domains" (Talmy, 1991, p. 482), determining how the events belonging to such conceptual domains are cognitively understood.

Each framing event must contain a number of conceptual elements, namely:

a. a '**figural entity**' which moves or stays put in relation to a series of
b. '**ground elements**,' for instance, a physical space wherein the event takes place;
c. an '**activating process**' of the figural entity (e.g., movement); and a
d. '**relating function**' through which the event occurs, most often a Path. (Talmy, 1991, pp. 482–483).

Any additional conceptual elements can be added to the macro-event by adding 'supporting events' to the framing event through a supporting or 'S-relation' such as '**Cause'** or '**Manner**.' (Talmy, 1991, pp. 482–483).

In addition to the framing event, Talmy (2000b) suggests the existence of a 'Co-event' which relates to the framing event within the macro-event. These Co-events may also "relate to a Motion event as its **Manner** or **Cause**" (pp. 8–9) or other S-relations like

a. '**Precursion**,' occurring when the Co-event goes before the Motion event but does not cause it, as in: "[glass MOVED onto the carpet]$_{\text{framing event}}$ WITH-THE-PRECURSION-OF [the glass splintered]$_{\text{co-event}}$" in the Motion event "glass *splintered* onto the table"[4] (p. 42; edited).
b. '**Enablement**,' in which the Co-event also goes before the Motion event, does not cause it, but "enables the occurrence of an event that causes the Motion," as in: "[I MOVED jellybeans into her sack]$_{\text{framing event}}$ WITH-THE-ENABLEMENT OF [I scooped up the jellybeans]$_{\text{co-event}}$" in the Motion event *"*I *scooped* jellybeans up into her sack" (p.44; edited).
c. '**Reverse enablement**,' the Co-event gets 'undone' in order to make the Motion event possible, as in "[I MOVED the dog TO FREENESS]$_{\text{framing event}}$ WITH-THE-ENABLING-REVERSAL-OF [(someone) had chained the dog]$_{\text{co-event}}$" (p. 44; edited) in the Motion event 'I *unchained* the dog.'
d. '**Concomitance**' wherein the Co-event "co-occurs with the main Motion event and is an activity that the Figure … additionally exhibits" but does not impact the Motion event, as in "[she WENT to the party]$_{\text{framing event}}$ WITH-THE-CONCOMITANCE-OF [she wore a green dress]$_{\text{co-event}}$" in the event "She wore a green dress to the party" (Talmy, 2000b, p. 46; edited).

[4] Henceforth: main verb in italics, satellite underlined.

e. '**Concurrent Result**,' in which the Motion event causes a Co-event that occurs at the same time, as in "[the rocket MOVED into the water]$_{\text{framing event}}$ WITH-THE-CONCURRENT-RESULT-OF [the water splashed]$_{\text{co-event}}$" in "the rocket *splashed* into the water" (p. 47; edited).
f. '**Subsequence**,' when the Co-event goes after the Motion event, and the Motion event either 'enables/causes/is the reason behind' the Co-event; for instance in "[they MOVED the prisoner into his cell]$_{\text{framing event}}$ WITH-THE-SUBSEQUENCE-OF [they locked the cell]$_{\text{co-event}}$" in "they *locked* the prisoner into his cell" (p. 47; edited).

Any of these S-relations may be present as conceptual elements conflated with Motion in any given Motion event, and will be given similar treatment than elements such as Path, Manner, Ground, or Figure.

As they all share the same key conceptual elements, and are present in most languages, framing events are key to the study of lexicalization patterns. Talmy defines five types of framing events, from temporal events to changes in state, but focuses his typology on events of Motion or Location (i.e., movement or the lack thereof). When applying the conceptual elements common to all framing events to events of motion, the resulting four elements are:

a. **Figure**: the entity that is moving or located
b. **Ground**: the entity which acts as a spatial reference point for the motion/location of the figure
c. **Path**: the path of motion of the figure
d. **Manner**: the manner of motion by which the figure moves along the path[.] (Croft et al., 2010, p. 2; edited)

In addition to these, 'fact-of-motion' or Motion is the conceptual element that conflates with one of the above in the main verb, depending on the pattern (Talmy, 1985, p. 71).

In his 'motion-actuating' and 'motion framing' typologies, Talmy classifies languages according to the various ways of mapping these conceptual elements onto verbal components through lexicalization for the purpose of expressing motion.

Motion Actuating Typology. Talmy's earlier 'motion actuating typology' (Talmy, 1985b as cited in, 2009, p. 1) classifies languages according to how they encode the conceptual elements of framing events through verbal elements, leading to different

lexicalization patterns. The author focuses on verb framing; that is, he classifies languages depending on which of the conceptual elements of a Motion event (e.g., Path) is realized through the main verb of the clause, and which is included through an S-relation or satellite; satellites being all "immediate constituents" of the verb phrase which are not "inflections, auxiliaries or nominal arguments," either a 'bound affix' or a 'free word;' and which relate to the main verb as "modifiers to a head" (Talmy, 1985, p. 102, 2000b, p. 222) Satellites are generally understood in opposition to both verbs and nominal/prepositional arguments of the main verb. Depending on what elements are mapped onto satellites as opposed to the verb root in different languages, varying lexicalization patterns arise.

In order for a pattern to be classified as widespread, Talmy states that it must be frequent "in occurrence in speech," present in everyday speech, and "pervasive" or able to express "a wide range of semantic notions" (Talmy, 1985, p. 62). He defines three major lexicalization patterns:

Motion + Manner/Cause. Languages of this type, such as English, are rich in motion verbs in which the salient conceptual elements are Manner or Cause (Talmy, 1985, p. 61). As the main verb expresses both Motion and Manner/Cause, the Path is indicated through a satellite. For example, in the clause "the rock *slid/rolled/bounced* down$_S$ the hill" (Talmy, 1985, p. 62; edited), the main verb (in italics) includes information about the Manner in which the rock moves, but the Path is included as a satellite (underlined).

Motion + Path. In languages where this lexicalization pattern is pervasive, such as Spanish, the main verb contains the Motion and the Path, while the Manner/Cause is included as an "adverbial or gerundive type constituent" (Talmy, 1985, pp. 68–69). In cases in which such a constituent does not fit the rhetorical style of the language, the Manner/Cause is "either established in the surrounding discourse or omitted" (p. 69). For example, in the Spanish clause 'la roca *bajó* la colina rodando$_{GER}$,' translated to 'the rock moved-down the hill rolling,' the main verb includes the path of the motion (i.e., down a hill), and the manner is added through a gerundive.

Motion + Figure. Languages belonging to this third lexicalization pattern, such as Navajo, mostly express Motion alongside Figure (Talmy, 1985, p. 61). English has a few verbs fitting into this pattern; for instance, in the sentence 'he *spit* violently at the wall,' the main verb

expresses both Motion and Figure, while Manner (i.e., 'violently') and Path are both expressed separately.

Variations. The three lexicalization patterns outlined by Talmy make up a typology of the existing 'core systems' of lexicalization in language; however, conflations of motion with unusual conceptual elements in the main verb, such as Motion + Ground or Motion do occur sporadically (Talmy, 1985, p. 75), and even conflations of more than two elements, for instance, Motion + Ground + Path, as in "I *shelved* the book;" 'multi-component conflation systems' "never form a language's major system for expression Motion," though (Talmy, 1985, p. 76)**.**

In addition to the lexicalization patterns outlined above, a Co-event might also be incorporated into the Motion event through relations such as Cause or Manner, encoding conceptual elements like Cause or Concurrent Result.

Motion-Framing Typology. In his 'motion-framing typology,' Talmy (1991, p. 515, as cited in, 2000b, 2009, p. 2) focuses on the 'core schema' of the macro-event, classifying languages in a binary typology according to "whether the core schema is expressed by the main verb or by the satellite" in a language's primary lexicalization pattern (Talmy, 2000b, p. 221). In motion events, the core-schema is often correlated with the Path; consequently, Talmy steps away from other conceptual elements introduced in his prior typologies. In the resulting typology,

> languages that characteristically expressed Path in the main verb (root) were called "verb-framed," and languages that did so in the satellite were called "satellite-framed." (Talmy, 2009, p. 2)

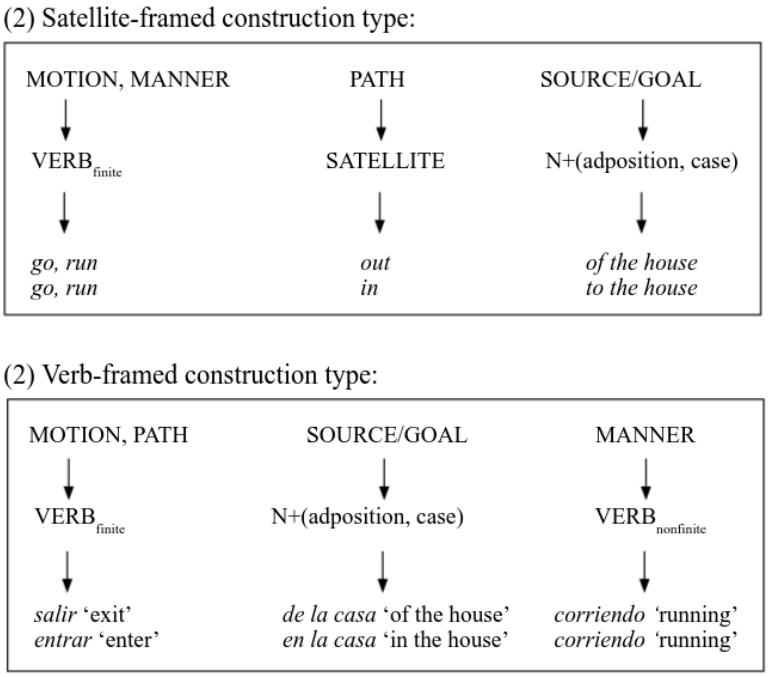
In order to illustrate this phenomenon, Talmy (2000b) compares English, a 'satellite-framed' (S-framed) language, and Spanish, a 'verb-framed' (V-framed) language:

> **English** - Satellite-Framing
>
> "The bottle *floated*$_{\text{MAIN VERB → MANNER}}$ <u>out</u>"$_{\text{SATELLITE → PATH (CORE SCHEMA)}}$
>
> **Spanish** - Verb-Framing
>
> "La botella *salió*$_{\text{MAIN VERB → PATH (CORE SCHEMA)}}$ <u>flotando</u> $_{\text{NON-FINITE VERB, GERUNDIVE → MANNER.}}$

TFS largely uses this classification as a typological background. Slobin (2000, pp. 111–112) suggests that in S-framed languages Manner is encoded in the main verb and Path in a satellite, whereas in V-framed languages the main verb contains Motion+Path and Manner is added through a satellite, as suggested in Figure 2..

Figure 2

Slobin's Approach to Motion Typology



Note. Adapted from Slobin, D. (2000). Verbalized events: A dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Mouton, p. 109.

In addition to verb framing, Slobin refers to two additional typological differences between S-languages and V-languages. Firstly 'salience,' introduced by Talmy (1985, p. 122), refers to the degree to which a component of meaning, due to its linguistic representation, "emerges into the foreground of attention or … forms part of the semantic background where it attracts little attention."

Secondly, Slobin proposes an additional dimension to Talmy's typology, proposing that V-languages only employ a Manner verb as the main verb in a 'path expression' when "no boundary-crossing is predicated" (Slobin, 2004, p. 7). For instance, verbs like "*enter, exit, cross"* that indicate changes of state are 'boundary-crossing events", as in[5]

[5] Examples from Slobin (2004, p. 7).

"Volar al árbol" ('*fly to the tree*')	**No boundary crossed**.
"Fly out of the hole"	**Boundary crossed** (in the hole/out of the hole); not possible in V-languages.

Furthemore, in V-languages a 'subordinate construction' is needed in order to express Manner, for instance, "exit the hole flying" (*'salir del agujero volando'*) (p. 7), which may lead to Manner not being 'economical' in V-languages (Slobin, 2003, p. 5)**,** and lead to V-language speakers avoiding the expression of Manner.

Whether the typological differences between S-languages like English and V-languages like Spanish result in differences in conceptualization is the main object of study of TFS research.

2.1.3. Motion in Thinking for Speaking

Slobin's Thinking for Speaking. As outlined in the previous section, Slobin (1987, 1996b, 1996a, 2000, 2003, 2004)[6] takes Talmy's framework and proposes that different languages' prevalent lexicalization patterns when describing motion events will result in distinct conceptualizations of reality; following his thesis that the language we use when communicating directly determines the simultaneous 'online' mental processes that take place. Although TFS research focuses mainly on verb framing, Slobin looks at additional differences between languages concerning 'discourse structures,' rhetorical style, and temporal description as reflected on verbal aspects (Slobin, 1996a, pp. 77–78, 2004, p. 2). The author argues that motion events exist simultaneously in a 'discourse frame,' meaning a narrated account of a 'journey,' and in a 'typological frame' wherein lie the grammatical limitations of the language (e.g. tense, aspect, and framing); in order to completely "characterize the linguistic coding of [motion] events, then, one must attend to usage, as constrained to typology" (Slobin, 1996b, p. 195).

Drawing from the typological differences outlined in Talmy's framework (see Section 2.1.2), Slobin proposes that the main difference between S-languages and V-languages is that, by encoding Manner in the main verb, S-languages "allow for an economical expression of manner of motion" and that

[6] Reviewed for this dissertation.

> as a consequence, these languages make habitual use of manner verbs when encoding motion events, and have developed large lexicons with many fine grained distinctions of manner, in comparison with smaller and less differentiated manner lexicons in V-languages. (Slobin, 2003, p. 5)

The author draws the resulting hypothesis that manner of motion will be a prevalent domain, with Manner verbs occurring often across discourse, and "speakers will have rich mental imagery of manner of motion" (p. 5).

To test this hypothesis, much of TFS Motion research uses the wordless cartoon book *Frog, Where are you?* (2003), depicting different animals involved in motion events (see Figure 3). Studies ask speakers of different languages to describe how the characters move in the storybook, commonly referred to as 'The Frog Story;' the resulting descriptions are known as 'Frog Narrations' (Slobin, 1996a, 2000, 2004).

Figure 3

Illustration of Motion Events in a Frog Story

Note. Taken from Slobin, D. I. (1996). From "Thought and Language" to "Thinking for Speaking." In J. Gumperz & S. Levinson (Eds.), *Rethinking Linguistic Relativity* (pp. 70–96). Cambridge University Press, pp. 72–73.

In addition to frog narrations, Slobin (1996a, 1996b, 2000) looks into literary, conversation, and translation corpora comparing S-languages and V-languages. All in all, the author finds that in frog narrations,

a. speakers of Spanish (V-language) show less motion verb variety than speakers of English (S-language) (Slobin, 1996b, pp. 199–204, 2003, pp. 7–8);

b. S-language speakers like English use more detailed motion verbs in frog narrations than V-language speakers, as reflected in increased use of phrasal verbs or 'Verb+Satellite constructions' such as "*pop+out, up"* to reflect Manner (Slobin, 1996b, pp. 199–204);
c. Path verbs are common in both S-languages and V-languages, but Manner verbs are increasingly present in S-languages, despite the verbs existing in V-languages as well (Slobin, 2000, p. 113, 2003, p. 3);
d. English speakers make reference to Ground elements more than Spanish speakers (Slobin, 1996b, pp. 199–204);
e. Spanish speakers "describe fewer segments of a journey" (Slobin, 1996b, pp. 199–204).

These results are substantiated by looking at literary and conversation corpora. Research involving novels is particularly interesting, as frog stories generally restrict the resulting narrations, but in written fiction, authors are free to focus on whatever 'semantic domains' they may consider relevant (Slobin, 1996b, p. 209). The research finds that

a. motion event descriptions are overall more frequent in English novels than Spanish (Slobin, 1996b, p. 207);
b. novels in S-languages show a greater variety of manner verbs (Slobin, 2000, pp. 113, 120, 2003, p. 8), English novels showing "greater lexical diversity" overall when compared to Spanish (Slobin, 1996b, p. 208);
c. S-languages use phrasal verbs to indicate Manner commonly, whereas they are rare in V-languages (Slobin, 2000, pp. 116–117);
d. authors of V-language novels don't use other ways of indicating Manner as often as expected in order to compensate for lesser variety of Manner verbs, for instance 'adverbs of manner' or "descriptions of motor behavior and body condition (*not looking where he went* …); descriptions of inner states (*agitated* …); descriptions of environmental conditions that affect manner of movement (*the snow was thick; the road was muddy*) (Slobin, 2003, p. 8);
e. English writers mention Ground more often in relation to the Path, and pay greater attention to the details in the Path (Slobin, 1996b, p. 208)**.**

Finally[7], Slobin also looks at translations, proposing that "it is necessary to go from the mental image produced by reading the source language to an appropriate rendering of that mental image in the target language" while using the formal elements available in both the source and target language (Slobin, 2000, p. 123), which may reflect differences in conceptualization. He finds that

a. in translations from S-language to S-language, Manner verbs are "retained," whereas in translations from S-language to V-language Manner verbs are "neutralized, diminished, or omitted (Slobin, 2000, p. 124);
b. V-language to S-language translations substitute 'simple path verbs[s]" for more expressive manner verbs (Slobin, 2000, p. 125);
c. when translating from Spanish to English, translators tend to add detail both concerning both Path and Manner (Slobin, 1996b, pp. 210–214);
d. translating from English to Spanish, there is a reduction of detail, with information being omitted unless it is extremely important for the narrative (Slobin, 1996b, pp. 210–214).

TFS further substantiates the impact of these results on 'mental imagery' by studying language reception, or 'Reading/Listening for Imaging.' For instance, Slobin (2000) asks speakers to English and Spanish to report 'mental imagery' of Spanish novel excerpts and their literal translations in English; in spite of Manner verbs being largely absent from the texts, most English speakers included Manner in their descriptions, but only a 'handful' of Spanish speakers did (Slobin, 2000 as cited in 2003, pp. 11–12). It appears that since S-language speakers are "habitually exposed to more elaborate and vivid descriptions of motion" their mental image of 'described events' "contains more information about manners of movement" than V-language speakers (p. 12).

Slobin's results seem to at least partially substantiate his claim that language influences the way we think and the way we conceptualize events around us. The research presented above shows that typological contrasts between S-languages like English and V-languages like Spanish—as first outlined by Talmy—result in differences in discourse; such as higher detail paid to motion overall and increased use of Manner verbs in S-languages,

[7] Research pertaining to language acquisition has not been analyzed in-depth for the purpose of dissertation, although some examples can be found in Slobin (2000, p. 121, 2003, pp. 9–11).

among others. These differences may reflect cognitive effects, such as Manner playing a larger role in S-language speakers' mental images of reality.

