The Pennsylvania State University
College of the Liberal Arts
Department of Applied Linguistics

Manner-Path Conflation through Co-speech Gesture:

An Examination of Motion Events Produced by Brazilian L2 Speakers of English

A paper in
Teaching English as a Second Language
By
Keenan M. Sweigart

© 2016 Keenan M. Sweigart

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Arts

May 2016

The following faculty members approve the MA paper entitled, *Manner-Path Conflation through Co-speech Gesture: An Examination of Motion Events Produced by Brazilian L2 Speakers of English, by Keenan M. Sweigart:*

_____ _____

Deryn Verity 4/19/2016
Senior Lecturer in Applied Linguistics
Director, ESL/EAP Programs
Academic Advisor

_____ _____

James P. Lantolf 4/19/2016
Greer Professor in Language Acquisition and Applied Linguistics
Director of the Center for Language Acquisition (CLA)
Co-Director of the Center for Advanced Language Proficiency Education and Research (CALPER)
MA paper Advisor

TABLE OF CONTENTS

Acknowledgements

I would like to extend a great deal of gratitude to my advisor Dr. James P. Lantolf, without whom I would have been unable to complete both the study that informs this paper and the paper itself. Throughout my studies in the department of Applied Linguistics, Dr. Lantolf's academic and intellectual support and care in guiding my analytical gaze will continue to inform the way that I approach both education and research. I also thank the Center for Language Acquisition for their generous financial support for my research and travels. Additionally, I would like to thank all of the faculty in the department of Applied Linguistics for creating a learning environment, which has truly facilitated my understanding of what it means to be both an educator and researcher.

Abstract

This study investigates the interface between speech and gesture in the cartoon narration of motion events by two advanced Brazilian-Portuguese speakers of English as a second language (L2). As a theoretical basis, this study is framed by both Slobin's (1996a, 1996b, 2003) thinking for speaking (TFS) framework as well as McNeill's (1992, 2005) growth point (GP) hypothesis. More narrowly, my interest is in the L2 participants' production of manner-path conflated gesture—manner and path co-expressed—in concert with a manner verb. Conceptually, the production of manner-path conflated gesture, in synchrony with a manner verb, does not follow the verb-framed language (V-language) typology (Talmy, 2000) of the speakers' first language (L1). V-languages conflate path of motion on the verb and optionally co-express manner in gesture or an adverbial, if at all. Satellite-framed languages, on the other hand, convey manner of motion on the verb and path on satellite phrases; additionally, manner-path conflated gestures are synchronized with a manner verb when manner is in focus. Through such L2 co-speech gesture production, the participants indicate a shift towards L2 TFS or hybridity of the L1 and L2 TFS patterns.

Keywords: manner-path conflation, thinking for speaking, co-speech gesture

Introduction

In the past two decades the field of cross-linguistic research has seen a blossoming of studies on the co-occurrence of speech and gesture. Working within Slobin's (1996a, 1996b, 2003) thinking-for-speaking (TFS) framework, studies have sought to determine whether the acquisition of a second language (L2) results in conceptual shifts in how one speaks about and, thusly, thinks about events in the world. In Slobin's (1996a, 1996b) seminal works the distinction is made between the Sapir-Whorf linguistic relativity hypothesis and a weaker form (TFS); the latter posits that a different type of thinking occurs during language production. To be sure, it is during language production that one's complete concept (i.e., mental image) is mobilized and mapped onto the semantically available structures of a given language. Through successful and constant use in the first language (L1), grammatical patterns of how to describe state, aspect, tense, and motion become entrenched (Tomasello, 2003). Once these patterns have become entrenched, the shifting of conceptual understanding and rearranging of semantic boundaries prove to be quite difficult in the acquisition of a L2 (Gullberg, 2009).

Indeed studies analyzing L2 TFS through speech production have yielded important findings in the field of cross-linguistic research. However, the analysis of speech alone presents us with but one pair-part in the TFS framework. In order to gain a more complete view of the TFS framework, McNeill and his colleagues (McNeill & Duncan, 2000; McNeill, 1992) argue that we must take into account the production of co-speech gesture. Within McNeill's (2000) framework, speaking and gesturing form a dialectical unity, which is referred to as the growth point (GP). The GP — spawning from Vygotsky's (1986) "psychological predicate" and "psychological subject"— consists of a segmented (speaking) and a synthetic (gesturing) semiotic resource, each a manifestation of a uniquely nuanced and co-expressive (but not redundant) characteristic of thought.

In the body of cross-linguistic research, a number of studies have been informed by Talmy's (1985) (and Aske's, 1989) domain of motion events to investigate Slobin's TFS hypothesis in both speech (e.g., Pourcel & Kopecka, 2006; Papafragou, Massey, & Gleitman, 2005; Zlatev & Yangklang, 2004) and speech and co-speech gesture (e.g., Stam, 2015; Brown & Gullberg, 2008; Kita & Özyürek, 2003; Özyürek, 2002; McNeill, 1992, 2001) — the latter is the focus of the current study.

A motion event, briefly, is the movement of an entity through space, which includes four key features: figure, motion, path, and ground. Talmy's proposed framework is particularly appropriate for the analysis of the TFS hypothesis, as it provides a system, with which one can systematically account for the mapping of semantics in each language. Indeed Talmy (2000) has not only provided the means to track semantics, but he has also classified two language typologies according to the way in which languages encode path of motion — verb-framed languages and satellite-framed languages (henceforth V-languages and S-languages, respectively). English, the prototypical S-language, encodes path of motion on the satellite (e.g., The cat ran out of the room), and Spanish, the prototypical V-language, encodes path of motion on the verb (e.g., "*El gato salió corriendo de la sala* 'The cat left running from the room'") (Choi & Lantolf, 2008, p. 193). Additionally, Slobin (2004) has extended Talmy's typology to include equipollently-framed languages (e.g., optional use in Chinese), in which both path and manner are expressed by equivalent grammatical forms. (e.g., "*na zhi mao pao le chu qu"* That cat ran+out+go).

The present study is situated within McNeill's (2000) GP hypothesis in relation to Slobin's (1996a, 2003) TFS framework. The goal of the study is to present evidence on the gesture-speech interface, as produced in L2 English motion event narrations by L1 speakers of

Brazilian Portuguese, with a focus on manner-path conflated gesture. Manner-path conflated gesture, briefly, refers to the co-expression of manner (manner of motion) and path (trajectory of motion) in one fluid gesture stroke. Importantly, the semantic feature with which a manner-path conflated gesture co-occurs is indicative of the TFS pattern (or mix of TFS patterns) of the speaker.

