4-11-2013

# The Effect of Motion Type and Modality in Word Learning in English

Samantha N. Emerson
*Georgia State University*, semerson2@student.gsu.edu

THE EFFECT OF MOTION TYPE AND MODALITY IN WORD LEARNING IN ENGLISH

by

SAMANTHA N. EMERSON

Under the Direction of Şeyda Özçalışkan and Gwen Frishkoff

ABSTRACT

Speakers of satellite-framed languages (S-languages such as English) express manner more frequently than speakers of verb-framed languages (V-languages such as Spanish) because S-languages use "satellite" phrases to encode path, leaving the verb free to encode manner (Talmy, 1985, Slobin, 2004). Gestures have also been shown to follow these cross-linguistic differences. While numerous studies have examined the effect of manner and path expressions on cognition cross-linguistically and in V-languages, less is known about these effects within S-languages. The current study examines encoding of path and manner events in English using a novel word-learning paradigm. Our results show that English speakers are less accurate at identifying words for manner than path—after controlling for the effects of learning and similarity of event pairs—regardless of the modality of learning (speech only *vs*. speech+gesture). Overall, our results suggest a path advantage in word learning even for S-language speakers.

INDEX WORDS: Talmy's typology, Motion events, Thinking for Speaking, Word learning, Gestures

THE EFFECT OF MOTION TYPE AND MODALITY IN WORD LEARNING IN ENGLISH

by

SAMANTHA N. EMERSON

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Masters of Arts

in the College of Arts and Sciences

Georgia State University

2013

THE EFFECT OF MOTION TYPE AND MODALITY IN WORD LEARNING IN ENGLISH

by

SAMANTHA N. EMERSON

Committee Chair: Şeyda Özçalışkan

Gwen Frishkoff

Committee: Chris Conway

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2013

## ACKNOWLEDGEMENTS

I would like to take this opportunity to thank several people whose efforts have been essential for the project. I thank my advisors Şeyda Özçalışkan and Gwen Frishkoff for making me revisit every statement I made again and again (and again and again and again) until everything was just right. I thank Michael Owren and Chris Conway for their help at various stages of my project. I thank Iria Romay-Fernández for taking me into her home (not once but twice!) and for the numerous hours she sacrificed to help me with my project. I thank Neil Gordon and Taylor Brooks for their technical help in developing the video and audio stimuli. I thank my partner Ryan Haggard for keeping me sane and for helping me with the making of the instructional videos. Last but not least, I thank my parents, Julie and Al Emerson, who raised me right and trusted in me. I am also grateful to the Language and Literacy initiative at Georgia State University for providing a fellowship that funded this work.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# 1 INTRODUCTION

We know from previous work that speakers of various languages differ in their expression of different components of spatial motion. Speakers of verb-framed languages (hereafter **V-languages**; e.g., Spanish) tend to express path of motion using verbs (e.g., *enter the house*). By contrast, they express manner (i.e., how one moves) less frequently, and when they do, manner tends to be encoded with optional constituents, such as adverbs (e.g., *enter the house slowly*) or subordinate verbs (e.g., *enter the house bouncing*). Speakers of satellite-framed languages (hereafter **S-languages**; e.g., English) express manner routinely in the verb (e.g., *run, crawl*) and typically use optional constituents outside the verb—i.e., "satellites"—to express path (e.g., *out, down*; Slobin, 2004; Talmy, 1985, 2000).

These cross-linguistic differences in lexicalization patterns also have consequences for speakers' underlying representations of motion events in these two groups of languages as evidenced in comparison and categorization tasks. As shown in previous work, S-language speakers have a stronger bias for manner of motion compared to speakers of V-languages and V-language speakers have a stronger bias for path of motion compared to the manner dimension of motion events (Gennari, Sloman, Malt, & Fitch, 2002; Kersten et al., 2010; Naigles & Terrazas, 1998; Slobin, 1996). By contrast, little is known about the relative sensitivity to path and manner events by speakers of S-languages. According to some researchers (Cardini, 2010; Papafragou, Massey, & Gleitman, 2002), English speakers should be more sensitive to manner than to path, because the verb is an "informationally privileged element" (Papafragou et al., 2002) and is often used to express manner in English but rarely encodes path. According to others (Gennari et al., 2002), English speakers should be equally sensitive to manner and path because both features are easily encoded in English, albeit by different linguistic devices (i.e., verb vs. satellite). Recent work suggests a third possibility: English speakers may be more sensitive to path than to manner because path is expressed more frequently when the total number of motion verbs and satellites are considered (i.e., when each instance of a motion verb and each instance of a motion satellite are combined; e.g., Emerson, Çörekli, & Özçalışkan 2013). In fact, Talmy (2000) suggests that for all languages, path is an obligatory component of each overt expression of a translational motion event, whereas overt expression of manner is optional.