Slobin proposes that "language plays a role in organizing conceptual space for the purposes of thinking for speaking," that is, that the actual differences between motion events in objective reality "exist in a graded and multidimensional psychological space, including factors of force dynamics, muscular pattern, rate, and social-emotional evaluation" and that S-languages, due to their typological characteristics, have divided this spectrum into more "lexical concepts" such as their wide Manner verb lexicon, whereas V-languages have not done so (Slobin, 2000, p. 122). From his studies, Slobin infers that when describing Motion, V-language speakers conceptualize Manner as separate from Motion, whereas S-language speakers process Motion+Manner as "a single conceptual event" and a resulting joint mental image (Slobin, 2003, p. 11).

Further Studies. Slobin's results have been substantiated by a number of studies by different authors. As the present dissertation focuses on the differences between English and Spanish, the articles presented here will follow the same focus.

Naigles et.al. (1998) study descriptions of pictures and videos depicting motion events by English and Spanish speakers; finding that English showed more variety of Motion verbs and that English speakers preferred Manner verbs over Path verbs, whereas Spanish speakers used Manner and Path verbs in more or less 'equal proportions.' When encountering boundary-crossing events speakers of both languages preferred Manner verbs. Spanish speakers also used 'bare'[8] verbs more frequently (e.g. 'they *fell*').

Similarly, Ibarretxe-Antuñano (2012, pp. 11–25) compares frog stories in English, Spanish and Basque, looking at (a) verb variety through verb type frequency; (b) among of detail dedicated to Ground; and (c) how much 'narrative attention' is devoted to Motion versus 'scene-setting' (e.g. characteristics of the surrounding space); finding that

a. English has greater verb type variety (58) compared to Spanish (27);
b. Spanish generally limits journey descriptions to 'one-ground-element-per-verb' whereas English mentions several elements;

[8] "with no elaboration of path beyond the inherent directionality of the verb itself' (Slobin, 1996a, p. 200, as cited in Cifuentes-Férez, 2008, p. 61)

c. English dedicates less narrative attention to scene-setting, letting locations be 'inferred' from the 'trajectory' of the Motion in several cases.

Cifuentes-Férez & Gentner (2006) conduct research around English and Spanish speakers' interpretations of novel verbs. The authors asked speakers to define novel verbs like "*ransined*," which were presented alongside both Path and Manner-related information, so the interpretations could go in either direction. The authors actively go against Slobin's hypothesis, giving Spanish speakers "frames that included a preposition, thus encouraging interpretation in terms of manner verb" and English speakers "bare transitive frames—which would normally be likely to invoke a path interpretation" (p.453); still, Spanish speakers' interpretations included a higher number of Path verbs than English speakers, and English speakers used more Manner verbs to define the novel verbs.

Research extracting conclusions on conceptualization merely from the linguistic attention paid to one domain or another generally substantiates Slobin's conclusions; however, studies focusing on cognitive aspects and involving non-linguistic tasks show more conflictive results.

Kersten et.al. (2010, pp. 644–646) use a 'category-learning task' in which speakers of English and Spanish had to classify animated (imagined) bugs, presented to them in pairs through animated videos. In each video, the path and manner of motion of each creature varied, as well as other factors like the background and the general appearance of the bugs; participants had to categorize the videos according to the species of the creatures, which sometimes depended on their trajectory (Path) and sometimes on the way they moved (Manner). Although speakers of both languages performed similarly when distinguishing the bugs according to their trajectory, "monolingual English speakers were more likely than monolingual Spanish speakers to notice the relevance of manner of motion," performing better in the manner-of-motion condition for differentiating the species (p. 645). These results suggest that the salience of Manner in English might lead English speakers to pay more attention to the domain in everyday life.

Hohenstein (2005) studied the impact of language on English and Spanish-speaking on 7-year-old children when subjected to 'video stimuli' through a 'match-to-sample task.' The participants first viewed a 'target video' showing a motion event and were asked to select similar videos; with English-speaking participants choosing videos with the same manner of

motion as the target video more often than Spanish participants. However, when performing a subsequent experiment that attached novel-verbs to the videos syntactically implying either Manner or Path (e.g. presenting the novel verb with a Manner satellite to imply Path), both English and Spanish-speaking participants fixated on Manner more often when the verb was syntactically implied to contain Manner no language effects were found for Manner or Path preference in the interpretations.

Gennari et.al. (2002) perform 'recognition memory' and 'similarity judgment' tasks with 46 English speakers and 46 Spanish speakers. Participants looked at 108 motion events—recorded in video—organized as triads of "36 targets and 72 alternates, two for each of the target events" (p. 60); the target video contained a motion event and the alternates showed differences concerning Manner of Path. The speakers had to shortly describe the target events, perform quick recognition tasks and judge the similarity of the events by quickly rating them; with the similarity task including a variation with 'linguistic encoding'. The linguistic tasks (descriptions) showed results consistent with Talmy's typology and Slobin's research, with English speakers preferring Path verbs and Spanish speakers prefer Manner verbs. However, in the non-linguistic Spanish participants only selected the 'same-path alternate' to the target more often when exposed to the verbal encoding condition, with English and Spanish participants performing similarly without the linguistic condition. These results imply that the differences might have stemmed from the increased 'weight' of Path in Spanish due to typological differences between the languages (p. 76), and not from general differences in conceptualization, as otherwise, the preference should have shown without the linguistic encoding. Therefore, the results of the non-linguistic tasks are less promising than the ones involving verbal descriptions of motion events concerning Slobin's hypotheses related to linguistic relativity.

The lesser frequency of research involving non-linguistic experimental evidence, as well as the conflicting results of the existing studies, have been the focus of much of the criticism received by Talmy and Slobin's framework, outlined in the following section.

2.1.4. Limitations and Criticism

Before delving into the research conducted by the author of this dissertation, it is necessary to mention the pitfalls and theoretical gaps in TFS research to date, which will serve as inspiration for the research and methodology questions tackled in the present study.

To begin with, Talmy's binary typology has received extensive criticism for 'leaving out' languages that fall outside of the binary typology; as well as neglecting intra-linguistic variation. (DeLancey, 1989; Schultze-Berndt, 2000; Slobin & Hoiting, 1994; Zlatev & Yangklang, 2004; as cited in Slobin, 2004). Croft et.al. (2010) and Berman & Slobin (2013) point out that the lexicalization patterns outlined by Talmy that serve as the typological framework for TFS are designed to reflect generalizations and not serve as complete characterizations of any given language. For instance, Talmy (1985) acknowledges that "English does have a certain number of verbs that genuinely incorporate Path … for example: *enter, exit, pass, rise, descend, return, circle, cross, separate, join*", and that these verbs also show "a Spanish-type pattern for the rest of the sentence" with Manner being "expressed in a separate constituent as in *The rock passed by our tent*," and presents the possible explanation that most such verbs "are not original English forms but rather borrowings from Romance [languages]" (p. 72); from this point of view, it would be expected to observe an effect of etymology on the lexicalization patterns of borrowings from Romance languages in English, although research in this regard is surprisingly lacking. As an example, Martínez-Vázquez (2013) briefly looks at the French borrowing *enter*, finding it to be common in English and not restricted to formal contexts (p. 146). Taking this into account, the study conducted later in this dissertation will attempt to incorporate etymology as a possible indicator of divergence from the expected lexicalization patterns of English and Spanish.

Thirdly, most of the research in the reviewed literature (see the previous section) employs frog narrations—collected from a limited number of participants—or smaller corpora of literary novels and translations. Crosslinguistic research studying everyday language with large collections of corpora seems to be uncommon. Reshöft (2011) uses a corpus-based methodology to compare English to Spanish, French, and Italian but employs smaller learner corpora, of 'about 200,000 words.' In contrast, Martínez-Vázquez (2013) compares the Corpus of Contemporary American English (COCA) to the CREA (Reference Corpus of Current Spanish) but develops a rather qualitative analysis, analyzing specific examples from the corpora. The present study will attempt to apply Slobin's TFS framework to large collections of corpora in English and Spanish, with the objective of assessing whether the author's generalizations are substantiated by everyday speech.

Lastly, TFS studies have also attracted controversy for prioritizing the 'speaking' over the 'thinking,' that is, not paying enough attention to differences in schematization, merely

tracing crosslinguistic differences in basic domain verbalization from language to language (Athanasopoulos & Bylund, 2013; Wessel-Tolvig & Paggio, 2016). So far, TFS research has focused mainly on the "crosslinguistic differences in the organization of information" between languages (Athanasopoulos & Bylund, 2013, p. 91). The part of Slobin's theory stating that speakers will naturally focus on whatever aspects of reality that are the most salient in their native language when describing events can be studied through linguistic tasks, and research in this regard has been productive (see Section 2.1.3), but there is an apparent lack of "studies that actually examine the cognitive aspects of speech production" (Athanasopoulos & Bylund, 2013, p. 91). To study the kind of 'on-line' thinking and schematization strategies that take place before or during active speech, it would be necessary to conduct research that prioritizes cognitive processes. Taking from psycholinguistic research, the study of onset time and reaction times, as well as behaviors occurring alongside verbal communication (e.g., gestures) are suggested as possible methodologies (Anasthopoulous & Bylund, 2013, pp. 93–94). In particular, research concerning co-speech gestures in relation to cognition has been fruitful and will be an accompanying methodology to TFS for the present dissertation, as outlined in Chapter 3.

2.2. Thinking for Gesturing

A large body of theory has been devoted to the study of gestures as relevant cognitive mechanisms. McNeill (1992, 2000, 2008) points out the importance of everyday gestures as meaning-bearing symbols that "fuel thought and speech" rather than being separate, co-existing entities; in this sense "gestures, language, and thought are seen as different sides of a single mental/brain/action process," being "active participants" in both speech and thought (McNeill, 2008, p. 3). McNeill (2008) interprets gestures, alongside their 'co-occurring speech' "as the embodiment of an individual's thinking at a specific moment," as gestures often lack any type of "conscious communicative effort" and yet seem to represent aspects of the event being described as conceptualized by the speaker (pp. 89-90). This framework is somewhat analogous to TFS. Where Slobin states that language and thought are inseparable, McNeill (2008) believes that "language is inseparable from imagery," and imagery "is embodied in the gestures that universally and automatically occur with speech," and which are "a necessary component of speaking and thinking" (p. 15). Therefore, the study of gestures can serve as an additional model for tackling research concerning motion events.

A categorization of gestures is proposed by McNeill (2008), drawing from Levy & McNeill (1992) and Duncan et.al., (1995). For the present study, we will focus on 'iconic gestures' namely those whose form and meaning are akin to the event being described in some way, co-occurring with the spoken description but not necessarily communicating the same aspects (McNeill, 2008, p. 39).

Figure 4

Illustration of an iconic gesture

Note. The Speaker lifts his hand "with the index finger extended, depicting the character rising and possibly the interiority of the pipe," while saying "*he tries going [up the inside] of the drainpipe."* Taken from McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. London: University of Chicago Press, p. 106.

Since iconic gestures encode an aspect of the motion event, they are likely candidates to show conceptual elements of motion events, such as Path or Manner.

2.2.1. Gesture and Motion

Gestures of Path. McNeill (2008) follows Talmy and Slobin in assuming that during communication, gestures "might differ across languages in predictable ways" (p. 195). Concerning motion events, S-languages largely express Manner through the main verb and Path through several satellites (many involving Ground elements), which results in S-language speakers producing one gesture per satellite, leading to an increased number of gestures.

Figure 5

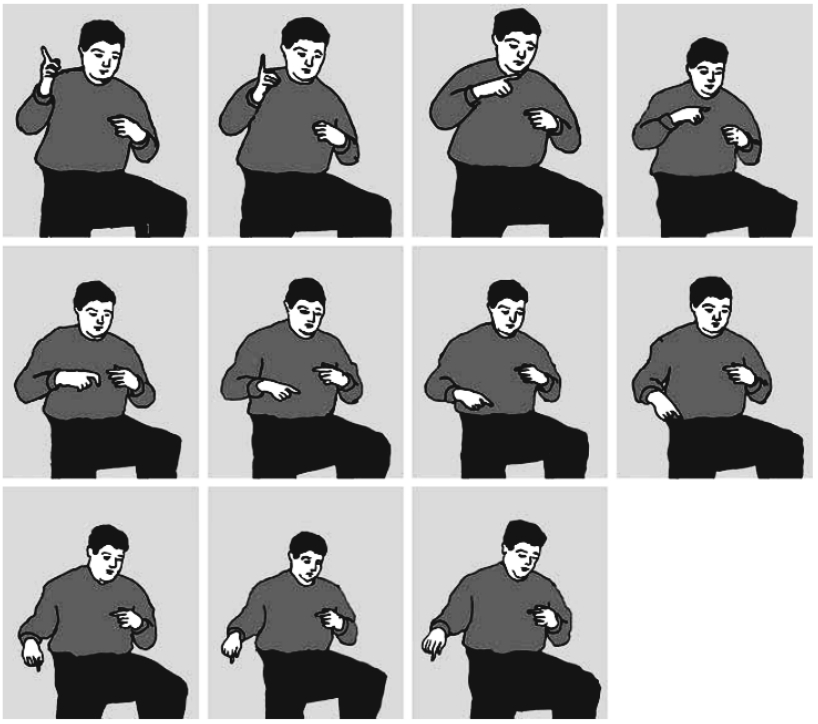
Gestures for Path in an English (S-language) speaker

Note. Speaker is describing the motion event "*and it rolls down but it rolls him out down the rain spout, out into the sidewalk into a bowling alley*." Adapted from McNeill, D. (2008). *Gesture and Thought.* London: University of Chicago Press, p. 197.

In contrast, V-languages pay less attention to the detail of motion events (see Section 2.1.3) and are able to encode Path in the main verb, leading to fewer gestures.

Figure 6

Gesture for Path in a Spanish (V-language) speaker



Note. The speaker produces a "single curvaceous gesture" to describe the same motion event presented in Figure 5, which was simply described as '*entonces se cae'* ("then he falls"). Adapted from McNeill, D. (2008). *Gesture and Thought.* London: University of Chicago Press, pp. 198–199.

Therefore, the number of gestures produced in relation to Path can be indicative of the typological differences between S- and V-languages impacting not only speech but also gesture, which may support further effects on schematization. The resulting theory would be that speakers of S-languages conceptualize Path as segmentable and substantially detailed, whereas V-language speakers pay less attention to the individual aspects of Path (e.g., Ground elements) and focus on trajectory. However, McNeill (1992) goes further in affirming that gestures can also serve as indicators of salience, that is, looking at gestures can provide supporting information on whatever aspects are highlighted in a particular speaker's mental image of reality.

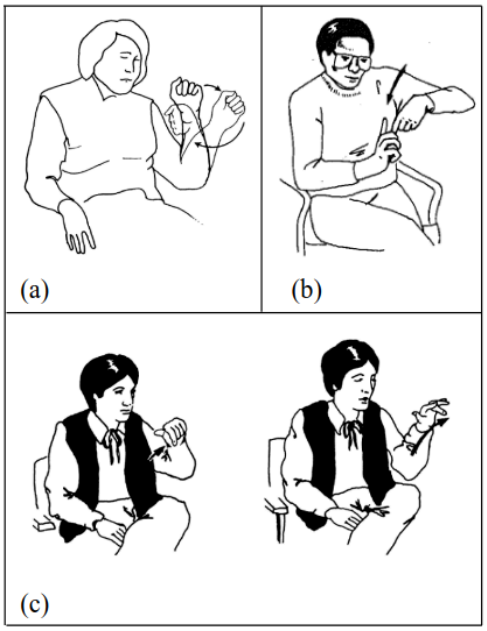
Gestures as Indicators of Salience. As speakers use more complex gestures when talking about an aspect of reality 'salient' to them, gestures can provide information on "what is important, relevant or salient to the speaker" when narrating an event; moreover, gestures might be 'better' than speech at indicating salience, as they lack grammatical 'obligatory elements' and are not bound by concrete rules (McNeill, 1992, pp. 125–128). The 'complexity' of co-speech gestures can be used as a measure of a particular conceptual element's (e.g., Manner) salience for a specific speaker when describing an event. For instance, the author finds that the verb '*to climb*' can be accompanied with (a) "a simple upward gesture" and (b) a "gesture with wiggling fingers … depicting the character rising and clambering movements" by two different speakers, showing a greater salience of Path for the first and Manner for the second speaker (p. 125). If these patterns are generalized, then one could hypothesize that they reflect typological effects on conceptualization.

Kita (1990, as cited in McNeill 1992, p. 126) provides the following method for measuring how complex gestures are, suggesting to add a point for each:

"use of two arms" (Figure 7-b)
"movement of the fingers" (Figure 7-c)
"change of hand shape during the stroke" (Figure 7-c)
hand position "other than a fist or open hand" (Figure 7-b)
hand in position "other than rest" (Kita 1990, as cited in McNeill, 1992, p. 126).

Figure 7

Examples of co-speech gestures



Note. Examples adapted from McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. London: University of Chicago Press.

2.2.2. Gesture Tagging

To conduct gesture-based research, researchers must first select a method for gathering and tagging gestures[9]. Levy & McNeill (1982, as cited in McNeill, 2008) suggest gathering gesture-data through elicited storytelling, in order to avoid interpreting gestures exclusively in relation to the co-occurring speech, comparing them instead to the source material (e.g., a cartoon). One such stimulus is the Frog Story, used extensively in TFS research. To tag gestures, McNeill's framework (2008, pp. 259–288) uses accompanying speech as the guide, transcribing speech for each speaker and then 'locating, identifying and coding' gestures, including gesture phases. The author implements an elaborate notation system. For the present dissertation, the relevant elements (McNeill, 2008, p. 262) are:

a. "[s]quare brackets show the gesture phrase relative to speech;"
b. non-stroke 'holds'[10] are underlined;
c. "[b]oldface shows the stroke phase[11]—the phase with semantic content and the quality of 'effort;'"

[9] Overview in Bressem (2013)
[10] Where the speaker begins the gesture but does not fully realize it yet.
[11] The gesture is carried out.

d. “self-interruption of speech” is marked with an asterisk;
e. “one or more slashes shows silent pauses;”
f. “audible breath pause[s]” are marked with #;
g. “filled pause[s] are marked in-between ‘< >’ symbols.