In this paper, I first outline Talmy's (1985, 2000) motion event typology and, in the following section, explain McNeill's GP hypothesis and its relevance in the present study. Next, I analyze the data, and, finally, I discuss the findings of the study and the implications for the importance of co-speech gesture research in L2 teaching and learning.

Theoretical Background

Motion Events in English and Brazilian Portuguese

Talmy's (2000) proposed framework accounts for the different types of lexicalization in motion events, in which he has identified four key features — figure, path, ground, and motion. The figure in a motion event is a moving or moveable object whose path is the focus, and path is the trajectory along which the figure moves. Both the figure and path are set in relation to the ground, i.e., the backdrop or object of orientation. The motion, as its name implies, expresses the motion of the figure. In a motion event, the main verb can be encoded with either non-specific path of motion, or it can optionally express motion in co-events, such as manner of motion (manner for short) or caused motion. Manner is denoted as the specific way in which a figure performs a motion, while caused motion describes the movement of a figure as set into motion by an agent (e.g., "The boy kicked the ball into the goal" where the ball is the figure, goal is the ground and the energy imparted by the boy's action imparts motion to the ball).

As noted above, Talmy (2000) distinguishes between two different language typologies based on their lexicalization of path of motion, both of which will be of focus here — S-languages and V-languages. In S-languages such as English (EN) (1), path is typically encoded on a satellite, a particle, or prepositional phrase. V-languages such as Spanish (SP) (2) and Brazilian Portuguese (BRP) (3), on the other hand, lexicalize path on the verb.

(1)EN- Sylvester climbed up the drainpipe.
(2)SP- Sylvester subió el canalón. "Sylvester ascended the drainpipe."
(Choi & Lantolf, 2008, p. 194)
(3)BRP- Sylvester subiu pelo cano. "Sylvester ascended the drainpipe."

While S-languages and V-languages follow typical patterns of lexicalization of path of motion, it is not out of the realm of possibility for language types to encode path of motion differently. For example, speakers of English can optionally encode path of motion on the verb, due to the Latin origins of some English verbs (e.g., "The boy descended the stairs"); however, it is infrequently used, especially in colloquial speech.

Lexicalization of optional manner of motion across S-languages and V-languages is another important distinction that is of focus in this study. In S-languages, manner is commonly expressed through the verb, whereas V-languages express manner, if present and in focus, by lexicalizing it "outside of the verb on an adjunct, an adverbial particle such as a gerund or a phrase" (Stam, 2015, p. 62). In the examples below path and manner are presented in a S-Language, English (4), and juxtaposed by two V-languages — Spanish (5) and Brazilian-Portuguese (6).

(4)EN- The cat rolls out of the drainpipe.
(5)SP- El gato sale rodando del canalón. "The cat exits rolling from the drainpipe."
(Choi & Lantolf, 2008, p. 194)
(6)BRP- O gato sai rolando do cano. "The cat exits rolling from the drainpipe."

In (4), the speaker expresses the manner in which the cat exits the drainpipe on the verb *rolls,* and encodes path of motion on the satellite *out of the drainpipe,* a prepositional phrase consisting of both a satellite and ground noun. The expression of manner through the verb and followed by path on a satellite, as Slobin (2000) has noted, indicates that for S-language speakers, manner and directed motion seem to be conceived of as "a *single conceptual event*, making it difficult to have a mental image of one without the other" (p. 132; emphasis in original). Manner production is not required, however, when manner is highly defocused. In the case of highly defocused manner, S-language speakers may direct their focus on path of motion (e.g., "He goes out of the drainpipe"). For V-languages, manner is conceived of as occurring in a specific location, i.e., not accumulating path. In both (5) and (6), the speakers convey path on the verb (*sale* and *sai* both meaning exits), and manner on an adverbial (*rodando* and *rolando* both meaning rolling). Salience of features across language typologies, as Slobin (2003) claims, is due to the "availability of a main-verb slot for manner verbs in S-languages, in contrast to a main-verb slot for path verbs in V-languages" (2003, p. 175). Slobin further states that the availability of manner in the main-verb slot in S-languages calls for a more habituated use in online[1] production of motion events. Moreover, S-languages are particularly rich with manner verbs and frequently express manner of motion through such verbs as crawl, dash, dart, dive, float, flip, etc. V-language speakers may also express manner on the verb. Largely the use of these manner verbs, however, consist of so-called basic manner verbs (e.g., walk, run, fly) that lightly define the type of motion, as compared to the complex manner verbs (e.g., roll, hop, skip) used in S-languages that heavily express the characteristic of the undertaken motion.

[1] This term refers to the type of speaking that affects or shapes thinking while being affected by thinking in a dialectical process.

Co-Speech Gesture

Growth Point

McNeill (2000) has presented a system of analysis through which a more holistic picture of a speaker's online thought is rendered visible. This system is predicated on the GP hypothesis, which denotes the simultaneous production of a synthetic imagistic unit (gesture) and a linear-segmented portion (speech) (see figure 1). Of the two modes, only one is considered to be representational— the linguistic— whereas the other is considered to be the "material image of meaning" (Vygotsky 1987, as cited by McNeill, 2005, p. 98) and bears a resemblance to the visuospatial action of the reality it is intended to depict (Negueruela & Lantolf, 2008). That being said, the gesture mode alone does not convey meaning; it is only with its co-occurring, co-expressive part of speech that the gesture bears meaning. This dialectically entwined unit is forged at the point in spoken production that synchronizes with the stroke phase of a gesture.