The goal of the current paper was to determine whether English speakers would show different patterns of word learning performance for events encoding path versus manner variations. We predict three possible patterns that English speakers might follow in learning words for motion event types: (1) manner of motion bias, (2) neither manner nor path of motion bias, or (3) path of motion bias. We also ask whether word learning is affected by modality. More specifically, we predict that English speakers will learn labels for events more readily when those labels are accompanied by iconic gestures that depict the manner or path component of the event than if the event labels are presented in the absence of such gestures. We test these predictions by using a learning paradigm that involves pairing pseudo-words with novel motion events. Events were presented as animated clips that depict different paths or different manners.

The thesis is structured as follows: We first outline the typology of motion events described by Talmy (1985, 2000) and explain how it may be framed within the Thinking for Speaking (Slobin, 1996) hypothesis. We then examine how the gestures of speakers from typologically distinct languages vary in language-specific ways. We next introduce the word learning paradigm and describe the methods and the results. We conclude the paper by discussing how these results relate to and extend prior work. We discuss some limitations of the study and offer recommendations for future research on cognitions related to motion events within and across languages.

# 2 BACKGROUND

## 2.1 Typology of Motion Events

The world's languages can be divided into two types based on their prototypical ways of expressing the different components of motion (Talmy, 1985, 2000). In ***verb-framed languages (V-languages),*** such as Spanish, the verb is typically used to express path of motion (e.g., *entrar* 'enter'; see example 1). Conversely, in ***satellite-framed languages (S-languages***), such as English, the path of motion is typically expressed outside of the verb, usually with a particle or a preposition, while the verb is reserved for the expression of manner of motion as in example (2):

(1) *El perro <u>entra</u> a la casa*

'The dog <u>enters</u> the house'

(2) *The dog* ***runs*** *<u>into</u> the house*

This cross-linguistic difference, however, reflects tendencies, not absolutes. Speakers of V-languages can—and sometimes indeed do—express manner of motion. This can be achieved by appending an optional manner adverbial or a subordinate clause, as in (3). However, because manner of motion requires the addition of an optional phrase or clause in V-languages, it is often omitted in V-languages (Slobin, 2004; Özçalışkan & Slobin, 1999; Naigles, Eisenberg, Kako, Highter, & McGraw, 1998).

(3) *El perro entra a la casa* ***corriendo***

'The dog enters the house **running**'

By contrast, in S-languages, verbs encode manner of motion as a default, leading to a wide variety of manner words that capture a broader and more fine-grained set of distinctions than that which is encoded by the limited set of manner verbs available in V-languages (Slobin, 2004). For instance, English speakers can express variations on jumping by using many different verb types, including *hop, skip, jump, leap, bound, bounce, pop, bob, vault,* and *spring,* while Spanish speakers have only three words to express variations on jumping—*brincar* 'jump', *saltar* 'jump', and *botar* 'bounce'. Furthermore, while S-languages have a greater variety of manner words and use them with greater frequency than V-language speakers, intra-typological variation also exists within S-language speakers with some S-languages (e.g., Dutch, German, English) producing manner at a greater rate and with greater variety than other S-languages (e.g., Mandarin, Thai, and Russian; Slobin, 2004).

In summary, compared to speakers of V-languages, speakers of S-languages express manner more frequently, using a greater number and variety of manner verbs.

## 2.2 Thinking and Talking about Motion in Typologically Distinct Languages

Cross-linguistic differences in the expression of motion may have implications for cognition, particularly during the verbalization of that motion event. According to the theory of linguistic relativity (i.e., the Sapir-Whorf hypothesis), language can shape or constrain our cognition (Sapir, 1929; Whorf, 1956). A strong version of this theory asserts that if a particular feature (e.g., a color, shape, or property of an event) cannot be verbalized, the speaker will be unable to conceptualize that feature. In this way, all forms of cognition would be restricted to only that which can be expressed in a particular language. Little if any research has validated this strong claim (see Chui, Leung, & Kwan, 2007 for a review).