Figure 8

Illustration of gesture tagging

Note. Gesture transcription is “so he gets a / hold of a big [oak tree / and he **bends it way back**]. Adapted from McNeill, D. (2008). Gesture and Thought. In Gesture and Thought. London: University of Chicago Press, p. 30.

2.2.3. Previous Research

A number of studies focus on the role of gesture for conceptualization during active speech. Alibali et.al. (2000; M. W. Alibali et al., 2001) and Alibali & Kita (2010) compare children’s linguistic productions when gestures are allowed versus when they are not, finding that “children who were prohibited from gesturing focused more on non-perceptual information … than children who were allowed to gesture” showing that “spontaneous gestures … may help speakers decide what to say” (Alibali et.al., 2001, p. 408). These results are promising in the sense that gesturing might not only be a tool for observing whether language impacts conceptualization, but that gestures on their own might impact mental processes as well, and should be included in TFS approaches.

A significant amount of research studying gestures in relation to speech and cognition looks at elicited narrations of motion events (Duncan, 2001; McNeill & Duncan, 2000; Parrill, 2011; Stam, 2006). Stam (2006) finds that Spanish speakers often use Path gestures with the main verb–usually also encoding Path–, and do not accumulate them, whereas in English speakers Path gestures both co-occur with the main verb and with satellites. This probably stems from typological differences, as Spanish cannot 'stack' satellites. McNeill & Duncan (2000) note two "alternating patterns" in English, one which focuses on the manner of motion by using Manner gestures with the main verb, and an opposite which uses Path gestures to 'downplay' manner (p. 150); in contrast, Spanish shows several instances of Manner encoded in a gesture without accompanying speech–which the authors define as 'manner fog'–and Manner gestures appearing with speech segments focused on other elements like Path. McNeill & Duncan (2000) find that Manner is likely to appear in Spanish when it is a focused component, or unusual in some way. In contrast to TFS-based hypotheses, Duncan's (2001) study shows that speakers of English and Spanish produce a similar number of Manner gestures, lacking significant differences in conceptualization; Duncan also finds that Spanish speakers are more likely to gesture for unusual manners of motion, and considers a possible 'compensation hypothesis' wherein Spanish speakers might use Manner gestures to compensate for lacks in speech. Research concerning gesture complexity and overall frequency of gestures in relation to English and Spanish' differing lexicalization patterns is infrequent.

These findings will partially serve as inspiration for the gesture experiment conducted later in this dissertation, which will attempt to provide supporting evidence, as well as implement some novel methodologies.

3. Data and Methods

Chapter 3 offers an overview of the data collection and analysis methodologies applied in the case study. The present dissertation employs a crosslinguistic, corpus and experiment-based typological methodology, following Talmy's typologies and Slobin's TFS framework, as well as McNeill's approach to gestures (see Chapter 2), in order to establish whether different languages lead to differences in motion event conceptualization. English and Spanish are used as examples of satellite-framed and verb-framed languages respectively. Additionally, the more specific questions to be studied arise from the criticism and limitations outlined in Chapter 2, such as:

a. whether the generalizations drawn by TFS (e.g., increased attention paid to Manner in S-languages) are substantiated in everyday speech and reflected in frequency variation in everyday speech, as shown through large corpora;
b. what is the relationship between etymological background and lexicalization pattern in borrowings from Romance languages in English, and whether an impact is observed in written versus spoken language; and
c. can one find substantial evidence of the 'thinking' in Thinking for Speaking, through a study of co-speech gestures when describing Motion in English and Spanish? Do the expected differences in detail and attention paid to Path/Manner outlined in the literature hold up in gesture analyses derived from frog narrations?

In order to answer, the approach is twofold. First, a corpus study is conducted, conducting large scale frequency searches on corpora of English and Spanish. Secondly, an elicited narration experiment is carried out, with a focus both on the resulting speech and gestures. The remainder of the present chapter explains the data collection and analysis process for both.

3.1. Corpus Study

Four corpora are employed for the corpus study, two for each language. This is done with the intention of having more frequency data, as the motion verbs studied will necessarily be restricted by the lexical limitations of each language. For English, the large Corpus of Contemporary American English (COCA) by Mark Davies (2013)**,** and the British National Corpus (BNC) in its web version (BNC-web, 2018; Consortium, 2007) were chosen. The COCA spans from 1990 to 2019, and contains a total of 1 billion words, compared to 100

million words and a span from 1975 to the present in the BCN; but the BNC has the added benefit of providing spoken data. For Spanish, Mark Davies' (2002) Corpus del Español Web/Dialects (henceforth CE), was chosen, alongside CORPES XXI, from the Spanish Royal Academy (2013). CE gathers data from 2010 to 2018, with about 2 billion words, and CORPES XXI from 2001 to 2020, and 175 million words. CORPES XXI was preferred to other Spanish corpora such as CREA, as it also offers the possibility of carrying out searches distinguishing spoken and written speech.

Cifuentes-Férez's (2008, pp. 296–431) lexica of Motion Verbs in English and Spanish serve as a guideline for the corpus searches. The author presents 375 English and 257 Spanish motion verbs, gathered from verb lists, dictionaries and thesauri (p. 136). The verbs are classified according to their conflation patterns (i.e., conceptual elements encoded in the verb), and the full list of factors taken into account for the tagging of the verbs can be found in Cifuentes-Férez (2008, pp. 138–142). For the present study, the verbs are further tagged according to their 'framing type' (i.e., V-framed or S-framed), depending on whether Path was encoded in the main verb or not, respectively, as in Talmy's Motion Framing Typology (see Chapter 2). Verbs were tagged as V-framed even when Path was encoded alongside a different conceptual element (e.g., Manner), as Talmy's Motion Framing typology focuses on whether the core schema (i.e., Path) lies on the main verb.

Before conducting the searches, verbs with ambiguous meanings or with two meanings encoding different conceptual elements were removed from the data, in order to exclude frequencies of use not relevant to Motion. Some examples of the ambiguous verbs which excluded are: *advance, charge, file, frisk, mince, nip, roar, rock, whisk, hurl, adelantar(se)* (get-ahead), *aligerar* (lighten/lighten-step), *divagar* (ramble/walk-around), *partir* (break/leave), *tender* (tend-to/lay), *mudarse* (change-clothes, move-house). The following verbs with two meanings encoding separate elements were also removed: *arise, dodge, moonwalk, rise, rocket, stalk, turn, echar(se)* (throw-oneself), *gatear* (crawl), *girar* (turn), *ladearse* (tilt), *levantar(se)* (get-up), *patear* (stomp), *rastrear* (track), *torcer(se)* (twist), *marchar(se)* (march/leave). After editing, 335 English motion verbs and 225 Spanish motion verbs remained.

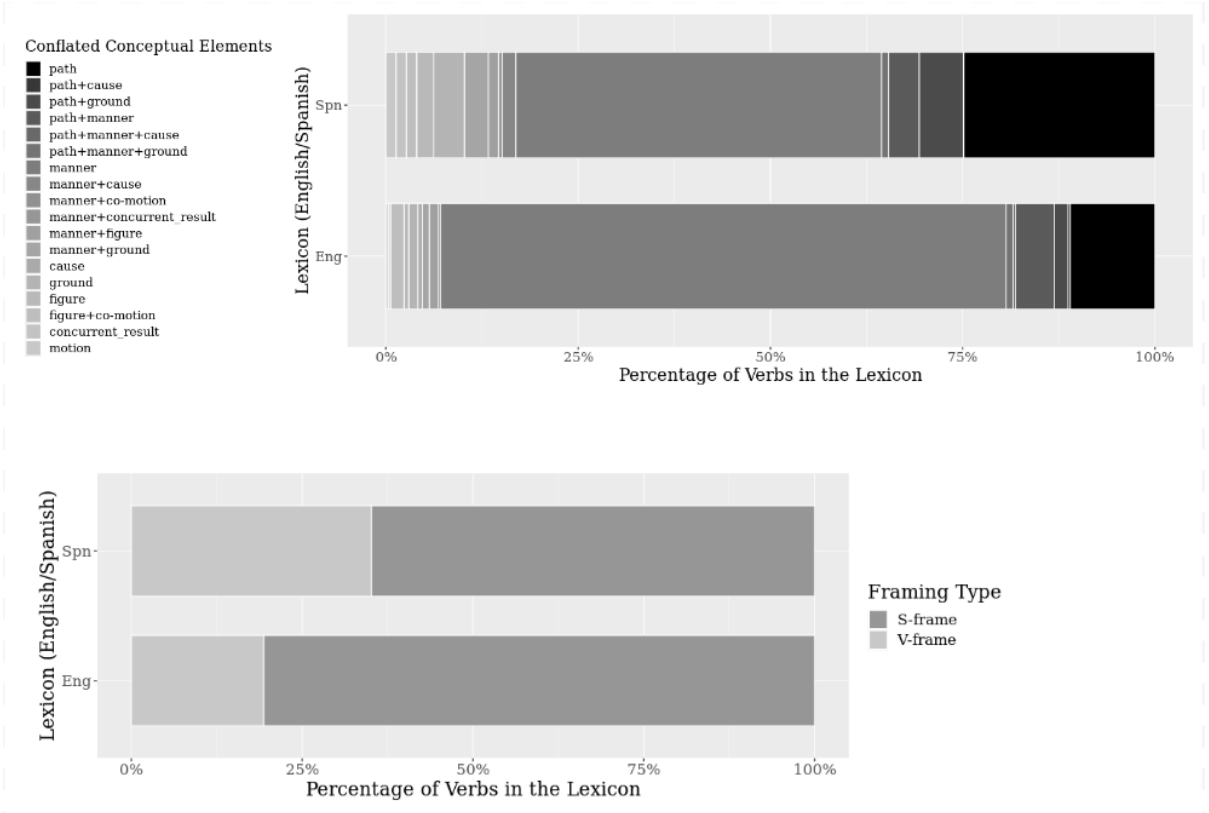
In order to account for the differences in size between corpora, the relative frequencies (per million words) of verb lemmas for the verbs searched in the corpora were put into a data collection table, alongside the framing and conflation type. In addition, frequencies for the

spoken and written speech sections were calculated separately, for the BNC and CORPES XXI. The word origin (from the Oxford English Dictionary)[12] of English verbs was found as well. A total of 1342 obserevations in total were gathered for English, and 899 for Spanish. Due to the large volume, the data is not attached to the present paper as an Appendix, but can be provided on request.

In order to analyze the verbs gathered, the study relies on frequency means, as edited lexica used as a base for the corpus searches contain a different number of verbs, and a different amount of verbs in each conflation type (e.g. Path + Manner) and verb framing type (e.g. S-framed).

Figure 9

Number of Verbs per Type from the (Edited) Lexicon in the Searches Conducted



Note. Henceforth when conflation types are presented, 'Motion' will be omitted unless it is the only conceptual element on the pattern (i.e. Motion + Manner + Path appears in the data as manner+path).

[12] (*Home*, 2022)

As can be observed in Figure 9 both the (edited) English and Spanish lexicon contain more S-framed verbs than V-framed verbs. Both lexica contain a higher number of Manner verbs than Path verbs. The English lexicon shows no instances of the following patterns:

a. Manner + Cause,
b. Manner + Concurrent Result;

whereas the Spanish lexicon lacks the patterns:

a. Cause,
b. Manner + Co-Motion,
c. Path + Cause,
d. Path + Cause + Manner.

Overall, these are 'minor patterns' in both languages (Cifuentes-Férez, 2008, p. 161), and the usage frequency of verbs within such types not allowing for crosslinguistic comparison should not present any significant obstacles to the analysis.

The corpus study seeks to explore the relationships between a number of variables, from a TFS perspective. The main goal is to observe whether the typological differences outlined by Talmy and Slobin have any effect on everyday speech; that is, whether English and Spanish show any significant differences concerning the frequency of motion verbs encoding Path or Manner, expecting English to show higher frequencies for Manner verbs, and Spanish higher frequencies for Path verbs. In addition, the effects of framing type and conflation type on the mode of communication (i.e., written or spoken) are also explored, from the starting hypothesis that V-framed verbs from Romance languages in English will show higher frequencies in written speech, not being preferred for everyday casual speech. These questions are further explored on Chapter 4, which also outlines the results of the study.

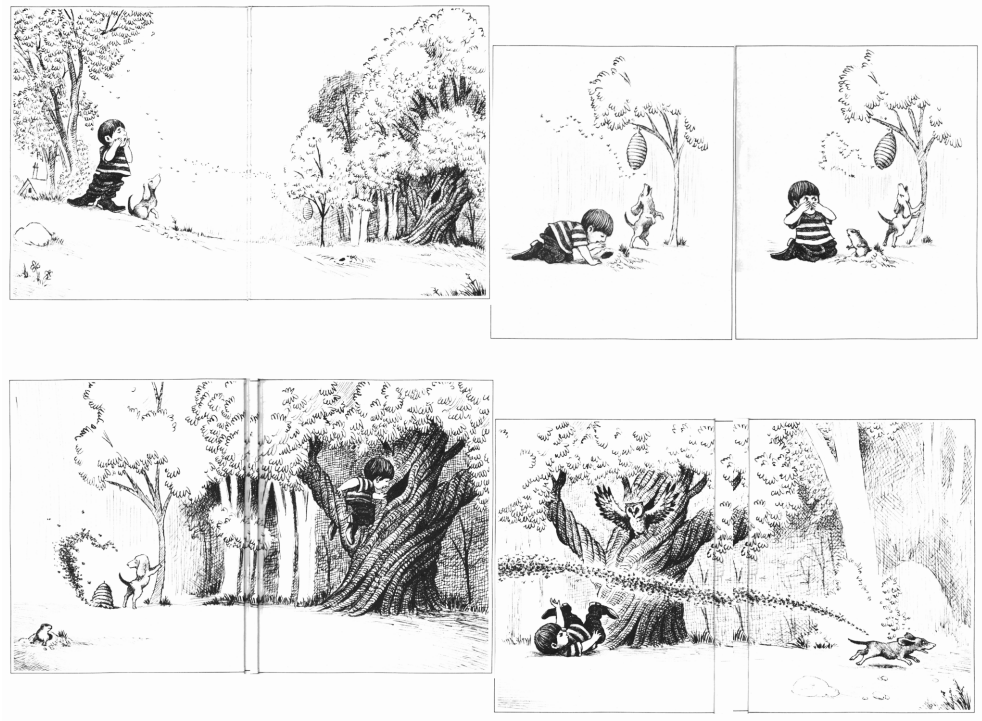
3.2. Experiment

The second experiment conducted in the present paper looks at video-recorded elicited frog narrations, both in terms of speech (i.e., motion verbs used) and gestures. The study examines whether the verbs and gestures in the frog narrations align with the lexicalization and conceptualization patterns described in the literature.

In order to obtain the narrations, a survey was designed using Phonic, a software providing tools for qualitative research, which allows for self video and audio-recording of the participants, and also provides built-in translation. The survey prompted participants to record themselves describing a 'frog story' (Figure 10), a cartoon depicting a number of characters in motion which has been widely used as an elicitation tool in motion event research (Slobin, 2004).

Figure 10

Excerpt from Frog Story Presented to Participants



In addition, they were asked to keep their torso and hands visible, in order for them not to record their face exclusively, which would obfuscate the gesture analysis. The full instructions can be found in Appendix I.

The survey was sent out to English and Spanish speaking participants, through friends and acquaintances, and social media (Facebook, Instagram). It received 26 usable responses, 13 from English and 13 from Spanish. The Spanish participants were all Spanish; and the English participants came from the United States (8), Australia (2), the United Kingdom (1), Canada (1) and Ireland (1). Neither of the participants was proficient in the other language, as bilingual speakers could skew the results.

The obtained videos were then manually processed, transcribing any speech segment describing a motion event accompanied by a gesture. An example of the resulting data can be seen below.

Figure 11

Example of gesture produced by a participant in the video data; anonymized

Once all data was compiled, the motion verbs found were tagged by:

a. speaker;
b. language;
c. motion verb used;
d. framing type of the verb (e.g., S-framed);
e. conceptual element(s) encoded in the verb;
f. whether the verb is bare (i.e., no further information given);
g. whether a satellite or a gerundive of manner accompanies the verb;
h. an orthographic transcription of the context, translated for Spanish.

The gestures found were transcribed generally according to McNeill's framework, and coded according to:

a. their focus in one or more of the conceptual components of a motion event (e.g. Path);
b. their complexity score, according to Kita (1990);
c. their specific characteristics (e.g. hand-shape), through a qualitative analysis.

The resulting data was entered into a data-collection table (Appendix III), including columns for:

a. the speaker (anonymized);
b. their native language (i.e. English/Spanish);
c. the length of the videos, to account for detail;
d. the full transcription (with gestures) of the motion event description, translated for Spanish;
e. the conceptual element(s) encoded in the gesture (e.g. Manner);
f. the gesture's complexity score.

The speech co-occurring with the gestures was not coded, as it varied widely and was not always surrounding a motion verb, therefore falling outside the scope of the study.

The motion verb data can be found in Appendix II, and the gesture data in Appendix III. The analysis and results of the data gathered are presented in Chapter 4, alongside examples.

4. Analysis & Results

Chapter 4 gives an overview of the results gathered from the dual methodology applied in the present paper, starting with the large corpora study, and concluding with the qualitative analysis derived from the elicited frog narrations in the gesture experiment.

4.1. Corpus Study

The corpus study conducted looks into the everyday usage frequency of motion verbs as attested in large English and Spanish corpora, from a Thinking for Speaking perspective. The frequency is studied in relation to three main variables[13]:

a. **Language type**, taking English as an S-language and Spanish as a V-language.
b. **The conflation patterns of conceptual elements** (e.g., Path, Figure) not across the language but within the individual motion verbs, as categorized by Cifuentes-Férez (2008)—further generalized by the author as V-frame (when the Path is encoded in the verb) and S-frame (when the component encoded in the verb is any other); following the hypothesis that English verbs will show greater frequencies for S-framed verbs and Spanish for V-framed verbs.
c. **Verb origin's** impact on both the framing type and whether the verb is used more frequently in **written or spoken language**; from the hypothesis that verbs coming from Romance languages will largely be V-framed and show greater usage frequencies in written English, whereas comparatively no difference between V- and S-framing in Spanish written and spoken language is expected.

These variables serve as the structuring guideline for the remainder of the chapter, as frequency counts, means, and illustrative graphs are presented for each; alongside both analytic and descriptive statistical measures. In addition, interesting arising patterns from the corpus data are explored, with examples offered.