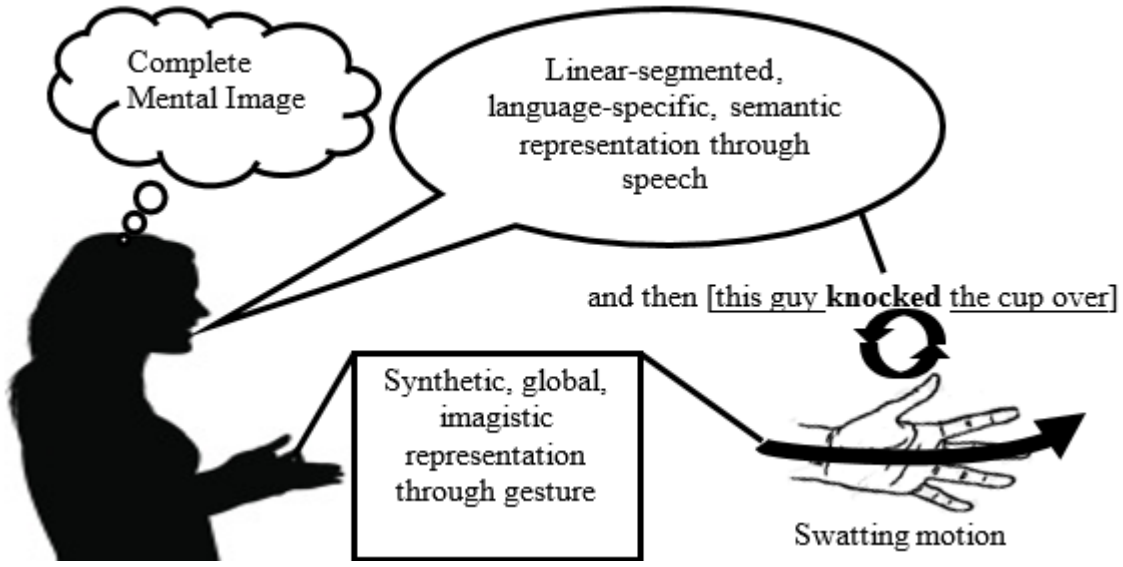

Figure 1: Growth Point

According to McNeill (2005) the stroke phase is the most important phase of the gesture, as this is the point in thought that is most salient for the speaker; the stroke phase is indicative of

salience of movement, as the movement is expressed through both a semantic and kinetic component. The stroke phase is part of a larger gesture phrase in which McNeill identifies six phases: preparation (optional), prestroke hold (optional), stroke (mandatory), stroke hold (optional), post stroke hold (optional), and retraction (optional).

Gesture and Motion Events

In a study on the TFS patterns of L1 speakers of English and Spanish through the GP hypothesis, McNeill and Duncan's (2000) findings suggest that speakers of S-languages and V-languages follow distinctive speech-gesture patterning according to language type — TFS is not only different in speech patterns. English speakers "synchronize their path gesture with verbal satellites, with ground noun phrases (NPs), or with a combination of both" (Choi & Lantolf, p. 198). Stam's data suggest that L1 English speakers may also produce "some path gestures on the verb or subject and verb with none on the satellite" (personal communication, December 29, 2015). Additionally, speakers of S-languages are found to have three options in expressing manner (when produced linguistically), known as manner modulation: 1) they may downplay manner, 2) they may conflate manner, or 3) they may conflate manner and path. In option one, the speaker may choose to downplay manner, i.e., they express manner linguistically on the verb and produce a co-occurring path gesture on a satellite — or the speaker might entirely downplay manner, in which case manner is neither lexicalized nor gestured. In option two, the speaker expresses manner through both a manner verb and a co-occurring manner gesture. Finally, in option three, speakers of S-languages may conflate manner and path in gesture, which is expressed across both a manner verb and satellite. Examples of such co-speech gesture expressions are provided in (7) to (9) below from McNeill and Duncan (2000) and Choi and Lantolf (2008).

(7)EN- [but it **rolls**] him out
Hand wiggles: manner information.
(8)EN- [and he rolls…**down** the drain spout]
Hand plunges straight down: path information only.
(McNeill & Duncan, 2000, p. 150)
(9)EN- [**ro**<**o**>**lls down** the street]
RH, cupped with a palm facing down, moves down to lower right, circular motion: manner + path information.
(Choi & Lantolf, 2008, p. 213)

As shown in (7), the speaker draws into focus the manner of motion that occurs in the event. In this excerpt, the salience of manner is expressed by both the speaker's production of the verb *rolls* and the co-occurring manner gesture — an example of manner conflation. In (8) the speaker downplays the manner through which 'he' (Sylvester) descends through the drainpipe, by expressing manner in speech and synchronizing a path gesture with a satellite. Through the co-expression of path in speech and gesture, the speaker seems to find the downward trajectory to be the most salient feature in the scene, which expresses a type of manner modulation. Finally in (9), the speaker expresses the salience of both the manner and the path, which are conveyed through the tightly packed unit of gesture and verbal meaning where the stroke begins on the verb *rolls* and extends through the satellite *down*.

V-languages on the other hand express motion events in an alternative fashion. To indicate path, the stroke phase of the gesture tends to co-occur with path verbs, ground NPs, or both, whereas manner is conveyed lexically through basic manner verbs, adverbials, gesture, or both, if it is mentioned at all. For V-languages a different synchronization occurs, namely a manner gesture may co-occur with a path of motion verb, a combination of path of motion verb and ground NP (as well as an optional adverbial) — referred to as a manner fog. Additionally, V-language speakers may express manner on the verb (but usually through a basic manner verb) with a co-occurring path gesture.

(10)BRP- [e ele **rola** pru]
[and he rolls towards]
[RH@waist fingers gripping LH, releases and moves outward in sweeping motion to @right of hip: path information]

(11)BRP- [e **sai** rolando]
[and exits rolling]
[RH@CC fingers extended, palm at body makes sweeping motion to @far right: path information]

(12)BRP- [**e sai pela rua . rolando com a** bola de boliche]
[and exits through the road rolling with the bowling ball]
[BH@CC, fingers extended, palms facing body make seven rotations, remaining @ CC: manner information]

In (10) the speaker of Brazilian Portuguese (abbreviated for convenience as BRP) expresses manner through the main verb *rola* 'roll' and produces a synchronized representation of the path that the figure travels along through a sweeping gesture. Under McNeill's (2005) framework, this is a type of manner modulation, much like the downplay of manner optionally indicated in S-language types. This, however, is uncommon in V-language types. Again, the downplay of manner is posited to evidence weaker salience of manner of motion on behalf of the speaker in S-languages, but this arguably shows stronger salience in V-languages, as manner is not typically coded on the verb. In (11) the speaker encodes manner on an adverbial, but expresses the salience of path of motion through the co-production of a path in both lexical and gestural form, which is quite common of V-languages. In (12) the speaker produces a manner fog through production of a manner gesture across both a path verb, ground NP, and adverbial. Indeed as the numerous options of manner and path GPs above display, the expression of salient features in motion events within and across language typologies can be quite dynamic.