The weaker view of linguistic relativity states that language biases speakers to focus more on features of the world that are more easily encoded verbally. For example, while the range of colors is actually a continuum, all languages divide it into discrete categories with some dividing it into fewer categories than others. Berlin and Kay (1969) found that basic color words ranged from two to 11 terms, and more recently Roberson, Davies, Corbett, and Vandervyver (2005) found that this number could range up to 12. Despite this division of color expression, these studies found that speakers frequently chose to arrange various color chips into a greater number of groups than basic color words in their language. These distinctions seemed to be based more on universal principles of perception and went against the strong version of linguistic relativity hypothesis. At the same time, Roberson et al. (2005) found that speakers did follow language specific patterns in how they chose to group the color chips. For example, those languages with separate words for blue and green were more likely to divide chips of those colors into different groups while those languages with only one term did not, showing a weaker influence of language on speakers' perceptions of color similarity.

An alternative version of this theory is the 'Thinking for Speaking' (Slobin, 1996) hypothesis. This hypothesis proposes that it is the act of verbalizing an event that causes the speaker to conceptualize it in a manner that is consistent with their language. Consequently—and in contrast to the weak version of linguistic relativity—without some form of verbalization of the event, there should be no influence of language on cognition. For example, Slobin (1996) notes that an event may have numerous features, such as definiteness (e.g., "**a** dog" vs. "**the** dog"), tense (e.g., “the dog **ran**” vs. “the dog **runs**”), progressive aspect (e.g., “the dog **is running**”), boundedness (e.g., “the dog **ran**”, which implies a beginning and end to the action, vs. “the dog **was running**”, which does not), perfectivity (e.g. in Spanish “el perro **corrí**” in which the dog ran at a particular moment vs. “el perro **coria**” in which the dog used to run), evidentiality (e.g., “the dog **ran**” vs. “the dog **must have ran**”), and visibility (e.g., in Turkish “kopek kaç-ıyor-**du**” in which the dog’s running was witnessed vs. “kopek kaç-ıyor-**mus**” in which the dog’s running is assumed). When we decide to speak about an event, we must make decisions about which features of the event will be linguistically encoded. Slobin proposes that the decision to verbally express certain features while omitting others is based on the relevance of a feature to the conceptualization of the event and also to how readily encodable that feature is in one’s language. This means that language may attune speakers towards encoding certain features of an event while leaving other features out. Therefore, in order to be able to verbalize an event, a speaker is likely to attend more to features that are linguistically easier to encode in their native language.

Recent work has examined how these two theories (i.e., the weak version of linguistic relativity and the Thinking for Speaking hypothesis) might apply to cognitive differences in perceiving and remembering various components of a motion event in speakers of typologically distinct languages. Two prior studies (Gennari et al., 2002; Papafragou et al., 2002) examined memory for motion events using a variation on the classic old/new memory task. Participants were asked to view a series of motion events with both manner and path components. After a delay period (i.e., approximately 10 to 20 minutes or two days, respectively), they were presented with a second set of motion events; half of these events matched the initially observed motion events (i.e., old stimuli) and the other half were altered so that either the path or the manner was different than the original stimulus (i.e., new stimuli). The participants' task was to identify whether each motion event was old (i.e., one that they had seen before) or new (i.e., one that they had *not* seen before; see Table 1 for additional details). The authors' prediction was that speakers would make more false alarms (i.e., incorrectly classify a new stimulus as old) if the stimulus feature that was altered was one that was not routinely encoded in their native language (e.g., speakers of V-languages should be more able to recognize a change in path than a change in manner). Contrary to both theories, neither study reported language-specific effects in participants' identification of events as new or old—even in the conditions where participants produced a verbal description during the first viewing of the stimuli. These results thus suggest that cross-linguistic differences in the expression of motion may not influence speakers' long-term memory for motion events.

Several other studies have used forced choice similarity judgments to explore motion event processing among speakers of S- and V-languages (Cardini, 2010; Gennari et al., 2002; Naigles & Terrazas, 1998; Papafragou et al., 2002). In these studies, participants were presented with motion events that were characterized by path and manner cues. After each event presentation, they viewed two alterations of the same event and were asked to choose the one that was "most similar" to the original event (see Table 1 for additional details). In each case, one of the choices involved an alteration of path (i.e., same-manner), and the other involved an alteration of manner (i.e., same-path). The authors hypothesized that speakers should choose the variant that was unchanged in the feature that they were more likely to express (e.g., speakers of V-languages should choose same-path events because they are more likely to verbalize path than manner). The studies by Cardini (2010) and Papagfragou et al. (2002) showed that speakers of S- and V-languages were equally likely to choose the manner or path alterations as being more