4.1.1. Verb Framing Effects and Usage Frequency

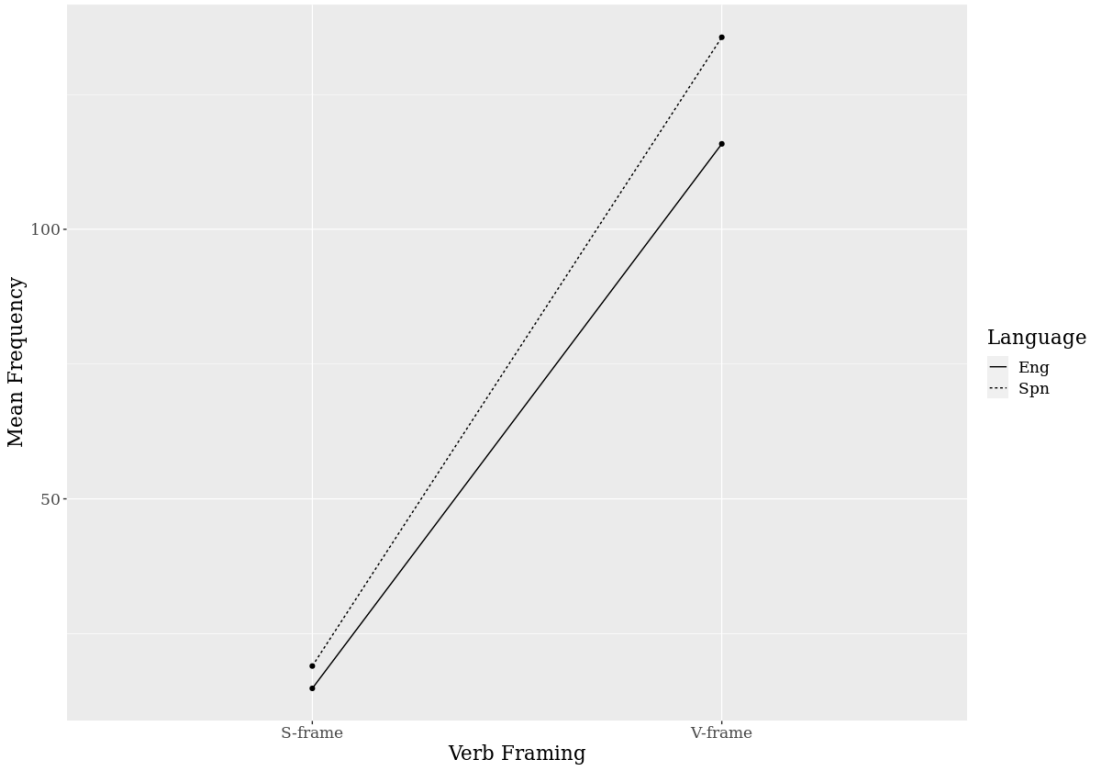
Two different techniques were applied in order to get a first look at the relative frequency of verbs belonging to each framing type, understanding usage frequency as a continuous variable and the framing types as categorical. Firstly, an interaction plot was

[13] Further elaborated-on in Chapter 3.

generated in order to observe how language impacts the mean usage frequency of each framing type (Figure 12).

Figure 12

Interaction Plot of Verb FramingFrequency per Language

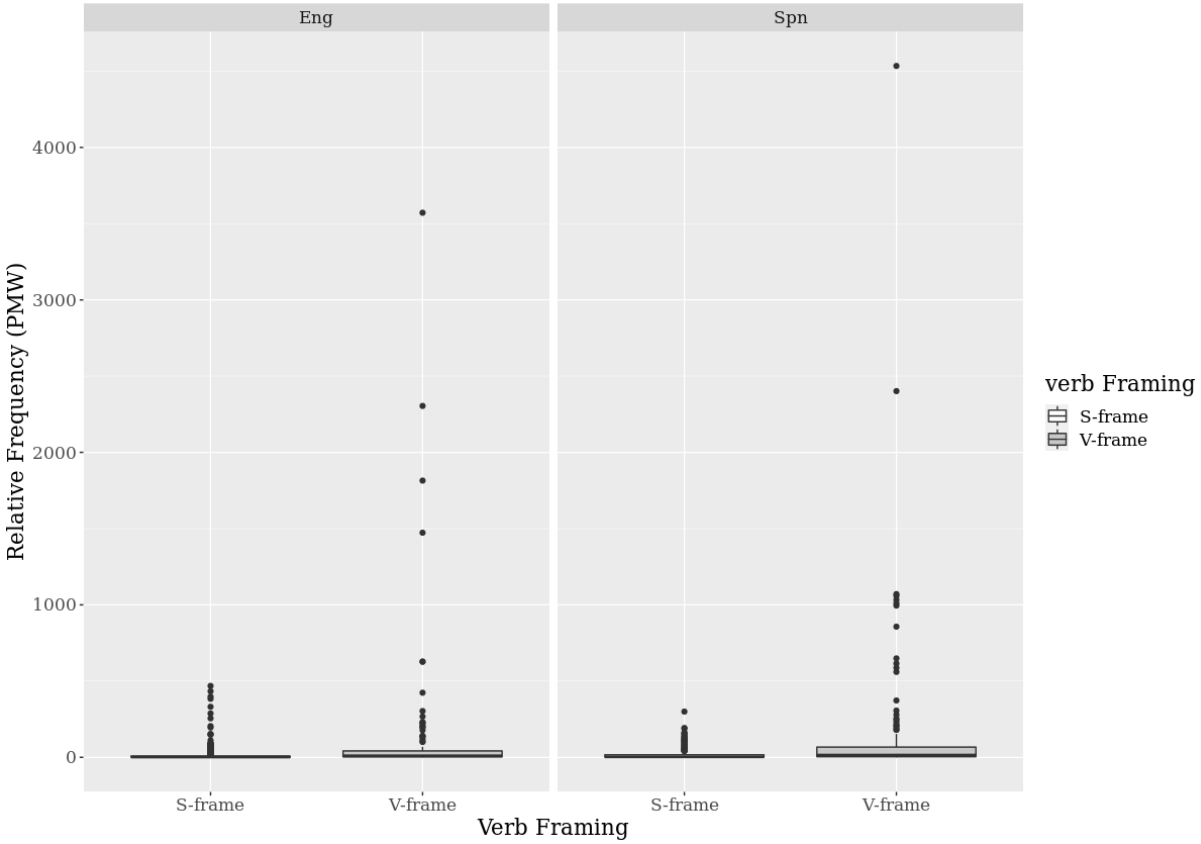


The interaction plot shows some interaction between the variables, as the lines generated are not completely parallel; however, the means are consistently lower for S-framed languages (manner of motion encoded in the main verb) across both English and Spanish, regardless of English being S-framed and Spanish V-framed. The mean relative frequency is also higher for V-framed verbs in both languages; although it is higher for Spanish than English, which is expected from Talmy and Slobin's framework.

In addition to the interaction plot, the results were visualized through a box plot, without any alteration to the y-axis (relative frequency PMW of the verbs), shown in Figure 13.

Figure 13

Box Plot of Verb Framing Frequency per Language



The graph contextualizes the results of the interaction shot, showing several outliers for V-framing which may skew the mean relative frequency, leading to inaccurate visualizations. If we look at the mean alongside the standard deviation for each framing type, as in Table 1, we can observe extremely high standard deviations in V-framed verbs for both Spanish and English, which also substantiates the presence of outliers.

Table 1

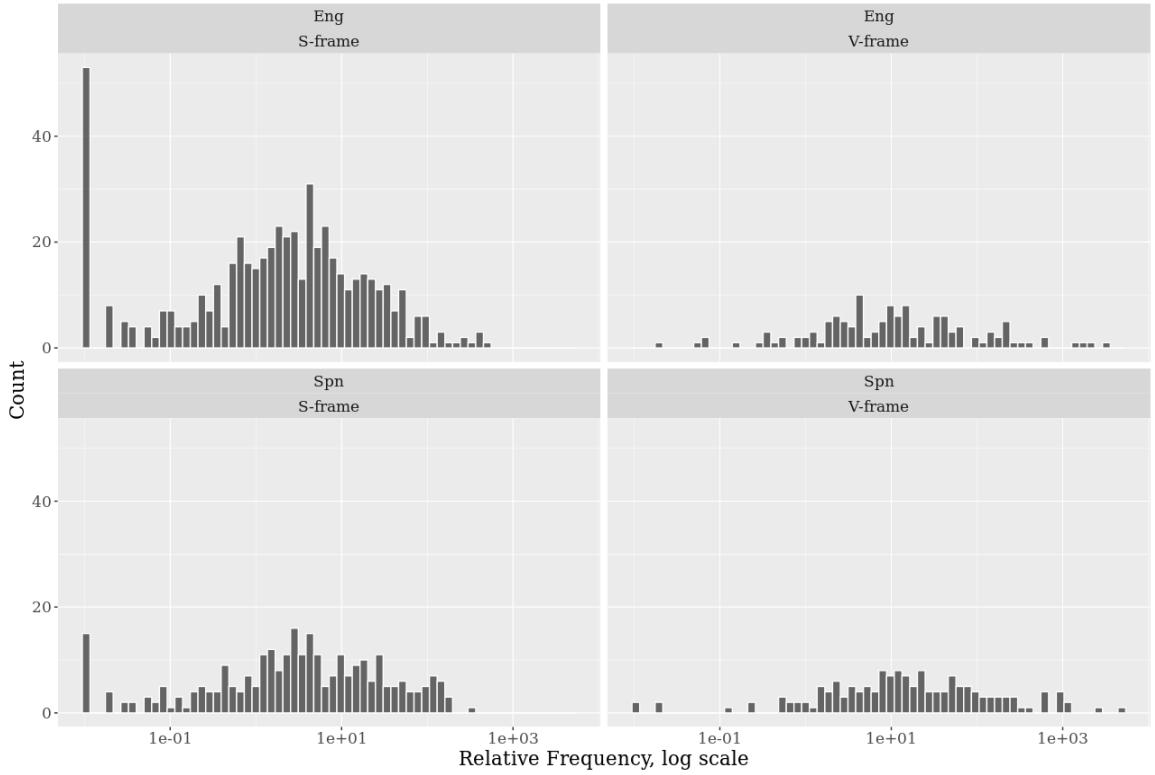
Mean and SD for each framing type per language

Framing Type	Language			
	English		Spanish	
	Mean[14]	SD	Mean	SD
S-framed	5.17	46.82	8.91	38.54
V-framed	26.54	425.58	39.78	450.57

The box plot also presents abnormally low values for S-framed verbs. When checking the distribution (Figure 14) we found that, interestingly, in English, the S-framing type contains a high number of verbs with a frequency of 0, which may also influence the low mean frequency.

Figure 14

Distribution (Histogram) of Relative Frequency of Framing Types per Language

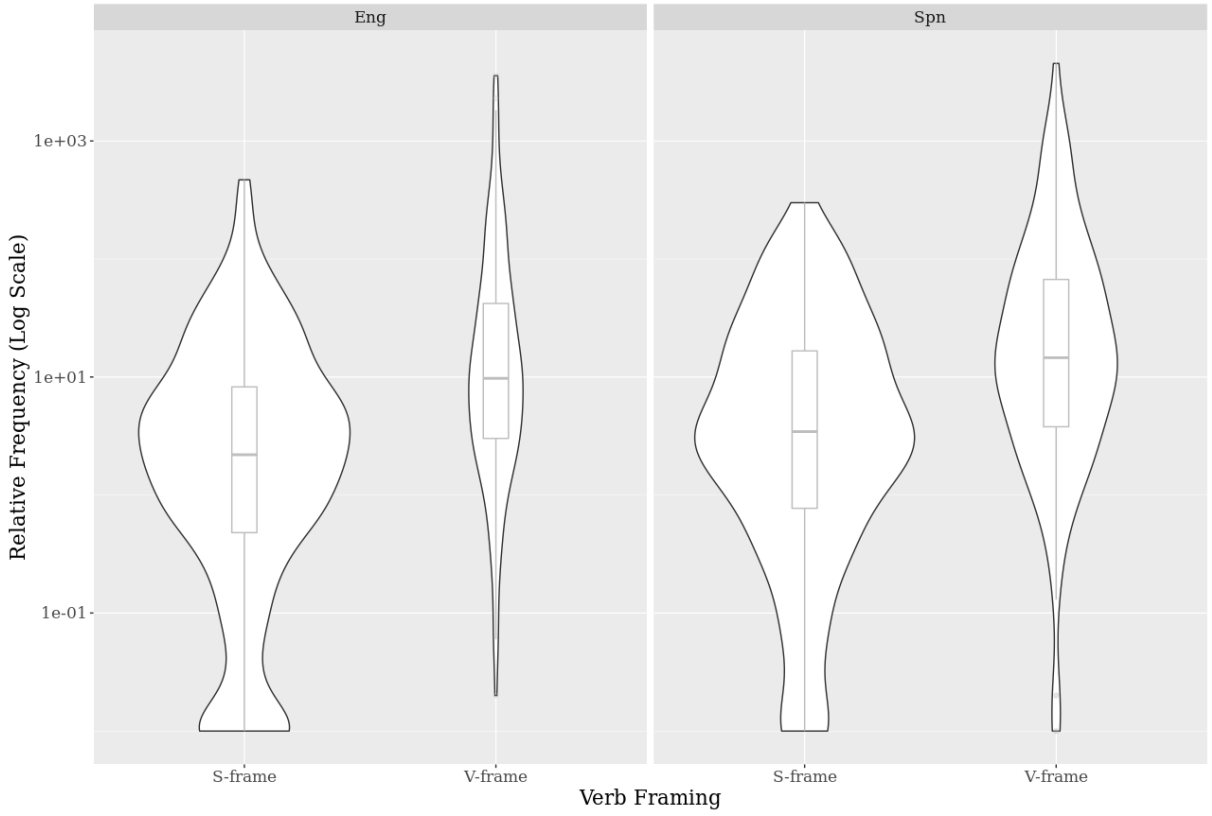


[14] With extreme outliers trimmed (0.1).

Figure 14 also shows that Spanish shows slightly higher frequencies of V-framed verbs; and English higher frequencies of S-framed verbs, in spite of the low outliers. When also looking at a box plot of the data with a logarithmic scale on the y axis so as to better deal with the wide range of values in the data, we find that the distribution of both framing types is quite similar in both languages (Figure 15).

Figure 15

Violin Plot and Box Plot of Verb Framing Frequency per Language



Note. The area of the violin plot varies depending on the number of observations (i.e. number of V-frame verbs in English is lower than in Spanish, therefore the violin is thinner).

Figure 15 shows that many S-framed English verbs have low-frequency values, and that V-framed verbs are more frequent in Spanish overall.

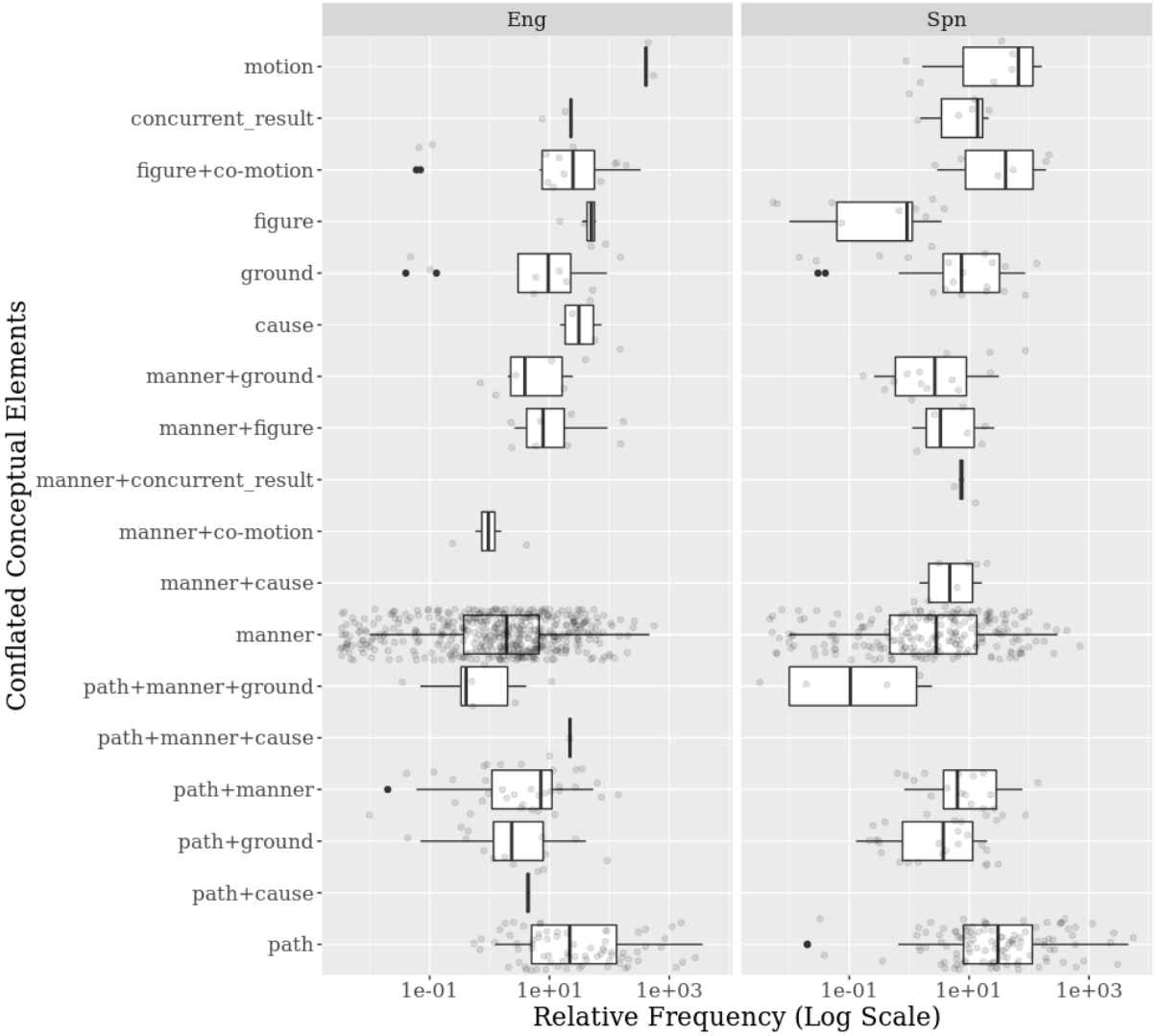
The results were statistically tested, to see the influence of language on the relative frequency of each framing type. The mean relative frequency for each group was tested in a two-way ANOVA test, finding a p-value=0.04 for framing type, and p-value=0.3 for

language; therefore, framing type is a significant predictor for usage frequency, but language is not significant. Applying a single-way ANOVA test for each language, framing type had a p-value=9.12e-08 for English, and p-value=1.37e-05 for Spanish; being slightly more significant as a predictor in English.