Gesture Unit

The work of Kendon (1972; 1980) has identified larger units of extended gesture usage that span across larger pieces of discourse. Kendon named these periods of non-retracted gesture production Gesture Units (G-unit, henceforth). Within McNeill's framework, a number of GP's

might occur within a G-unit, and as such are accounted for as a successive and referential narrative, in which the point of reference is realized in the interaction of the gesture and speech co-occurrences. Moreover, the production of multiple GPs within a G-unit should be understood as a cumulative description of thought.

In example of the connection of GPs within one G-unit, one of the participants in the current study conveys the idea of Tweety Bird throwing a bowling ball into a drainpipe and the ball descending through the pipe. While initially representing both the bowling ball with her left hand (loosely cupped) and the drainpipe with her right hand (narrowly cupped), the participant states "and he puts the bowling ball into the pipe." During the first GP, the speaker brings her left hand to her right hand in an arced trajectory. In co-speech gesture production of the bowling ball being put into the pipe, the speaker makes available a reference point in the iconic (McNeill, 2005) gestures representing both the bowling (left hand) and the pipe (right hand). Without retracting her gesture, the participant extends the referential gestures produced in the first GP to a second GP in stating "and it goes down the pipe," during which she continues to represent the pipe with her right hand (narrowly cupped) and the bowling ball (now narrowly cupped) with her left hand. During the second GP, she shows the path of the bowling ball as it descends through the pipe.

In accounting for both the imagistic and linguistic semiotic features produced by the speaker at the unfolding GPs within the G-unit, we are provided with a greater description of the speaker's complete mental image of both the ball being put into the pipe, the manner in which the ball was put in the pipe, and the path the ball took while descending through the pipe.

Path Accumulation

Another extension of the GP is realized in the accumulation (or lack thereof) of path across language typologies. What is meant by the accumulation of path in speech alone is that speakers of S-languages can accumulate path satellites that express boundary crossing under one verb. In McNeill's (2005) work, an English speaker provides a narration of a cartoon stimulus in (13) below.

(13) 1 and it goes down
2 but it rolls him out
3 down the rain spout
4 out
5 into the sidewalk
6 into a bowling alley and he knocks over all the pins.
(McNeill, 2005, p. 197)

In the speaker's description, they produce five satellites on one verb (*rolls*). The verbs that the speaker produces describe four instances of boundary crossing. When analyzing the gesture, in addition to speech, McNeill (2005) notes that the speaker produces five, segmented path gestures in co-occurrence with the five path satellites, as indicated by the underlined words in (13).

V-languages in comparison to S-languages do not allow for the accumulation of satellites expressing boundary crossing on a single verb, rather a new verb must be provided (or the verb must be restated) for each mention of boundary crossing. Likewise, the co-speech gesture in V-language expression of boundary crossing differs from that of S-language speakers, in that V-language speakers often mark the trajectory of the figure in a single, curvilinear gesture.

Manner & Path Co-Speech Gesture Research

A number of studies in the past 20 years have analyzed L2 TFS patterns of speakers within same and different language typologies (e.g., Özyürek, 2002; Kellerman & van Hoof, 2003; Negueruela, Lantolf, Rehn Jordan, & Gelabert, 2004; Stam, 1998, 2006, 2008, 2010, 2015). These studies have sought to understand whether speakers are able to shift from L1, V-

language TFS to L2, S-language TFS, as is the focus in this study. Additionally, a number of studies by Stam (1998, 2006, 2008, 2010, 2015) have focused on observing the shifts in TFS during SLA. Indeed Stam (1998) suggested that through an analysis of the co-production of speech and gesture by L2 learners, "we can see evidence of language transfer as well as the beginning of 'thinking-for-speaking' in English" (Stam, 1998, p. 618).

In both Kellerman and van Hoof (2003) and Negueruela et al. (2004) findings showed that Spanish speakers of L2 English were unable to entirely make the shift to L2 TFS, as indicated by their gesture. In Negueruela et al. (2004) they found that speakers had made the shift for path but not in manner.

Stam (2006) found that L2 learners (L1 Mexican Spanish) did not accumulate path components, as is typical in English. The learners, however, showed increases in their production of gestures with path satellites and verb+satellite, much like L1 English speakers. In analysis of path and manner, Stam (2008) concluded that the advanced L2 learners (L1 Mexican Spanish) in her study used the same verbs as L1 English speakers, with a few cases of manner verbs. Additionally, the gestures that accompanied speech in the L2 seemed to follow those of L1 English speakers in path, but not manner production; moreover, the co-speech gestures frequently lacked conflation in comparison to those of L1 Mexican Spanish. Stam concluded that the participants had not yet fully made the shift to L2 TFS. In Stam (2006, 2008), she concluded that learners were in a transitional period between the L1 and L2 TFS patterns, and that a further analysis of highly proficient learners would yield a better understanding of the developmental trajectory learners follow in SLA.

Drawing on one highly proficient participant from Stam (2006), Stam (2010) elicited data using the same stimulus. She analyzed manner as well as path in co-speech gesture production.

The study found that after nine years (1997-2006) of 50/50 use of Spanish and English the L2 speaker had begun using more satellites in speech and had begun gesturing more frequently with ground NPs and less frequently with verbs; in addition, the speaker's gestures covered more speech and were less segmented. The learner, however, showed no signs of a shift in how she expressed manner linguistically nor gesturally. Choi and Lantolf (2008) also found that Korean L2 learners of English followed L2 TFS patterns for encoding path, but did not show signs of a shift in TFS when expressing manner in the L2.

In Stam (2015), data was again elicited from the same participant as in the previous studies (Stam, 2006, 2008, 2010). According to the data, the participant was found to follow L2 TFS patterns when discussing path both linguistically and gesturally, though she sometimes expressed path on the verb. The speaker also began to produce manner in some ways similar to that of the L2, in that she began "following the English pattern of rarely having manner gestures without manner in speech" (Stam, 2015, p. 96). Indeed, as proposed, the learners' shift towards L2 TFS patterns are observable through gesture-speech production — while the learner's TFS did not shift entirely to the L2 pattern, a gradual change had occurred. Moreover, Stam (2015) concluded that SLA, as analyzed through TFS in motion events, seems to fall on a continuum in which learners acquire path patterning in both gesture and speech first and manner (if at all) second.

Finally, Özyürek (2002) is the only study to my knowledge that has discovered L2 TFS shifts in manner patterns. In her study, Özyürek analyzed the shift in TFS in L1 Turkish—a V-language— to L2 English. As shown in her quantitative data, L2 English speakers (L1 Turkish) expressed manner-path conflation in both gesture and speech. On the other hand, no qualitative data were provided, so it is unclear with which part of speech the stroke of the gesture fell, i.e.,

whether the manner-path conflated gesture coincided with the manner verb and path satellite. According to Stam's (2015) proposed ease of feature acquisition, one might assume that the L1 Turkish participants in Özyürek (2002) had fully made the shift to L2 TFS in motion event scenarios.

To summarize the research studies above, there seems to be a trend towards the development of L2 TFS patterns when describing path. On the other hand, there is little evidence provided regarding the development of L2 TFS patterns in the description of manner.

Method

Participants

The two participants in this study are L2 English speakers, whose L1 is Brazilian Portuguese. Both participants are drawn from a larger set of data collected in the summer of 2015 that consists of 15 participants in total — 10 L2 English (L1 Brazilian Portuguese) speakers, and 5 Brazilian Portuguese monolinguals. In line with Choi and Lantolf (2008, p. 202), English monolingual participants were not recruited for the larger data set, as "robust data on L1 English narrations elicited by the same stimulus cartoon is readily available in several published works by McNeill and his colleagues (e.g., McNeill & Duncan, 2000)".

Of the larger group of participants, only two were selected for analysis in this study, due to their production of manner-path conflated gestures. The first participant, Alice (pseudonym), is a Ph.D. student in architecture at a large research university in the northeastern region of the U.S. She had lived in the U.S. for little over a year at the time of data collection. Danerys (pseudonym), the second participant, is a Ph.D. candidate in biochemistry at a private university in southern Brazil. Danerys had made brief visits to both the Republic of Ireland for one month

of language study, during which she lived with an L1 English family; Danerys also visited the U.S. for one month on vacation.

Data Collection

Following McNeill's (1992) data elicitation process, all participants were prompted to watch a cartoon stimulus (described below). The participants were asked to remember the cartoon in as much detail as possible, as they would provide a post hoc narration of the cartoon's contents to a native speaker of the language of narration (a co-researcher and myself), which would be audio-visually recorded. Participants were not informed that the study concerned gesture or TFS. The bilinguals were prompted to recall the cartoon in both English and Brazilian Portuguese, whereas the monolinguals were cued simply to recall the cartoon. All narrations were provided immediately after the participants had watched the cartoon stimulus.

The participants were sourced from two different settings — an immersion and a non-immersion setting (a large U.S. research university and a private Brazilian university). The data were collected in the two different language environments in order to account for the exposure to L1 English production of speech and co-speech gesture. In the recruitment process additional requirements were made with regard to participant eligibility. In order to ensure that participants were above a certain level of English proficiency[2], a minimum score of 80 on the Internet-based TOEFL was required. Finally, all participants in the immersion group were expected to have lived in an English-speaking country for at least one year.

[2] The point value listed is also an admission requirement for non-native English speakers at the same U.S. university where five of the participants were recruited.

Cartoon Stimulus

The stimulus used to elicit narrations was a Tweety Bird cartoon titled *Canary Row* (Freleng, 1950), which has been widely used in motion event studies (e.g., McNeill, 1992, 2005; Stam, 2006). The contents of the clip are as follows: In the clip, Tweety Bird (a canary) sits in his cage on the windowsill of a tall brick building singing an Irish folk song. Meanwhile, Sylvester (a cat) on the street below climbs up the building on the outside of a drainpipe attached to the building. When Sylvester reaches the top, he moves his finger to the tune of Tweety Bird's song. When Tweety Bird notices this, he flies out of the cage and into the apartment and Sylvester follows. At that point, we see Sylvester flying out of the window and hear a crash signifying that he has landed on a group of trashcans. We then see Granny (Tweety Bird's owner) shaking an umbrella and scolding Sylvester. The next scene begins with Sylvester deliberating how he will capture Tweety Bird. He then notices the drainpipe he had previously climbed. This time, however, he decides to climb up inside of the pipe. He enters the pipe and begins to climb. Tweety Bird notices this and retrieves a bowling ball from the apartment and drops it down the drainpipe from the top — the drainpipe is oddly not attached to anything. The ball falls through the pipe until it collides with Sylvester, at which point we see the pipe bulge out. We then see Sylvester launch out of the pipe with a much rounder figure. We assume that he has ingested the bowling ball. Sylvester then helplessly rolls down the street and into a bowling alley, where he crashes into the pins as indicated by the sound of the event.

Coding

The audiovisual recordings were used in the analysis of semantic coding, as well as co-speech gesture occurrence. In order to provide a precise transcription of the recordings, the clips were watched at a quarter of the regular speed using VLC media player, and the transcriptions

were rigorously checked for their faithfulness to the recording. In this study, I focus on the speakers' production of GPs (synchronization of speech and gesture), as this is the point in the motion event that is most salient to the speaker. More specifically, I analyze the occurrences of manner-path conflation in gesture, which conveys both the manner of motion, while simultaneously conveying change of location.

The transcription methods used in the analysis of the data are adopted from McNeill (1992, 2005), as provided in the appendix. The methods that McNeill has prescribed identify a number of sequences in the gesture phrase. The gesture phrase itself is defined as the full extent of the gesture; from the initial movement of the hand from a resting position to its return to a resting position. As explained above, within the gesture phrase a number of phases can be defined—preparation, pre-stroke hold, stroke, post-stroke hold, and retraction. Of the phases in the gesture phrase only one is required, namely the stroke.

Data Analysis

The analysis considers the GPs produced by two participants during the portion of their narrations in which they describe Sylvester exiting from the drainpipe and descending the street towards the bowling alley. L2 narration of this point in the scene is of particular interest for the analysis of the participants' verbal and co-speech gesture production, as it is at this point in the narration of the scene that participants tend to mark a clear distinction between V-language and S-language TFS patterns and reveal their underlying conceptual understanding of the features in a motion event.

When narrating this scene, speakers of S-languages will commonly mark the salience of manner and path by producing a manner-path conflated gesture in concert with a manner verb and a path satellite (as described in (9) above). On the other hand, speakers of V-languages do

not note the salience of manner and path in a condensed unit as in S-languages, rather V-language speakers tend to convey path only or manner only gestures simultaneously with path verbs. This is optionally followed by an adverbial conveying manner, when manner is not co-expressed gesturally with the verb.