Going deeper into the individual conflation patterns of conceptual elements within motion events, the data was visualized through box plots—as violin plots are difficult to read due to the high quantity of conflation patterns—adopting the same strategy for dealing with outliers as with framing types and applying a logarithmic scale, in this case on the x-axis.

Figure 16

Box Plot of Conflation Type Frequency per Language, Observations in Jitter



A few interesting patterns can be noted from Figure 16; Spanish uses more Motion verbs, showing more preference for the zero-conflation pattern than English. Spanish also shows more verbs—and more frequent—conflating Motion with Figure. The other patterns show surprisingly similar distributions, with no significant differences observable between Manner and Path.

4.1.2. Verb Origin and Framing

After studying verb framing, the present study wanted to observe whether word origin has any relation with the framing type of motion verbs in English. Table 2 shows the percentage of verbs belonging to either framing type that were calculated for all word origins found in the data (see Chapter 3).

Table 2

Counts and Percentage of Verb Origin per Framing Type in English

	Framing Type				
	S-framed		V-framed		
	count	percentage*	count	percentage	
Word Origin					
Germanic	*131*	48%	*26*	40%	
Romance	*58*	22%	*32*	49%	
Word Formation	*36*	13%	*4*	6%	
Other	*16*	6%	*1*	2%	
Unknown	*30*	11%	*2*	3%	Sum
Sum	*271*	100%	*65*	100%	336**

* Of all type-verbs. Rounded up.

** Motion verbs in the English lexicon.

From these numbers, we see that roughly half of S-framed verbs in English come from a Germanic language and less than a fourth from Romance languages. The process summarized as ‘word formation’—wherein a non-verb is borrowed and then becomes a verb through a word-formation process in English, such as derivation—is also moderately salient in S-framed languages. In contrast, half of the V-framed languages come from Germanic languages and half from Romance languages, with the remaining word origins being much

less salient for this framing type. To summarize, 49% of V-framed languages in the English lexicon come from Romance languages, in contrast to 22% of S-framed languages; however, roughly half of the V-framed verbs have a Germanic language origin. In addition, 13% of S-framed verbs are borrowed from Romance languages as non-verbs, as opposed to only 6% of V-framed verbs.

The next step was to observe whether V-framed verbs (particularly those coming from Romance languages, e.g., *enter*) were more frequently used in English written speech; which as further explained in Chapter 3, would serve as an indicator of the verbs being used more formally and not being widespread in everyday language. The results are presented below:

Figure 17

Violin Plot of Frequency per Framing Type and Word Origin

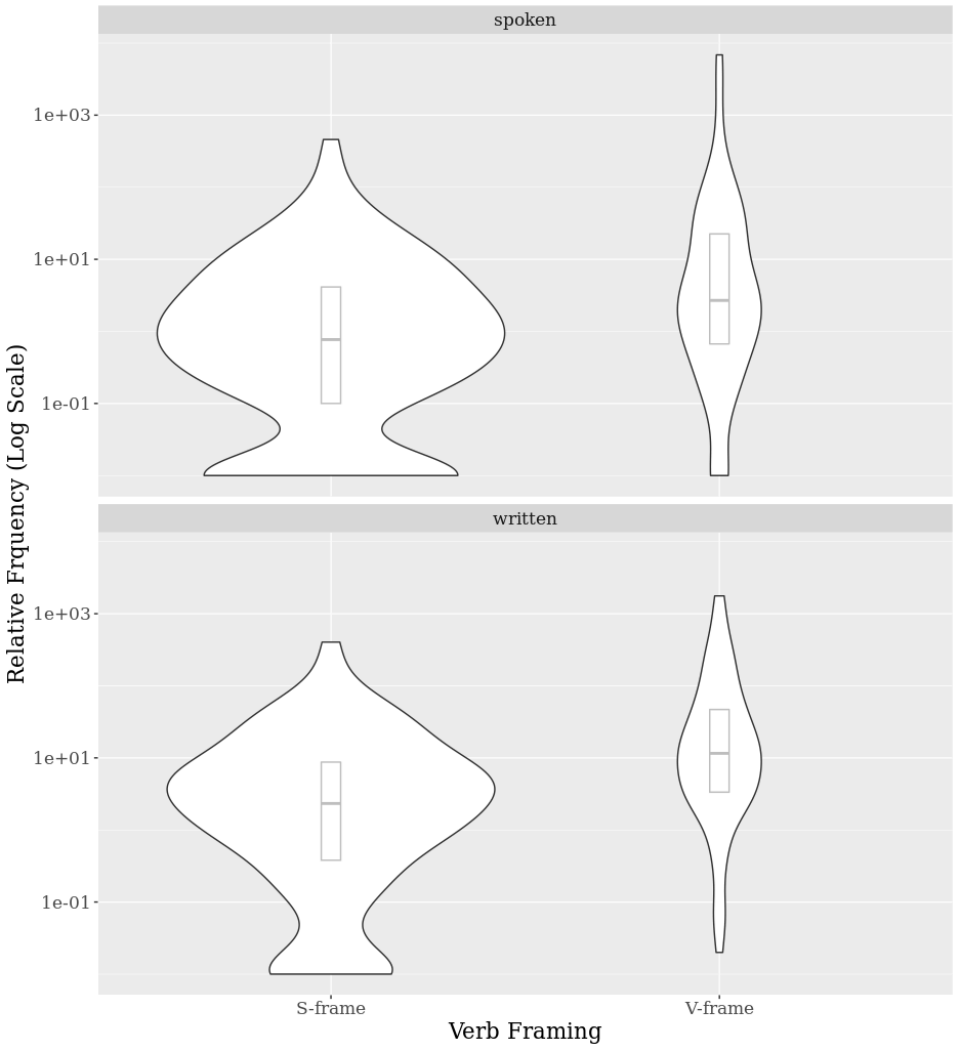
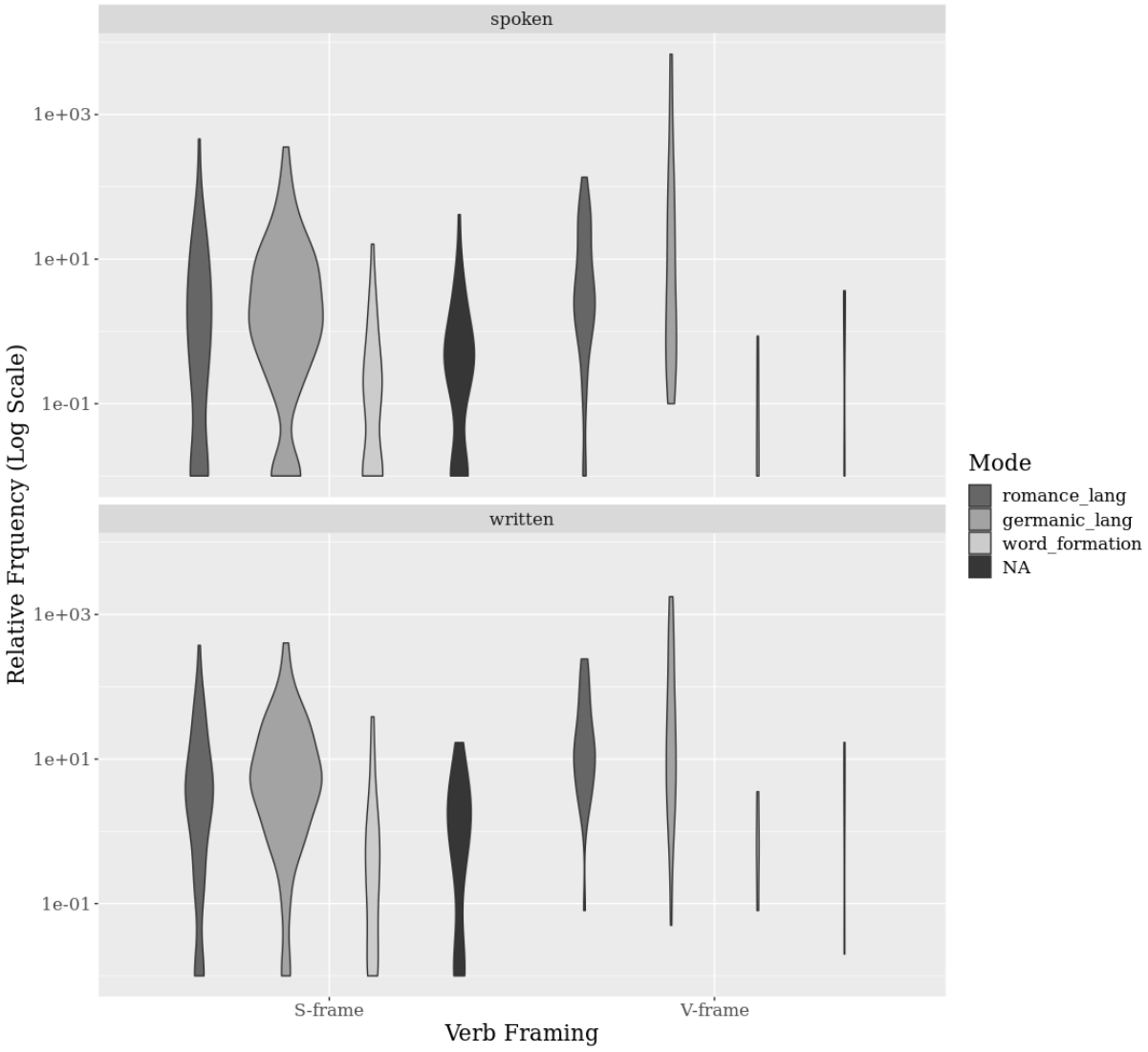


Figure 17 shows little difference between the distribution of verb framing in written and spoken speech, with V-framed verbs actually reaching higher frequencies in spoken language in English. This pattern extends onto the individual verb origins (Figure 18), as the two languages show very similar patterns, with no significant distinctions between the distribution in either mode.

Figure 18

Relative Frequency per Word Origin in Written and Spoken English



4.1.3. Verb Framing in Written and Spoken Language

. The same process was repeated to compare the frequency of verb framing types in written and spoken speech in both English and Spanish. An interaction plot (Figure 19) showed no significant interaction effects.

Figure 19

Interaction Plot of Frequency, Framing Type and Mode in English and Spanish

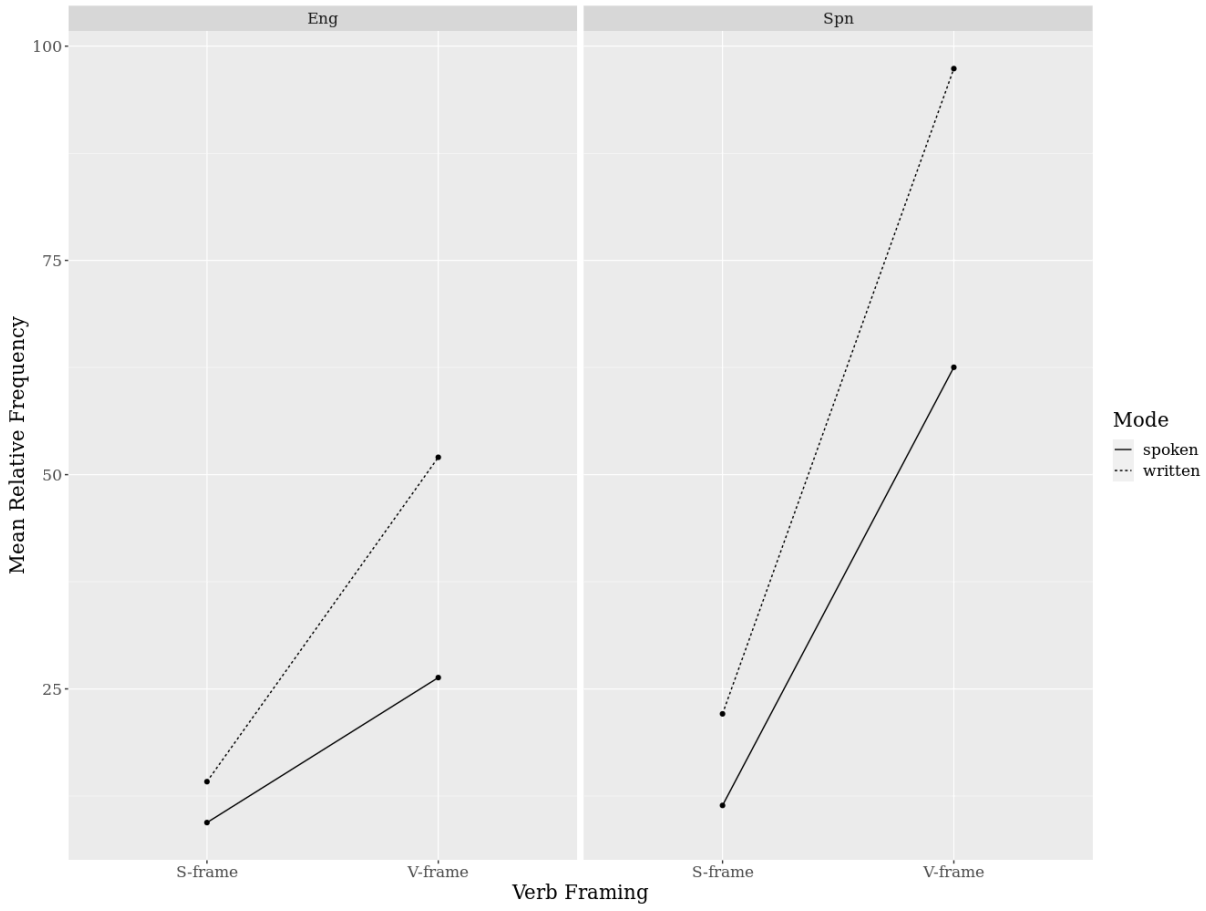
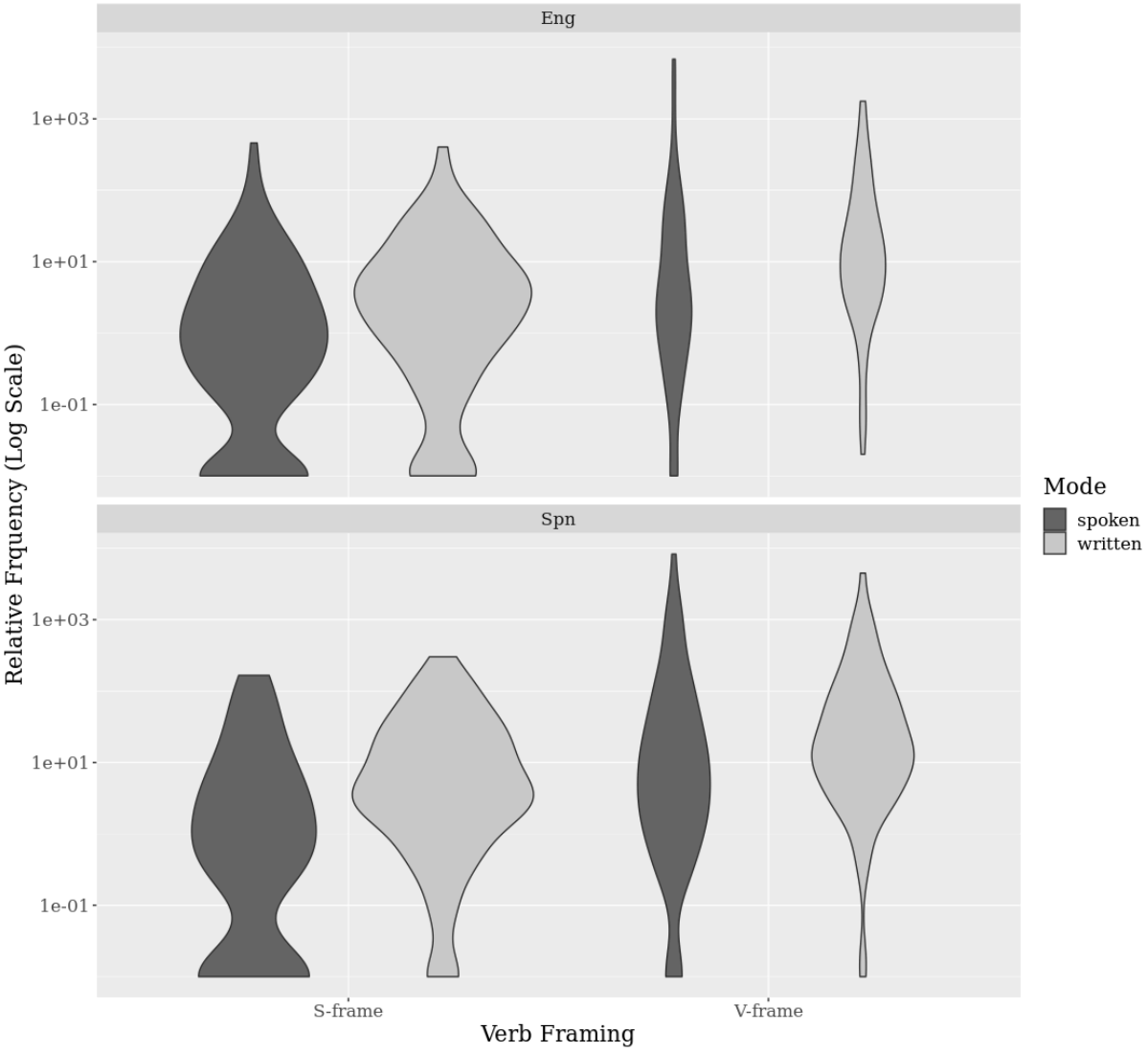


Figure 20

Violin Plot of Framing Type Frequency in Written and Spoken English and Spanish



Looking at a violin plot (Figure 20) we can also see that the languages have very similar distributions, with V-framed verbs showing only a slightly higher mean frequency for written speech in both languages, and V-framed verbs reaching the highest frequencies overall in spoken English.

4.2. Experiment

A similar strategy to Section 4.1 was followed in order to structure the results of the gesture experiment. The experiment addressed a series of sub-questions, drawn from TFS research conducted by Slobin and others, and ultimately based on Talmy's typologies of motion events. The four main aspects to be studied were:

a. whether English participants produced more **detailed descriptions** of motion events, as shown through the length of the descriptions, the number of verbs and gestures produced, differences in verb diversity (verb types/tokens), and rates of bare verbs in both languages;
b. **the conceptual elements** encoded in the verbs and gestures used by the participants and their variation according to language, from a TFS perspective (a focus on Path and Manner);
c. the effect of language type on the conceptualization of motion, through the individual characteristics of the gestures found, and their overall **complexity,** as a nonlinguistic measure;
d. other **interesting patterns** arising from the data.