While much research has been conducted on L1 V-language speakers' shift towards L2 S-language TFS patterns (e.g., Özyürek, 2002; Kellerman & van Hoof, 2003; Negueruela, et al., 2004; Stam, 1998, 2006, 2008, 2010, 2015), only one study (Özyürek, 2002) has found advanced L2 speakers of S-languages to be able to produce manner-path conflated gestures with a manner verb—again, no transcriptions were provided. Contrary to Özyürek (2002), the majority of studies focusing on shifts in TFS patterns show that L1 V-language participants have yet to internalize L2 S-language meaning when expressing manner-path conflation in speech and gesture, let aside manner of motion. In a similar vein to Özyürek's (2002) findings, the analysis of the participants' narrations below indeed evidences the production of manner-path conflated gestures. The data here, more specifically, are qualitatively analyzed for the co-speech gesture production of the participants. The GPs formed in these gesture-speech productions, however, are typical of neither L1 Brazilian Portuguese nor L2 English TFS.

In excerpt 1, Alice—an immersion learner of English—begins the description of Sylvester's path of motion by stating the motion verb *goes* and initiates the preparatory phase of a gesture, to which she quickly reorients and produces a manner-path conflated gesture in line 35. This manner-path conflated gesture in particular coincides with both the newly stated figure *ball* and the verb *rolls*. Of particular importance, here, is that the verb, with which the gesture syncs and creates a GP, conveys manner. Bearing in mind that V-language speakers highly

infrequently produce complex manner verbs such as *roll,* Alice's production of a manner verb, alone, evidences a shift towards L2 TFS.

A further, noteworthy indication of a shift towards L2 TFS is provided in the manner-path conflated gesture that co-occurs with the manner verb *roll*. The production of this type of GP is much like the pattern found in S-language TFS, in that both manner and path are present in gesture, and manner is present in speech (as detailed in (9)). What does not resemble L2 TFS patterns in the GP produced by Alice in line 35 is the lack of a path satellite following the manner verb. As noted earlier, the production of a manner verb followed by a path satellite in S-language TFS seems to be such a tightly packed unit that the two seem to be conceptualized as a single unit of thought. Alice's production of a manner-path conflated gesture in concert with the verb roll, on the contrary, seems to account for an alternate conceptual understanding of the verb *roll*, i.e., *roll* carries the meaning of both the manner 'roll' and the non-descript path taken by the figure. Such a construal of conceptual meaning seems to form a new type of modulation that Lantolf and Sweigart have coined as a "path fog" (personal communication between Lantolf and Sweigart, January, 26, 2016), i.e., the production of a manner verb with a coinciding manner-path conflated gesture—in such a co-speech gesture production, path is present in gesture, but is not present in speech. This term derives from McNeill's manner fog (2005). The meaning provided at the GP indicates towards hybrid TFS patterns. It is plausible that this meaning might have arisen for the reason that Alice abruptly stops to indicate the location of the ball inside of Sylvester's stomach, which is a salient feature in her view of the motion event.

In line 33 Alice begins the narration of this event by placing Sylvester as the figure of the motion event. After initiating a basic motion event frame, she shifts to a manner of motion event in line 35, in which the bowling ball is placed in the figure position, then indicates the location of

the bowling in relation to Sylvester in line 38. Following this description, Alice indicates her understanding of the scene not as a standard motion event, rather a caused manner of motion event in line 41. Here, Sylvester is defocused by the energy imparted by the bowling ball and made the patient of the event. Such a view of this particular scene as a caused manner of motion event has, to my knowledge, yet to be documented and might have implications for the type of TFS that occurs during the narration of the event.

Due to Alice's background as an architectural engineer, her view of the world is seen through the lens of scientific knowledge and causes a more rigid, methodical way of narrating events that might be less salient to the average speaker. Indeed, what unfolds in the narrated scene is a caused motion event, as Sylvester (the patient) is carried down out of the pipe and forced down the hill into a bowling alley by the energy imparted by the ball (the agent). What is shown in the scene, however, defies the laws of the natural world, as the rolling of the bowling ball that occurs within Sylvester does not cause Sylvester himself to roll, rather Sylvester remains in an upright position, as he is brought down the hill by the ball. As such, Alice's description of the scene is rather precise.

From the preparation phase in line 33 to the post-stroke hold in line 41, Alice's hands do not retract, thus indicating a larger G-unit consisting of three GPs. The three GPs in total provide a greater understanding of Alice's complete mental image during her narration of the scene, i.e., the manner of the ball rolling along with its non-descript path, the location of the ball inside of the cat, and the descript path of the ball and the cat in a caused motion event. Being that Alice deploys a number of GPs without periods of retraction separating each, Alice seems to be working through her L2 narration of the event. Certainly she knows what she has viewed in the scene, but the part that becomes difficult is how one encodes the information witnessed onto the

semantic features of an L2. Moreover, Alice must question how one encodes the more rigid scientific information within the scene onto the semantic features of an L2. To be sure, it seems that Alice is not only thinking for speaking, rather she also seems to be thinking during speaking, i.e., she works through her L2 description through private speech (Flavell, 1966) — the type of speech produced with function of helping oneself regulate thinking during a cognitively burdening task— as well as communicating her message to an interlocutor through social speech (Vygotsky, 1987).

Excerpt 1: Alice

Figure 2.1 Line 33: "he goes um"

33 and (.) [he goes <u>um:</u>]
34 [BH@Chest move to @L-Shldr,fingers extended, palms facing right]

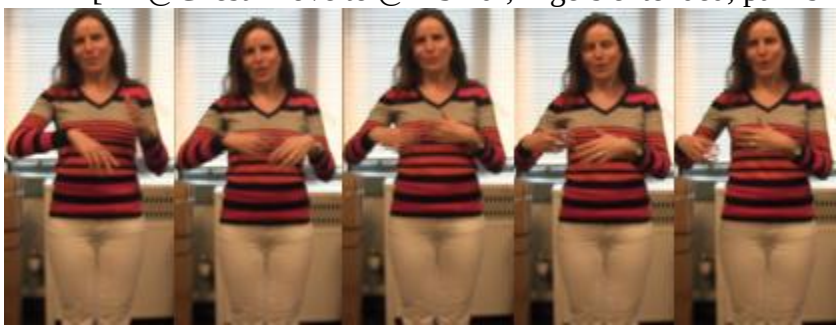

Figure 2.2 Line 35: "the ball rolls"