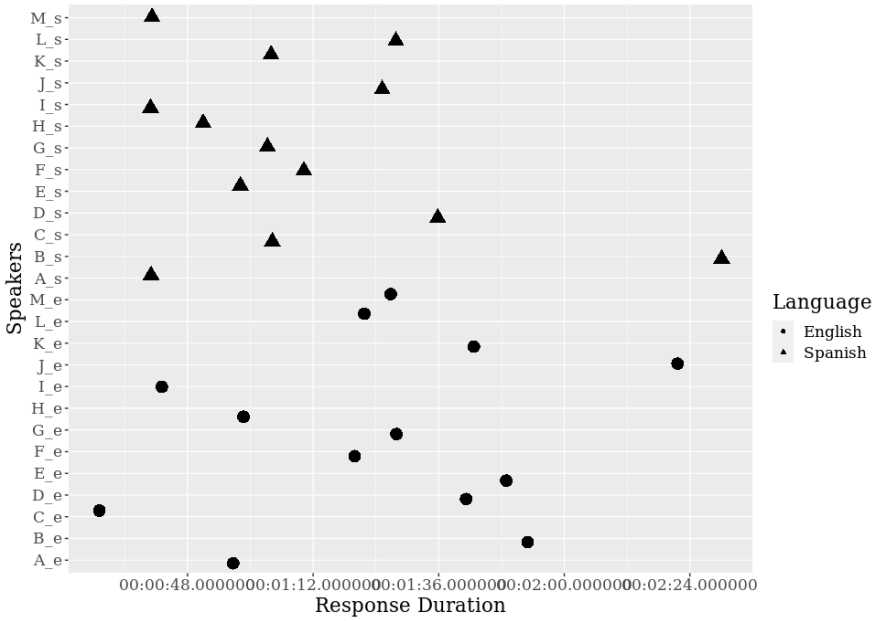
4.2.1. Motion Detail

In order to establish whether one language group described motion in more detail than the other, three variables were looked into: video length, verbs and gestures used (counts), and verb types per token through a type/token ratio (TTR).

Starting with the video length, Figure 21 shows a first look into the data, in which response duration seems to be slightly higher for English speakers, with Spanish speaker's video lengths clustering on the left of the chart (lower duration).

Figure 21

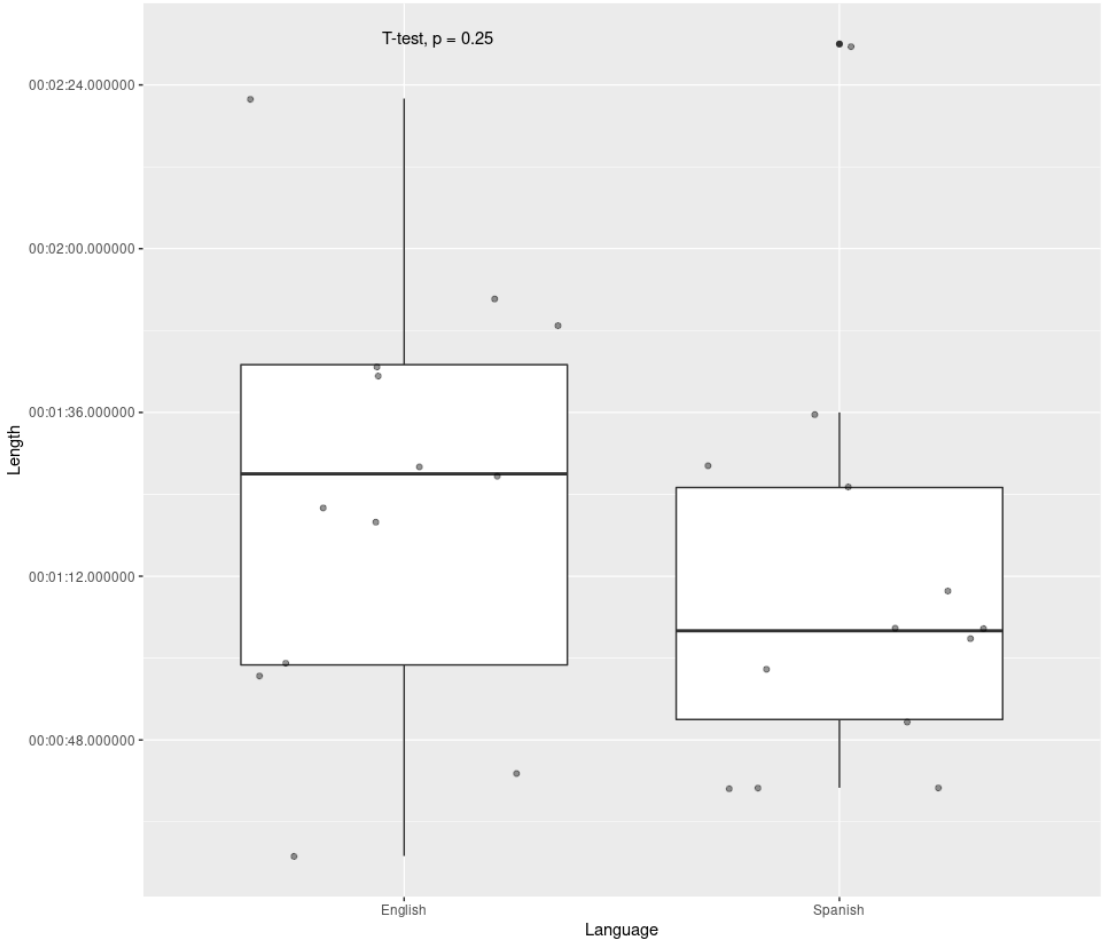
Response Duration per Speaker of Each Language



An examination of the mean response duration (Figure 22) shows that English participants recorded longer videos overall, although with p value = 0.25 the difference is not statistically significant.

Figure 22

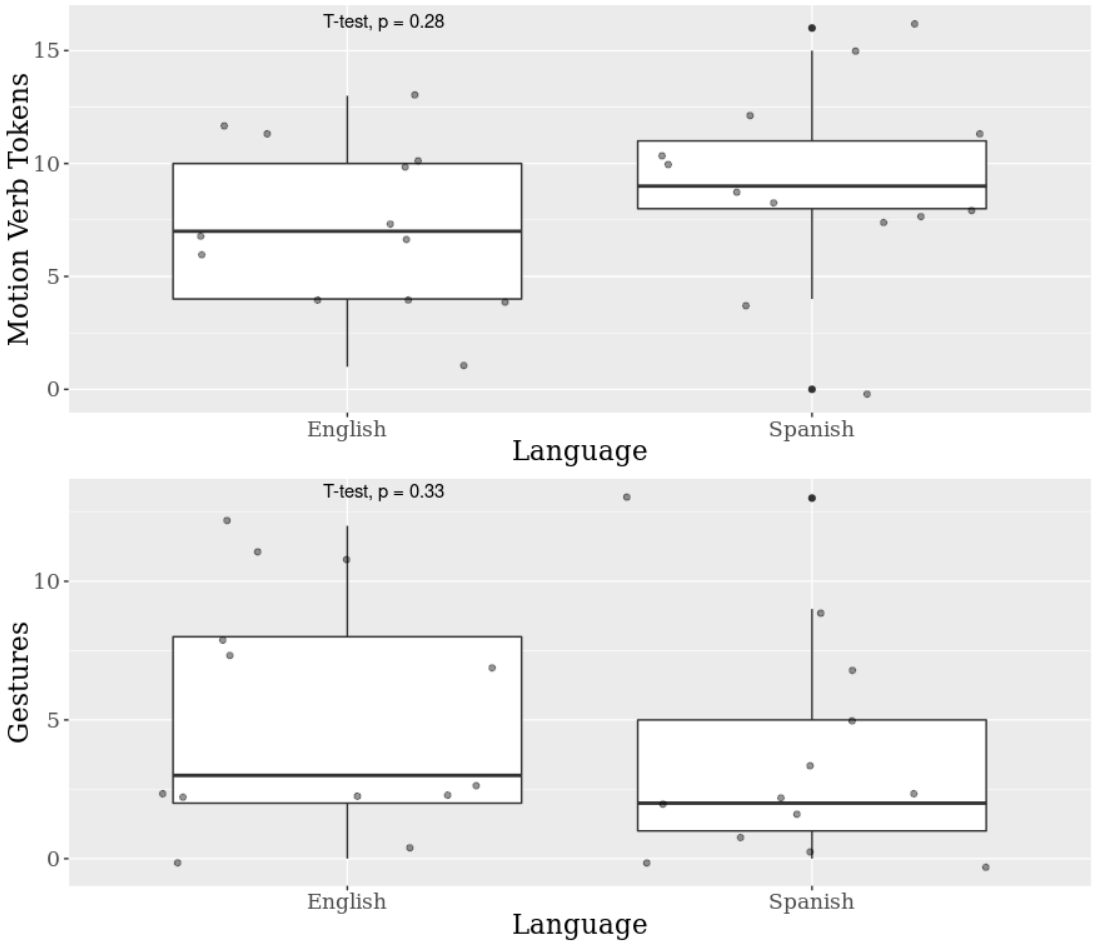
Mean Response Duration per Language



Examining the mean number of verbs and gestures for each language (Figure 23) yields similarly inconclusive results, as there is no significant difference other than a slightly higher number of verb tokens for Spanish.

Figure 23

Mean Verb and Gesture Tokens per Language



If we look into the types (Table 3), we can see that English and Spanish use a very similar number of unique verbs.

Table 3

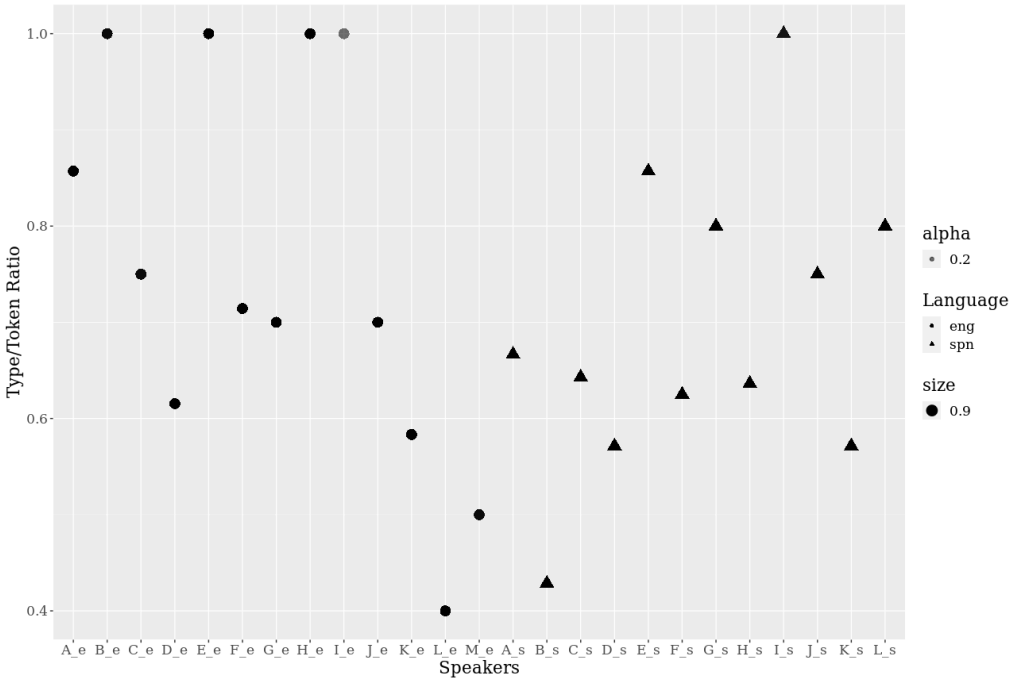
Verb Types in English and Spanish Elicited Frog Narrations

English			**Spanish**		
V-framed	S-framed/Other		V-framed	S-framed/Other	
arrive, chase, climb, come, enter, fall, follow, go, head, pass	*burst, buzz (around), fly, get, go, jump, lad, pop, run, rush, shake, swarm, swing, walk*		*acercar(se)* (come-closer), *apoyar(se)* (lean-on), *caer* (fall), *entrar* (come-in), *escapar* (run-away), *huir* (run-away), *ir* (go), *levantar(se)* (get-up), *pasar* (pass), *perseguir* (chase), *salir* (go), *seguir* (follow), *subir(se)*(climb), *trepar* (climb), *venir* (come)	*agachar(se)* (crouch), *balancear(se)* (swing), *correr* (run), *mover* (move), *pasear* (walk-around), *revolotear* (fly-around), *saltar* (jump-up), *tirar* (throw-away), *volar* (fly), *zarandear* (shake)	
					Total
10	14	**24**	15	10	**25**

However, the list of verbs alone is not a good indicator of the detail put into the descriptions of motion, as it does not account for how many times each verb is repeated. In order to calculate the detail put into the descriptions of motion more accurately, a type/token ratio (TTR) score was calculated for each speaker; dividing the number of verb lemmas per the total number of verbs used in their narrations (Figure 24)

Figure 24

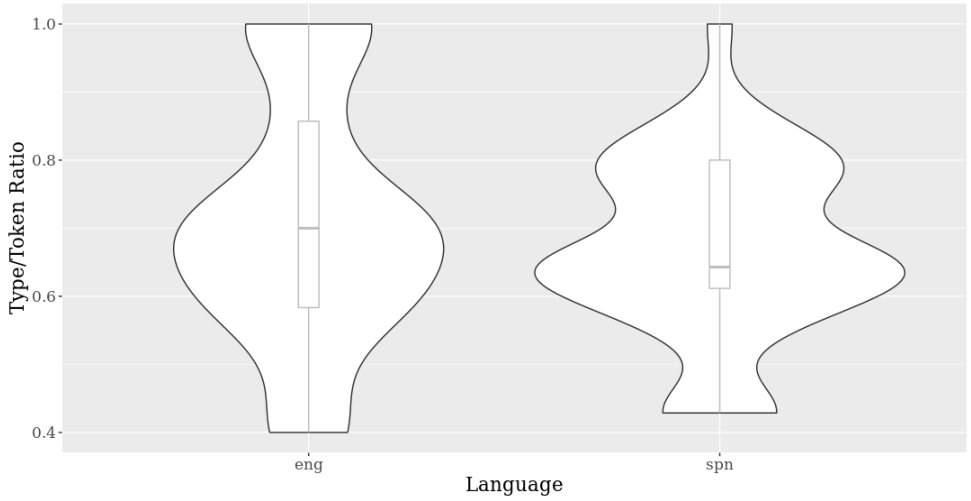
Type/Token Ratio per Speaker



Observing the mean TTR per language, we see that it is slightly higher in English (Figure 25)

Figure 25

Type/Token Ratio per Language; T-test, p= 0.12.



In addition to tokens and TTRs, the rates of bare verbs per total number of verbs used for each language were calculated (Table 4), showing that Spanish contains a much higher percentage of bare motion verbs, which may indicate less attention to Motion.

Table 4

Bare Verbs per Total number of Verbs

English		Spanish	
count	percentage*	count	percentage
5	5%	22	20%

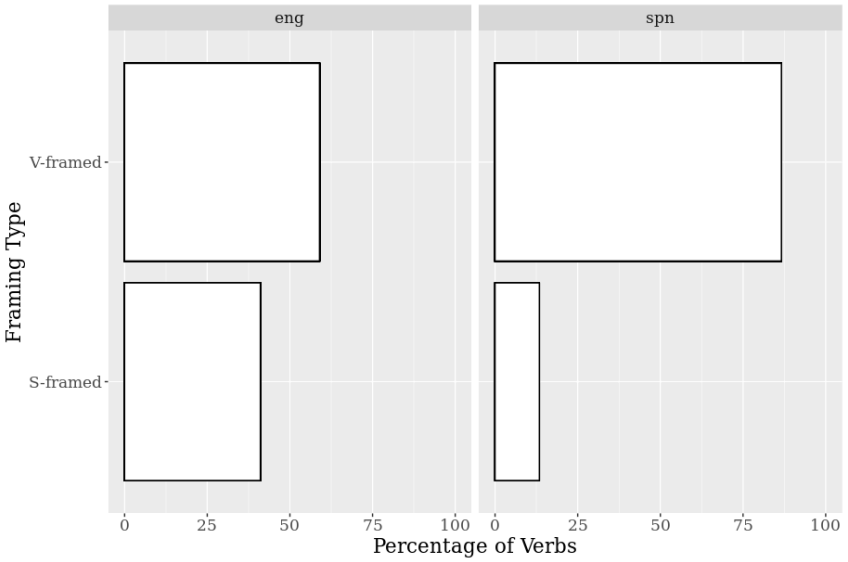
* Of all verbs from that language.

4.2.2. Conflation Patterns in Verbs and Gestures

The framing and conflation type frequencies were also calculated for frog narration verbs and gestures (Figure 26), finding that V-framed verbs were more frequent in both English and Spanish, although S-framed verbs were more frequent in English narrations. Manner verbs are less frequent in Spanish verbs, and Path verbs more frequent.

Figure 26

Framing and Conflation Type Percentages per Language in Frog Narration Verbs



No differences in the general distribution of the conflation patterns were seen for the frog narration gestures (Figure 27).