35 [**the ball rolls**]
36 [BH@L-Shldr, fingers extended, palms facing right, move in 3 rotations (end over end)
37 to @R-Arm, fingers extended, palms facing @ body: manner + path information]

Figure 2.3 Line 38: "inside of him"

38 [**inside of hi:m**]
39 [BH in semi-cupped form, palms facing upward; RH@R of R-hip; LH@Stomach;
40 BH move in a circular motion vertically: manner + deictic information]

Figure 2.4 Line 4: "him down a"

41 [and bringing **hi:m down a** sloped street]
42 [BH@Chest; RH palm facing downward, fingers extended moves through
43 @Stomach to @ far right of R-waist; LH palm facing center, fingers extended
44 moves through @Stomach to @R-waist]

In excerpt 2, Danerys— a non-immersion learner — produces the verb *goes* followed by the path satellite *into*, with which she deploys a path gesture (line 49). In the following line (line 53), Danerys produces the ground noun phrase *the street* with a coinciding manner gesture. Danerys then states the adverbial *rolling* which is accompanied by a manner-path conflated gesture.

Excerpt 2: Danerys

Figure 3.1 Line 49: "into"

49 and it goes [**into** <u>the</u>]
50 [RH@R-shldr moves to @R-hip; LH@L-shldr moves
51 to @L-hip; BH palms facing at body to palms
52 facing down: path information]

Figure 3.2 Line 53: "the street"

53 [**the street**]
54 [BH make a rotating motion; RH@R-hip; LH@L-hip: manner information]

Figure 3.3 Line 55: "rolling"

55 [**. ro:lling**]
56 [RH@R-hip moves outward making three rotating motions
57 to @in front of chest: manner + path information]
58 until it reaches a bowling station and
59 it made a sound of a strike

Following the verb *goes* (line 49), Danerys produces a path gesture on the path satellite *into*.

While the verb and satellite on line 49 adhere to lexical patterns in both V- and S-languages,

Danerys' co-expression of path follows L1 English TFS patterns. In line 55, Danerys expresses

Sylvester's manner of motion in gesture on the ground NP *the street*, which follows neither L1 nor L2 TFS patterns—her gesture is formed in one full rotation by both hands. Danerys then forms a manner-path conflated gesture with her right hand; starting close to her chest, her hand makes three small outward rotations that depict Sylvester's rolling descent of the street.

In excerpt 2, Danerys produces a number of GPs in one clause. While only one verb is produced, Danerys' creation of three different GPs seems to paint an entirely different picture than that provided in speech, which is not typical of V-language speakers. With the stroke phase on line 49, she marks the path traveled by Sylvester in co-production with the satellite *into*, which matches L1 English TFS patterns. Following this, Danerys deviates from both L1 and L2 TFS patterns by indicating manner on the ground NP *the street* (line 53). In line 55, Danerys produces a manner-path conflated gesture on the adverbial *rolling*. While the use of manner gestures has been observed in co-production with adverbials in V-languages, the production of manner-path conflated gestures have not been observed. This GP alone conveys the understanding of the adverbial *rolling* as expressing path in addition to manner, which is indicative of neither L1 nor L2 TFS patterns. This GP, on the other hand, does portray a conceptual understanding, on behalf of Danerys, that the expression of motion events in the L2 are not completed in the same way as the L1, thus shifting her conceptualization of semantic boundaries.

As regards the amount of GPs employed in one clause by the speaker, McNeill (2005) has provided a brief discussion on the implications of GPs across language types through the example of path accumulation by S-language speakers. Such accumulations are instantiated through the multiple uses of satellites with co-occurring gesture on one verb. Indeed such examples do exist in the accumulation of path, but the GPs in excerpt 2 display a different type

of focus during online thought, namely one directed at the salience of multiple features (not just path). The accumulation of GPs under one G-unit in excerpt 2 seem to build on one another in succession until the GP is formed that truly expresses the visuospatial image in the speaker's thinking. That being said, Danerys seems to be somewhere in between L1 and L2 TFS. Moreover, this working through of the scene verbally and gesturally is indicative of private speech; therefore Danerys seems to be thinking during speaking, not thinking for speaking.

Discussion and Conclusion

In both of the excerpts above, the speakers produce TFS patterns that are unlike those of the L1 or L2. Rather than expressing one type of TFS or the other, the speakers' production of GPs indicate movement towards hybrid L1-L2 TFS patterns. Given that the speakers produced GPs on two different parts of speech, yet producing manner-path conflated gestures, they indicate different conceptual understandings of the manner of motion *roll*. The learner's conceptual and personal understanding (or *smysl*, as Vygotsky called it), as contextually formed meaning, are rendered visible at the site of the GP. Moreover, through additional analysis of the multiple GPs within both G-units, it appears that the participants are working through their explanation of the narration in private speech. The production of private speech (or thinking during speaking) in the excerpts above is representative of the shifting of semantic boundaries from the L1 with the goal of L2 semantic alignment/development.

To summarize, Alice's GP in excerpt 1 indicates that she has not made the shift to L2 TFS, but she indeed seems to be in transition to L2 TFS. Evidence for this transitional phase is provided in Alice's production of a manner-path conflated gesture that coincides (at least partially) with the manner verb *rolls*—beginning on *the ball*. While this does not conform tightly to L1 English TFS, it is quite distinct from L1 Brazilian Portuguese TFS. Again, Alice's

understanding of the verb *roll* as a manner-path conflated verb, as expressed through co-speech gesture, marks a new type of GP, namely a path fog.

Within the larger G-unit, Alice's production of multiple GPs is indicative of her use of private speech and, as such, her developmental process of conceptually understanding L2 TFS patterns. Danerys, also, seems not to have made the shift to L2 TFS, but is arguably on her way as evidenced by the fact that she employed a manner-path conflated gesture, although it does not coincide with a verb. Building on the claims in Stam (2015) these patterns indicate that the speakers are making a transition from L1 TFS to L2 TFS. With further longitudinal data, a second analysis would yield a better understanding of the type of progress the participants have made in their shift in L2 TFS, if they indeed continue to.