Figure 27

Percentage Of Frog Narration Verbs Per Conflation Type In Each Language

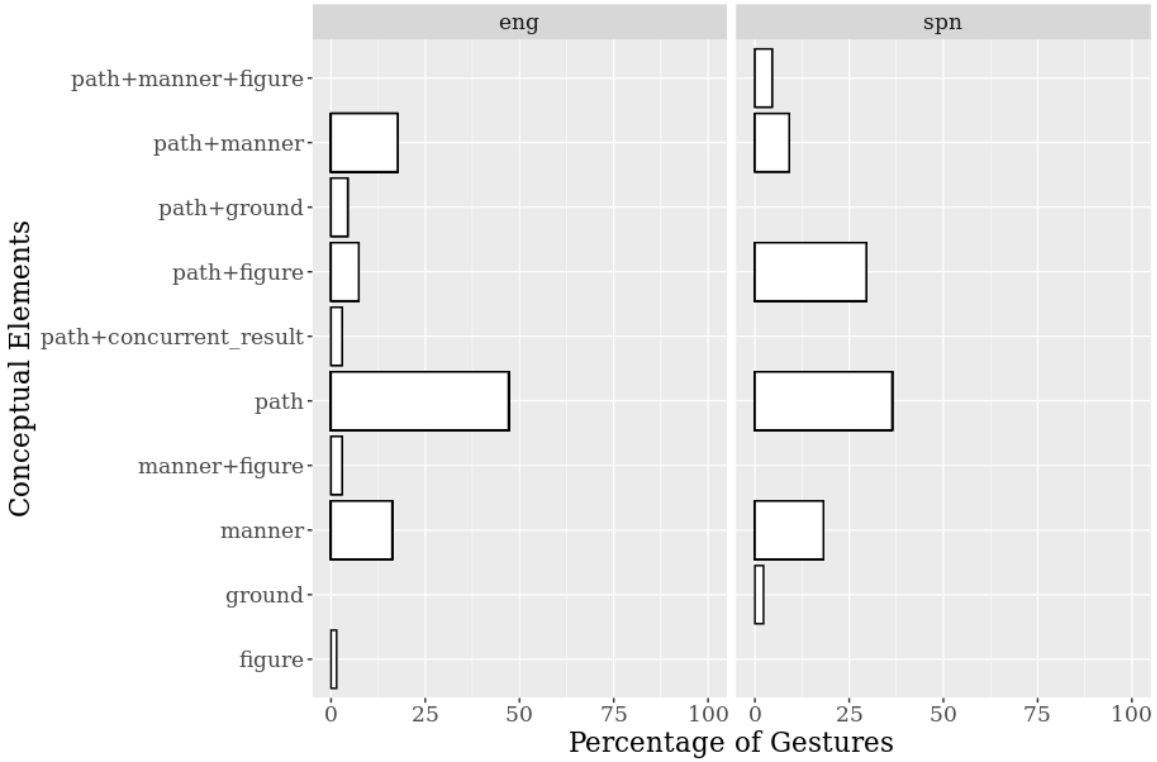
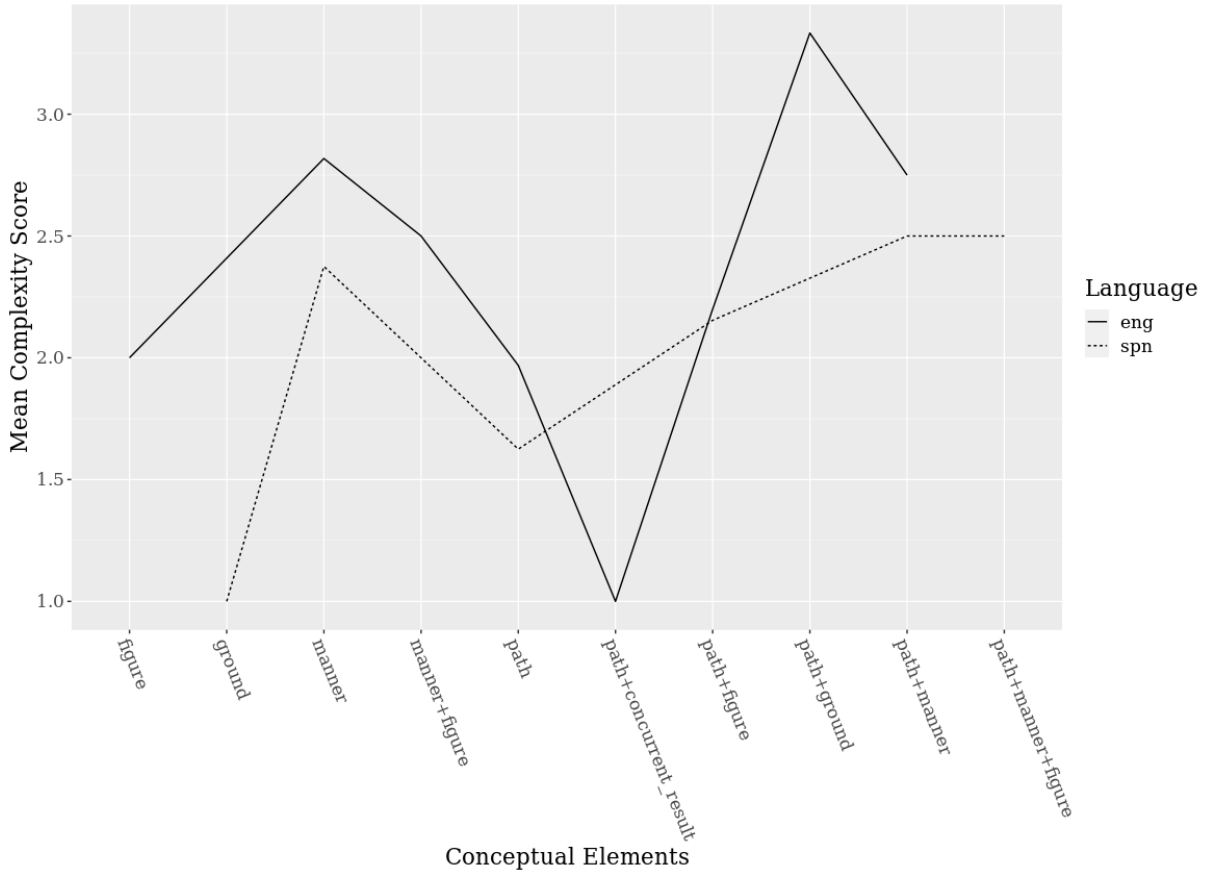


Figure 28

Complexity Scores per Conflation Type in Frog Narration Gestures



To conclude, the gestures' complexity scores were plotted against the conflation types for each language, in order to observe the amount of overall complexity for each language, and the complexity for each conflation type (Figure 28). Results show that Manner gesture complexity scores were higher in English, but the scores dipped on Path gestures. Spanish gestures showed generally less complex gestures, with the highest scoresappearing for Manner and Path+Ground.

4.2.3. Qualitative Analysis

A qualitative analysis of the verbs and gestures found in the elicited frog narrations reveals a few interesting patterns, subsequently outlined.

Verbs of Motion

Concerning the verbs issues with satellite and framing type classification are found. Many English verbs which are generally classified as Path-encoding (Cifuentes-Férez. 2008), such as '*come,' 'go,'* or '*fall'* from the frog narrations further specify the path through satellites in half of all recorded V-framed verbs. For example

e1: an owl has **come** out of the tree (Appendix II; speaker D_e[61]),
e2: to **go** up the tree (Appendix II, speaker D_e[64]),
e3: the beehive **falls** down (Appendix II, speaker G_e[111])

These constructions are preferred both to using the Path-encoding verb without a satellite, specifying the Ground through an adverbial of place (AP)—as in '*an owl has come from the tree*,' '*the beehive falls*;' and to using other Path-encoding verbs which further specify the Path, such as '*an owl exits the tree,*' '*to ascend the tree*,' or '*the beehive descends'* which are not present in the frog narrations at all in English, but which resemble constructions in Spanish. For instance, the same motion events are described in Spanish as

e4: de la corteza **sale** un búho; *from the bark* ***comes-out/exits*** *an owl* (Appendix II, speaker C_s[51])
e5: se sube a un árbol; *he* ***goes-up*** *a tree* (Appendix II, speaker G_s[120]),
e6: la colmena de abejas **cae**; *the beehive* ***falls*** (Appendix II, speaker H_s[130].

Secondly, Talmy's typology generally differentiates satellites from noun phrases, prepositional phrases, and verbs. However, Spanish contains some particles which are present, although not predominant, in the frog narrations obtained. For instance, gerundive adjunct constituents are present in Spanish as a way to further specify Manner, and seem to act as satellites, like

e7: el perro **sale** corriendo; *he* ***goes-out*** *running* (Appendix II, speaker I_s[142]),
e8: se **escapa** corriendo; *he* ***runs-away*** *running* (Appendix II, speaker G_s[121]).

Lastly, 29% of all English Manner-encoding verbs found in the frog narrations appear without a satellite, instead specifying the Path through other components or not at all:

e9: some bees **buzzing** in a line (Appendix II, speaker G_e[103]),

e10: there are bees **flying** overhead (Appendix II, speaker K_e[166])

e11: the dog is **jumping** on the bees (Appendix II, speaker L_e[185])

Gestures of Motion

Gestures showed some crosslinguistic variation. In most instances of Path verbs in Spanish, the gestures occurred with the main verb, usually also Path-encoding, with only very few exceptions, such as

e12: [el búho que hace / **cucú** #]; *[the owl that goes / **cuckoo**]* (Appendix III, speaker I_s[103]

iconic gesture (Path+Figure); moves hands to the front with palms facing the camera, fingers spread;

e13:[y **se ven** también] unas abejas por el aire; [and **there's** also] some bees in the air (Appendix III, speaker J_s[90])

iconic gesture (Path), left index finger slides from right to left in a line.

In contrast, almost half of the Manner encoding gestures in Spanish appeared without any reference to manner in spoken speech, as in

e14: [decide **subirse** a un árbol]; *[decides to **go-up** a tree]*(Appendix III, speaker I_s[76])

iconic gesture (Manner); wiggles hands up and down slightly alternating, palms partially open, emulating the fingers while climbing;

e15: [**mueve** el árbol la colema cae]; *he [**moves** the tree the hive falls]* (Appendix III, speaker E_s[72])

iconic gesture (Manner); shakes fists back and forth.

No patterns of use were found for Path gestures in English speakers, with the participants using them with Path verbs, Manner verbs, non-motion descriptions and other elements of speech. In contrast, English participants seemed to prefer realizing Manner gestures with Manner verbs, or without motion verbs altogether, like in

e16: he's [looking **up** <um> into the hive] (Appendix III, speaker K_e[12]

iconic gesture (Manner); hands form claws and move upwards and to the front in scratching motion,

This leads us to our second point; the several instances in the data in which motion events were not described through motion verbs, but were accompanied by gestures that encoded conceptual elements of motion events. For example,

e17: [the boy is **forced** onto his back] in shock (Appendix III, speaker A_e[6])

iconic gesture (Path); moves hands back towards self with palms open.

The verb '*force'* would not ordinarily be recognized as a motion verb, but the co-occurring gesture encodes Path, emulating the boy being pushed away. Similarly,

e18: debió de asustar [al niño // porque]; [*it must have scared [the child // because]* (Appendix III, speaker B_s[19])

iconic gesture (Path); moves both hands up and outwards with palms open.

The verbal description (i.e, '*to scare*) does not explicitly refer to motion, but the gesture imitates the act of an animal popping out, indicating that the speaker was thinking about the motion event but prioritized verbalizing the kid's reaction.

Further complicating the tagging of motion gestures in relation to the accompanying speech, several instances of two gestures being used for the same motion event were found. For instance,

e19: [a **gopher** or something] (Appendix III, speaker G_e[60])

iconic gesture (Path+Manner); moves hands apart and opens them;

e20: [pops **up** and he's like] (Appendix III, speaker G_e[61])

iconic gesture (Path+Manner); moves hands apart with palms open towards the camera.

Examples 19 and 20 would be coded as a single motion event in speech, with the '*gopher*' being the figure and the main Manner-encoding verb being '*pop*', however, the speaker produces two very similar distinct gestures, one after the other; seemingly conceptualizing the action as two separate events, the gopher first appearing and the gopher popping out.

In contrast, some examples also showed events expressed through two motion verbs but one single gesture,

e21: [**mueve** el árbol la colmena cae]; *[he **moves** the tree the beehive falls]* (Appendix III, speaker E_s[46])

iconic gesture (Manner); shakes fists back and forth.

Example 21 would be understood in speech as two separate motion events: (a) he **moves** the tree, and (b) the beehive **falls**. The speaker seems to conceptualize it at a single event, in which the framing event is the moving of the tree through shaking (encoded in the gesture), and the co-event encoding Concurrent result is the resulting fall of the beehive.

Finally, some continuous gestures were present in the data. These were marked as different gestures whenever they contained two distinct strokes, but where clearly connected and referred to interrelated motion events, as in

e22: [la colmena **de abejas** cae]; *[the hive **of bees** falls down]* (Appendix III, speaker H_s[78])

iconic gesture (Path+Figure); moves right hand down, forming a claw;

e23: [y empieza **a perseguir** al perro]; [*and starts **chasing** the dog]* (Appendix III, speaker H_s[79])

iconic gesture (Path+Figure); moves right hand from right to left in swiping motion, forming a claw.

These gestures occur one after the other and are similar in their stroke characteristics. Two distinct strokes are produced but they seem to reference one motion event, with the framing event being the chasing of the dog, and the co-event encoding Cause being the falling of the beehive. However, the events would normally be coded as separate in speech.

To conclude, it is worth noting that in multiple instances in either language, the stroke of the gesture (in bold) did not occur with the main verb, instead accompanying other aspects, commonly the Figure. This fell outside the scope of the present study but remains a point of interest for further research.

5. Discussion

In the previous chapter, the analysis of the data and the subsequent results were presented, following the structure of the research questions. In the present chapter, these results will be contrasted with the reviewed research literature on motion verbs and gestures, presented alongside Chapter 2 of this dissertation. In addition, the results will be explained and contextualized according to the over-arching working hypotheses drawn from TFS.

5.1. Verb Framing and Conflation Types in English and Spanish

TFS research finds that the typological differences between English and Spanish, as categorized by Talmy and others, are reflected in usage patterns when describing motion events. Slobin (1996a, 2000, 2004), Naigles et.al. (1998) and Ibarretxe-Antuñano (2012) find that Spanish speakers put less detail into motion event descriptions, employing less and less varied motion verbs; and that whereas Path verbs are used in both English and Spanish, Manner verbs are used more in English. However, most of these studies look at frequencies of use in frog narrations from a limited number of participants, as well as literary, translation and conversation corpora of a few thousand words.

Through the twofold methodology carried out in the present study, data similar to the existing literature was gathered from frog narrations; but the author also compiled extensive motion verb frequency data from large corpora, giving a window of insight into everyday usage without the limiting aspect of working with few participants.

Verb Framing. First looking at effects of framing on verb frequencies, the results were not conclusive. Corpus data shows very similar distributions for S-framed and V-framed verbs in English and Spanish, with V-framed verbs reaching overall higher frequencies in both English and Spanish, and S-framed verbs showing a high amount of frequencies close to zero in the data. When looking at the lexicon, we can deduce that the zero-frequencies are due to the high volume of specialized manner verbs in English (e.g., *cancan, somnambulate*), whereas V-framed verbs seem to be more restricted and widely used in both languages. Concerning the mean relative frequency of the framing types, V-framed had higher mean relative frequencies than S-framed verbs in both the English and Spanish corpus. No significant differences between the languages concerning the mean frequency per framing type were found, and language was not a statistically significant predictor. Looking into the individual conflation patterns, it was noted that Spanish shows verbs conflating no other

conceptual element with Motion, whereas English shows almost no instances of zero-conflation; Spanish also exhibits a higher frequency of Motion + Figure verbs compared to English. No differences in frequency distribution between Motion + Manner and Motion + Path verbs were observed in the languages.

In the frog narrations, V-framed verbs were more frequent in both English and Spanish; but S-framed verbs were more frequent in English; similarly, verbs from the Path + Manner conflation pattern were present in both languages but more frequent in English, indicating that Spanish prefers using Path verbs conflating no other conceptual elements. S-framed verbs being more common in English is expected from a TFS focus, but could be explained symply by the typological differences between the languages.

Quantitative gesture data gathered from the frog narrations revealed that English speakers used Path gestures more than Spanish speakers, and Manner gestures roughly the same amount. Path + Manner gestures were also more frequent in English frog narrations. This falls within the line of McNeill's (2008) research, suggesting that as Spanish speakers are able to encode Path in the main verb, and do not understand Path as a series of segments due to the inability to stack satellites, they naturally produce less Path gestures.

When looking at gesture complexity, it was found that English had generally higher complexity scores; although both Spanish and English participants produced less complex gestures for Path than Manner, Spanish showed lower scores for Path. If we take gesture complexity as an indicator for salience, as suggested by McNeill (1992) who stated that speakers will use more complex gestures with conceptual elements salient to them, it would seem logical that if Motion is generally a more salient domain in English than Spanish, scores will be higher in English; however, Spanish showing lower complexity scores for Path would then go directly against a TFS hypothesis, suggesting instead that Manner is more salient for speakers of both English and Spanish.

Stam (2006) finds that Spanish speakers use Path gestures with the main verb, normally V-framed as well, whereas English speakers don't always produce the gesture with the main verb, instead gesturing with a satellite or another element. McNeill & Duncan (2000) note that English speakers use both Manner and Path gestures with the main verb, whereas Spanish speakers often use Manner gestures without an accompanying Manner verb, and sometimes no speech at all (defined as 'manner fog'). Duncan (2001) notes that English and

Spanish speakers produce Manner gestures similarly. Through a qualitative analysis of the gestures, the data shows that in Spanish, gestures occurring with Path verbs were almost always stroked alongside the main verb, substantiating Stam's (2006) results. Manner-encoding gestures in Spanish occurred without Manner verbs half the time, falling within Duncan (2000), although the pattern was not completely prevalent and instances of Manner gestures occurring with Manner verbs were still observed. English participants almost always produced Manner gestures alongside Manner verbs, in contrast. These crosslinguistic differences could in theory substantiate a compensation hypothesis, as Spanish speakers utilized Manner gestures alongside fragments of speech including different elements. However, more research would be needed to substantiate this claim.

Motion Detail. Slobin suggests that the amount of detail paid to Motion differing in various languages results from differences in conceptualization. Looking at the detail of motion descriptions through the length of video responses to the gesture experiment, results show that English responses were on average longer, although the difference is not statistically significant at p=0.25 (T-test). Spanish participants used more motion verbs on average than English speakers (p=0.28), although English participants produced more gestures (p=0.33); again, the differences are not statistically significant. Spanish and English narrations contained roughly the same number of verb types; but when accounting for the lemmas' frequency through a type/token ratio (TTR), it was found that English speakers used more types per token overall, whereas Spanish participants repeated the verbs more often (p=0.12). Finally, 20% of all Spanish motion verbs were bare (no further elaboration on any semantic component, e.g., the toy *fell*), in comparison to 5% of English verbs.

These findings overall seem to substantiate that English speakers describe motion in increased detail when compared to Spanish speakers, as they show longer duration responses and increased TTRs, and use much fewer bare verbs. The differences did not hold up to statistical testing, however, and no distinct patterns were found on the number of verb tokens and gestures produced for either language.