Conceptual Development as a Remedy

For L2 instructors and researchers, this area of cross-linguistic research provides both an interesting and intricate understanding of target difficulties in the acquisition of a L2. Indeed the literature and this study provide evidence of learners' difficulties with such nuanced features in the acquisition of a L2 (especially L2 S-language speakers of a L1 V-language background) through the GP — as noted, the GP is the site at which a speaker's local, conceptual understanding is rendered visible. However, to date there has been little done to mediate learner's conceptual understanding of these nuanced features in the L2. Though some have argued for the importance of gesture in L2 teaching; but how do we teach it and, furthermore, how do we assess it?

One possibility is not to teach gestures directly, but to teach the concept of manner along with manner verbs. If these are mastered, then it might well follow that learners will use manner-path conflated gestures appropriately—an indication of a shift in TFS from L1 to L2. Certainly,

the acquisition of a L2 requires somewhat major shifts in conceptual understanding and the re-positioning of semantic boundaries (Gullberg, 2009). Given that the difficulty lies in the shift of conceptual understanding in the L2, implementing concept-based language instruction (CBI or CBLI) (Negueruela, 2003) would presumably remedy this problem.

References

Aske, J. (1989). Path predicates in English and Spanish: A closer look. *Proceedings of the 15th annual meeting of the Berkeley Linguistics Society*, 1-14.

Brown, A., Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *SSLA, 30,* 225-251.

Choi, S., & Lantolf, J. P. (2008). Representation and embodiment of meaning in L2 communication: Motion events in the speech and gesture of advanced L2 Korean and L2 English speakers. *Studies in Second Language Acquisition, 30* (2), 191-224.

Flavell, J. (1966). La langue privé. *Bulletin de Psychologie, 19,* 698-701.

Freleng, I. (Director). (1950). *Canary Row* [Animated film]. New York, NY: Time Warner.

Gullberg, M. (2009). Reconstructing verb meaning in a second language How English speakers of L2 Dutch talk and gesture about placement. *Annual Review of Cognitive Linguistics, 7*(1), 221-244.

Kellerman, E., & van Hoof, A. M. (2003). Manual accents. *IRAL, 41,* 251-269.

Kendon, A. (1972). Some relationships between body motion and speech. *Studies in dyadic communication, 7*(177), 90.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. *The relationship of verbal and nonverbal communication*, *25*, 207-227.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago University Press.

McNeill, D. (2001). Imagery in motion event descriptions: Gestures as part of thinking-for-speaking in three languages. *Proceedings of the Twenty-Third Annual Meeting of the Berkeley Linguistics Society, 23*, 255–267.

McNeill, D. (2005). *Gesture and thought.* Chicago, IL: University of Chicago Press.

McNeill, D. & Duncan, S. (2000). Growth points in thinking-for-speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 141-161). New York, NY: Cambridge University Press.

Negueruela, E. (2003). A sociocultural approach to teaching and researching second language: Systemic-theoretical instruction and second language development. Unpublished doctoral dissertation: The Pennsylvania State University.

Negueruela, E., Lantolf, J. P., Rehn Jordan, S. R., & Gelabert, J. (2004). The "private function" of gesture in second language speaking activity: A study of motion verbs and gesturing in English and Spanish. *International Journal of Applied Linguistics, 14,* 113-147.

Negueruela, E., & Lantolf, J. P. (2008). The dialectics of gesture in the construction of meaning in second language oral narratives. In McCafferty, S., & Stam, G. (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 88-106). New York, NY: Routledge.

Özyürek, A. (2002). Speech-gesture relationships across languages and in second language learners: Implications for spatial thinking for speaking. In B. Skarabela, S. Fish, & A. H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 500– 509). Somerville, MA: Cascadilla Press.

Papafragou, A., Massey C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98,* B75-B87.

Pourcel, S., & Kopecka, A. (2006). Motion events in French: Typological intricacies. Unpublished ms., University of Sussex & Max Planck Institute for Psycholinguistics, Brighton, UK & Nijmegen, The Netherlands.

Slobin, D. (1996a). From 'thought and language' to 'thinking for speaking'. In S. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70-96). New York, NY: Cambridge University Press.

Slobin, D. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195-219). Oxford: Clarendon Press.

Slobin, D. (2003). Language and though online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157-192). Cambridge, MA: MIT Press.

Slobin, D. I. (2004). The many ways to search for a frog. *Relating events in narrative. Typological and contextual perspectives,* 219-257.

Stam, G. (1998). Changes in patterns of thinking about motion with L2 acquisition. In Santi, S., Guaïtella, Cavé, C., et Konopczynski, G. (Eds.) *Oralite et gestualite: Communication multimodale, interaction,* (pp. 615-620). Paris: L'Harmattan.

Stam, G. (2006). Thinking for speaking about motion: L1 and L2 speech and gesture. *IRAL-International Review of Applied Linguistics in Language Teaching, 44*(2), 145-171.

Stam, G. (2008). What gestures reveal about second language acquisition. In McCafferty, S., & Stam, G. (Eds.), *Gesture: Second language acquisition and classroom research* (pp. 88-106). New York, NY: Routledge.

Stam, G. (2010). Can an L2 speaker's patterns of thinking for speaking change?. In Z. Han & T. Cadierno (Ed.), *Linguistic relativity in SLA: thinking for speaking* (pp. 59-83). Bristol, UK: Multilingual Matters.

Stam, G. (2015). Changes in thinking for speaking: A longitudinal case study. *The Modern Language Journal, 99 (S1),* 83-99.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. *Language typology and syntactic description, 3,* 57-149.

Talmy, L. (2000). *Toward a cognitive semantics: Typology and process in concept structuring.* (Vol. 2). Cambridge, MA: MIT Press.

Tomasello, M. (2003). *Constructing language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

Vygotsky, L. (1987). *The collected works of L. S. Vygotsky, Vol. 1: Problems of general psychology. Including the volume thinking and speech.* R. W. Reiber & A. S. Carton (eds.). New York: Plenum Press.

Zlatev, Jordan & Peerapat Yangklang (2003). A third way of travel: the place of Thai in motion-event typology. In Strömqvist & Verhoeven (Ed.), *Relating events in narrative: Typological and contextual perspectives.* (pp. 159-190). UK: Psychology Press.

Appendix

Gesture and Speech Coding

Gesture	Speech
Adapted from McNeill's (1992) convention	Adapted from Jefferson Jefferson (1984) convention
gesture phrase [preparation <u>pre-stroke hold</u> **stroke** <u>hold</u> retraction]	: elongation of a syllable (.) micro-pause in speech
LH left hand RH right hand BH both hands CC center center (@ at body part)	

McNeill's (2005) Space Manikin