5.1.1. Word Origin, Spoken and Written language

The present study attempted to contribute some data studying the influence of word origin on verb framing. As expected, verbs of Romance origin made up half of the V-framed verbs in the English lexicon, as opposed to 22% of S-framed verbs. This supports the

hypothesis that V-framed verbs are borrowed into S-languages without losing their framing type. A second resulting hypothesis was that V-framed verbs of Romance origin would seem exceedingly formal to S-language speakers, which would not prefer them in everyday communication, leading to V-framed verbs being more frequent in written than spoken communication.

However, the results obtained don't substantiate this hypothesis, as no significant differences in frequency distribution were found for S-framed and V-framed verbs of any word origin in either written or spoken language. If anything, the data shows a slightly higer mean frequency for V-framed verbs in written communication.

5.2. Gestures and Conceptualization

Some interesting patterns deviating from framing were found in the gesture data. For instance, motion events were often encoded through gestures, without accompanying overt references to motion in speech. Furthermore, instances of two gestures being used for a single verbal motion event were observed, as well as one gesture for two verbalized motion events. The lack of complete correlation between the motion events present in the source data (frog story), the motion events verbalized, and the motion events expressed through gesticulation, indicate that there is a separate 'mental picture' of a motion event that does not exclusively rely on language. This goes against TFS' framework, suggesting innstead that although there may be differences in conceptualization across languages, these differences cannot be observed exclusively in speech, and should be substantiated through cognitive elements such as gestures.

5.3. Obstacles and Drawbacks

Some obstacles were encountered throughout the duration of the present study. Firstly, being able to use Cifuentes-Férez's (2008) lexica of motion verbs in English and Spanish was invaluable, as it provided a comprehensive list, as well as the conceptual elements of motion encoded in the verbs. However, as the lexica aim to be complete overviews of the motion systems in either language; they contain highly specialized vocabulary, particularly S-framed verbs, which showed values close to 0 frequency in the corpora, and skewed the results.

Furthermore, issues with the tagging of the verbs arose from the qualitative analysis of frog narrations. For instance, a marked difference was observed between V-framed languages

encoding Path in the main verb *absolutely*, such as '*enter'* or '*descend*', almost all coming from Romance languages; and V-framed languages such as '*come*' which in almost all occasions further specify the Path through a satellite. These verbs would both have been coded as 'V-framed' verbs in the present study, in spite of the fact that verbs like '*descend*' almost never take satellites, in spite of the language technically allowing it (e.g. '*descend down from the tree'*) and that verbs like '*come*' seem to need satellites to compare their meaning. This issue makes it difficult to compare the rates of V-framed and S-framed verbs in languages with different prevalent lexicalization types. As a solution, V-framed and S-framed should in the author's opinion only be applied to languages and not to individual verbs. Instead, verbs could be placed in a cline of Path and Manner salience, as suggested by Ibarretxe-Antuñano (Antuñano, 2009) for languages.

Furthermore, several Manner verbs were found in English frog narrations occurring with no satellites, therefore seemingly not encoding the Path. This would go against Talmy's assumption that the Path is the core schema of a motion event, as there are several instances where it isn't encoded in English motion verbs. In this sense any of the other conceptual elements, such as Manner or Ground, could also serve as the core schema.

6. Conclusions

The main research question to be answered in the present dissertation was whether hypothetical differences in motion event conceptualization stemming from the typological differences between English and Spanish, as outlined by TFS, would be present in everyday speech through a large corpora study; and could be substantiated through a smaller-scale gesture experiment, as a way to survey the 'Thinking' in 'Thinking for Speaking'. From an analysis of gaps in the literature, some sub-questions arose, such as whether any effects of typology could be seen on framing type frequency in written and spoken speech.

All in all, the results from the dual study conducted in this dissertation were highly conflictive. Concerning differences in verb frequency per framing type, as reflective of differences in conceptualization (Path/Manner salience) in English and Spanish:

a. V-framed verbs were more frequent than S-framed verbs in the English and Spanish corpora. No significant crosslinguistic differences in the distribution on the frequency of V-framed or S-framed verb were found in the corpora;

b. framing type was a statistically significant predictor of verb frequency in the corpora, with V-framed verbs being significantly more frequent across both corpora;
c. in frog narrations, V-framed verbs were also more frequent in both languages, although S-framed verbs occurred more frequently in English than Spanish;
d. concerning detail in motion descriptions, English participants produced longer responses, and had overall higher type/token rations, contributing to the hypothesis that English pays more attention to Motion as a conceptual domain than Spanish. However, the differences in response duration and type/token ratio were not statistically significant, and further research in this direction would be needed in order to affirm anything with security.

On the variation relying on verb origin (e.g. Romance languages) as well as mode of communication:

a. it was noted that half of all V-framed verbs on the Cifuentes-Férez (2008) lexicon of English come from Romance languages, substantiating the hypothesis that V-framed verbs can be borrowed into an S-framed language without losing their lexicalization patterns; however,
b. no significant differences were found across English and Spanish, with both languages showing similar distributions of framing type frequency for written and spoken speech, and for verb origin.

Lastly, regarding the value of gestures as tools for studying motion event conceptualization:

a. some patterns regarding motion framing were found. English speakers used Path gestures more than Spanish speakers, suggesting that as Spanish speakers are able to more readily encode the Path in the main verb, they don't have a need for incorporating the conceptual element into gestures. Furthermore, since English speakers verbalize Path through satellites, and multiple satellites can be attached to a main verb, Path is understood as smaller segments and facilitates adding distinct gesticulations to each segment.
b. English had consistently higher gesture complexity scores, although the patterns (complexity score per conceptual element) did not vary between the languages, and Spanish speakers had lower scores for Path. Either we take McNeill's assumption that score complexity serves as an indicator for salience, and that would indicate that Path

is not as salient in Spanish as TFS literature suggests; or gesture complexity does not necessarily indicate salience, but rather can vary widely depending on the speaker. Again, further research would be needed to clarify.

c. Spanish speakers produced Path gestures with Path verbs, but Manner gestures separately from Manner verbs. This could indicate that Path is understood in Spanish as an intrinsic part of the motion event, whereas Manner, is a separate event that sometimes co-exists with the motion event. English participants, in contrast, produced Manner gestures alongside Manner verbs most of the time, possibly indicating that Manner is inseparable from the motion event in English. Although these findings would support TFS conclusions, the data gathered from the frog narrations is insufficient to make definite claims.
d. lastly, gestures are revealed as cognition-fueling mechanisms that do not always need to accompany a motion verb in order to encode a motion event, as several instances were found of motion gestures occurring without references to motion in the speech, which focused on a different aspect (e.g., the Figure's emotional reaction to the motion event). This seems to support the existence of a 'mental image' separate from, or not completely relying on language, which goes partially against Thinking for Speaking. In order to clarify the roles of speech and gestures in motion conceptualization, further studies should incorporate gestures (alongside other nonlinguistic elements such as eye-tracking, perhaps) and step away from verbal descriptions only.

All in all, the present study contributes some data to the research questions proposed by Thinking for Speaking. A great part of the data goes against the theory, as typological differences do not carry over to verb frequency in large corpora, and gesture data does not reveal any exaggerated preference for Manner in English and Path in Spanish. However, several interesting patterns arising from gesture data offer promising directions for future TFS research, and Motion research in general.

References

Alibali, M., Kita, S., & Young, A. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and Cognitive Processes - LANG COGNITIVE PROCESS*, *15*, pp. 593–613.

Alibali, M. W., & Kita, S. (2010). Gesture highlights perceptually present information for speakers. *Gesture*, *10*(1), pp. 3–28.

Alibali, M. W., Kita, S., Bigelow, L. J., Wolfman, C. M., & Klein, S. M. (2001). Gesture plays a role in thinking for speaking. *The International Oralité et Gestualité Colloquium (ORAGE) 2001*, pp. 407–410. https://pure.mpg.de/rest/items/item_3020901/component/file_3020902/content.

Antuñano, I. (2009). Path Salience in Motion Events. *Crosslinguistic Approaches to the Psychology of Language*, p. 423.

Athanasopoulos, P., & Bylund, E. (2013). The 'thinking' in thinking-for-speaking: Where is it? *Language, Interaction and Acquisition*, *4*(1), pp. 91–100.

Berman, R. A., & Slobin, D. I. (2013). *Relating events in narrative: A crosslinguistic developmental study*. Psychology Press.

BNCweb (CQP-Edition). (2018). http://bncweb.lancs.ac.uk/.

Boas, F. (1911). *Introduction to Handbook of American Indian Languages*. University of Nebraska Press.

Bressem, J. (2013). Transcription systems for gestures, speech, prosody, postures, and gaze. In C. Müller, C. Alan, E. Fricke, S. Ladewig, D. McNeill, & S. Tessendorf (Eds.), *Body—Language—Communication 1,* pp. 1037–1059. De Gruyter Mouton.

Chomsky, N., Hauser, M. D., & Fitch, W. T. (2005). The evolution of the language faculty: Clarifications and implications. *Cognition*, *97*(2), pp. 179–210, 211-225.

Cifuentes-Férez, P., & Gentner, D. (2006). Naming motion events in Spanish and English. In *Cognitive Linguistics 17*(4), pp. 443–462.

Consortium, B. N. C. (2007). British national corpus, XML edition. *Oxford Text Archive. Http://Hdl. Handle. Net/20.500*, *12024*, 2554.

Croft, W. (2009). The role of domains in the interpretation of metaphors and metonymies. In *The role of domains in the interpretation of metaphors and metonymies*, pp. 161–206. De Gruyter Mouton.

Croft, W., Barðdal, J., Hollmann, W., Sotirova, V., & Taoka, C. (2010). Revising Talmy's typological classification of complex event constructions. *Contrastive Studies in Construction Grammar*, *10*, pp. 201–236.

Davies, M. (2002). *Web/Dialects* [Corpus]. Corpus Del Español. https://www.corpusdelespanol.org/web-dial/

Davies, M. (2013, present). *Corpus of Contemporary American English* [Corpus]. English-Corpora.Org. https://www.english-corpora.org/coca/

DeLancey, S. (1989). Verb agreement in proto-Tibeto-Burman. *Bulletin of the School of Oriental and African Studies*, *52*(2), pp. 315–333.

Duncan, S. (2001). Co-expressivity of Speech and Gesture: Manner of Motion in Spanish, English, and Chinese. *Annual Meeting of the Berkeley Linguistics Society*, *27*(1), pp. 353–370.

Duncan, S. D., McNeill, D., & McCullough, K.-E. (1995). How to transcribe the invisible: And what we see. *Kodikas*, *18*(1–3), pp. 75–94.

Ebeard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2022). *Ethnologue: Languages of the World.* (25th ed.). SIL International. http://www.ethnologue.com.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), pp. 49–79.

Hohenstein, J. M. (2005). Language-Related Motion Event Similarities in English- and Spanish-Speaking Children. *Journal of Cognition and Development*, *6*(3), pp. 403–425.

Humboldt, W. (1903). *Gesammelte Schriften*. Berlin, Behr. http://archive.org/details/gesammelteschrif04humbuoft

Ibarretxe-Antuñano, I. (2012). Linguistic typology in motion events: Path and manner. *Anuario Del Seminario de Filología Vasca 'Julio de Urquijo'. International Journal of Basque Linguistics and Philology*, pp. 1–39.

Kaye, A. S. (2009). Linguistic Relativity. In S. Chapman & C. Routledge (Eds.), *Key Ideas in Linguistics and the Philosophy of Language*, pp. 116–119. Edinburgh University Press.

Kersten, A. W., Meissner, C. A., Lechuga, J., Schwartz, B. L., Albrechtsen, J. S., & Iglesias, A. (2010). English speakers attend more strongly than Spanish speakers to manner of motion when classifying novel objects and events. *Journal of Experimental Psychology: General*, *139*(4), pp. 638–653.

Lakoff, G. (1987). *Women, fire, and dangerous things*. University of Chicago Press.

Langacker, R. W. (1976). Semantic Representations and the Linguistic Relativity Hypothesis. *Foundations of Language*, *14*(3), pp. 307–357.

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cognitive Science*, *10*(1), pp. 1–40.

Lenneberg, E. H. (1953). Cognition in Ethnolinguistics. *Language*, *29*(4), pp. 463–471.

Levin, B., & Rappaport Hovav, M. (2019). Lexicalization patterns. *Oxford Handbook of Event Structure*, 395–425.

Levy, E. T., & McNeill, D. (1992). Speech, gesture, and discourse. *Discourse Processes*, *15*(3), 277–301.

Living Languages, 2022. (2022). [Map]. SIL International. https://silintl.carto.com/builder/852e178e-e26e-11e5-88b4-0e674067d321/embed

Lucy, J. A. (1992). *Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis*. Cambridge University Press.

Malt, B., Gennari, S., & Imai, M. (2010). Lexicalization Patterns and the World-to-Words Mapping. In *Words and the Mind: How Words Capture Human Experience*, pp. 29–57.

Martínez-Vázquez, M. (2013). Intralinguistic variation in the expression of motion events in English and Spanish. *Lingue e Linguaggi*, *9*, pp. 143–156.

Mayer, M. (2003). *Frog, Where Are You?* (Illustrated edition). New York: Dial Books.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press.

McNeill, D. (2000). *Language and gesture* (Vol. 2). Cambridge University Press.

McNeill, D. (2008). Gesture and Thought. In *Gesture and Thought*. University of Chicago Press.

McNeill, D., & Duncan, S. D. (2000). Growth points in thinking-for-speaking. *Language and Gesture*, *1987*, pp. 141–161.

McWhorter, J. H. (2014). *The Language Hoax: Why the World Looks the Same in Any Language*. Oxford University Press.

Moravcsik, E. A. (2013). *Introducing Language Typology*. Cambridge University Press.

Naigles, L. R., Eisenberg, A. R., Kako, E. T., Highter, M., & McGraw, N. (1998). Speaking of motion: Verb use in English and Spanish. *Language and Cognitive Processes*, *13*(5), pp. 521–549.

OED. (2022). Oxford English Dictionary. https://www.oed.com/

Parrill, F. (2011). The relation between the encoding of motion event information and viewpoint in English-accompanying gestures. *Gesture*, *11*(1), pp. 61–80.

Pederson, E. (2009). Linguistic Relativity. In H. Narrog & B. Heine (Eds.), *The Oxford Handbook of Linguistic Analysis*, pp. 663–678. Oxford University Press.

Pederson, E. (2010). Cognitive Linguistics and Linguistic Relativity. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford Handbook of Cognitive Linguistics*, pp. 1012–1044. Oxford University Press.

Pinker, S. (2007). *The Language Instinct: How the Mind Creates Language* (Reprint edition). Harper Perennial Modern Classics.

Real Academia Española. (2013). *Corpus del Español del Siglo XXI (CORPES)*. Banco de Datos(CORPES XXI) [En Línea]. https://apps2.rae.es/CORPES/

Reshöft, N. (2011). Converging evidence in the typology of motion events. *Converging Evidence: Methodological and Theoretical Issues for Linguistic Research*, *33*.

Rosen, S. (1999). The Syntactic Representation of Linguistic Events. *Glot International*, *4*(2).

Sapir, E. (1921). *Language: An introduction to the study of speech*. Harcourt, Brace.

Sapir, E. (1929). The Status of Linguistics as a Science. *Language*, *5*(4), 207–214.

Schultze-Berndt, E. F. (2000). *Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language*. https://repository.ubn.ru.nl/bitstream/handle/2066/147040/mmubn000001_325820651.pdf

Slobin, D. I. (1987). Thinking for Speaking. *Annual Meeting of the Berkeley Linguistics Society*, *13*(0), pp. 435–445.

Slobin, D. I. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics*, *1*(1), pp. 7–25.

Slobin, D. I. (1996a). From "Thought and Language" to "Thinking for Speaking." In J. Gumperz & S. Levinson (Eds.), *Rethinking Linguistic Relativity*, pp. 70–96. Cambridge University Press.

Slobin, D. I. (1996b). Two Ways to Travel: Verbs of Motion in English and Spanish. In M. Shibatani & S. Thompson (Eds.), *Grammatical Constructions* (pp. 195–219). Clarendon Press.

Slobin, D. I. (2000). Verbalized Events: A Dynamic Approach to Linguistic Relativity and Determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for Linguistic Relativity*, pp. 107–139. John Benjamins Publishing.

Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. *Language in Mind: Advances in the Study of Language and Thought*, *157192*.

Slobin, D. I. (2004). The Many Ways to Search for a Frog: Linguistic Typology and the Expression of Motion Events. In *Typological and Contextual Perspectives*, pp. 219–257. Psychology Press.

Slobin, D. I. (2005). *Linguistic representations of motion events: What is signifier and what is signified*, pp. 307–322. John Benjamins Publishing.

Slobin, D. I., & Hoiting, N. (1994). Reference to movement in spoken and signed languages: Typological considerations. *Annual Meeting of the Berkeley Linguistics Society*, *20*(1), pp. 487–505.

Song, J. J. (2011). Setting the stage. In *The Oxford Handbook of Linguistic Typology*.

Stam, G. (2006). *Thinking for speaking about motion: L1 and L2 speech and gesture*. *44*(2), pp. 145–171.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and lexical descriptions 3,* pp. 57–149. Cambridge University Press.

Talmy, L. (1991). Path to Realization: A Typology of Event Conflation. *Annual Meeting of the Berkeley Linguistics Society*, *17*(1), pp. 480–519.

Talmy, L. (2000a). *Toward a Cognitive Semantics: Concept Structuring Systems I: Vol. I*. MIT Press.

Talmy, L. (2000b). *Toward a Cognitive Semantics: Concept Structuring Systems II: Vol. II*. MIT Press.

Talmy, L. (2009). Main verb properties and equipollent framing. In J. Guo, E. Lieven, N. Budwig, S. Ervin Tripp, K. Nakamura, & Ş. Özçalişkan (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, pp. 389–402. Lawrence Eribaum.

Trabant, J. (2000). How Relativistic Are Humboldt's "Weltansichten"? In M. Pütz & M. Verspoor (Eds.), *Explorations in Linguistic Relativity*, pp. 25–45. John Benjamins Publishing.

Wessel-Tolvig, B., & Paggio, P. (2016). Revisiting the thinking-for-speaking hypothesis: Speech and gesture representation of motion in Danish and Italian. *Journal of Pragmatics*, *99*, pp. 39–61.

Whorf, B. L. (1956). *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). MIT Press.

Zlatev, J., & Yangklang, P. (2004). A third way to travel. *S. Strömqvist & L. Verhoeven (Éds), Relating Events in Narrative: Typological and Contextual Perspectives. Mahwah, NJ, Lawrence Erlbaum Associates*, pp. 191–218.