*Table 1. Methodological Differences in Old/New Memory and Forced Choice Similarity Tasks*

| | Papafragou et al. (2002) | Gennari et al. (2002) | Cardini (2010) | Naigles & Terrazas (1998) |
|---|---|---|---|---|
| **Languages** | V-Language: Greek<br>S-Language: English | V-Language: Spanish<br>S-Language: English | V-Language: Italian<br>S-Language: English | V-Language: Spanish<br>S-Language: English |
| **Stimuli** | Photographs | Video Clips | Video Clips | Video Clips |
| **Cognitive Task** | Memory Task<br>Similarity Task | Memory Task<br>Similarity Task | Similarity Task | Similarity Task |
| **Verbal Task** | Verbal description of target photograph for Memory Task, but not for Similarity Task | Participants either (1) produced a verbal description of the video, (2) repeated nonsense syllables, or (3) simply view videos prior to tasks | Verbal description of target video after completion of the task (i.e., no benefit of linguistic description on task) | No verbal descriptions were made by participants. Experimenter made prompts syntactically manipulated to be (1) manner-biasing, (2) path-biasing, or (3) "neutral" |
| **Language-Specific Results** | No | (1) Yes<br>(2) No<br>(3) Reversed pattern | No | Yes<br>Also syntactic effect for conditions (1) and (2) |

similar to the original event when *not* asked to verbalize the original event prior to making their similarity judgment. In contrast, Gennari et al. (2002) found that when participants were asked to produce a verbal description of the event before making their decision, Spanish speakers (V-language) were more likely to choose same-path events as being more similar to the original event than English (S-language) speakers. This suggests that for Spanish speakers path was the more salient feature. Similarly, Naigles and Terrazas (1998) also found the same language-specific pattern for participants' judgments when the original video was presented with a verbal description containing a novel nonsense verb. These latter two studies indicate the conceptualization of an event (as reflected in judgments of similarity) is affected when verbalizations are present during the encoding process—but not when verbalizations are absent. Thus it would appear that the act of verbalizing an event has short-term effects on cognition—at least for judgments of similarity—for motion events, thereby providing support for the Thinking for Speaking hypothesis over the weak linguistic relativity hypothesis.

### *2.2.1 Learning labels for motion events in typologically distinct languages*

Two studies have examined how adult speakers of typologically distinct languages learn new motion words with an emphasis on how the motion events are lexicalized. The first was Naigles and Terrazas (1998) who, in a similarity task like those discussed above, showed participants motion videos paired with verbal descriptions (i.e., spoken sentences) that contained novel pseudo-verbs (e.g., ,*kradding*, *mercando*). The descriptions encoded the motion event using one of three grammatical constructions: (1) **path biasing** constructions**,** which involved a *transitive frame* for English (e.g., "She's *kradding* the tree") and either a *transitive frame* or an *intransitive frame with a generic path preposition* for Spanish (e.g., "Ella está *mercando* <u>a</u>l árbol" 'She is *mercing* <u>to/on/at/from/by/in</u> the tree'); (2) **manner biasing** constructions**,** which involved an *intransitive frame with a path-rich preposition* (e.g., in English: "She's *kradding* toward the tree,"; in Spanish: "Ella está *mercando* hacia el árbol" 'She is *mercing* toward the tree'); or (3) **neutral** constructions**,** which involved an *intransitive verb with no prepositional phrase* (e.g., in English: "She's *kradding*"; in Spanish: "Ella está *mercando*" 'She is *mercing*'). After hearing these statements, participants were asked to choose which of two subsequent videos (changed in either manner or path) depicted the same type of motion (i.e., *kradding*/*mercando*) as the previous video. The results showed language-congruent performance, with English speakers showing a preference for same-manner videos and Spanish speakers showing a preference for same-path videos. However, at the same time, the study also showed that the type of linguistic description (i.e., path-biasing, manner-biasing, or neutral) also influenced speakers' choices with participants choosing the same-path videos more often than same-manner videos in response to path biasing descriptions and same-manner videos more often in response to the manner biasing and neutral descriptions. In a separate study, Feist (2010) suggested that the reason why the "neutral" statements (e.g., "She's *kradding*") also elicited a greater number of same-manner responses was because the statements included an explicit subject (i.e., "she") and because the presence of a path and ground tend to be correlated (and both were absent in these "neutral" statements), manner was the only remaining feature that the verb could be used to express.

The second study that examined novel word learning for motion events was by Kersten et al. (2010). In this study English (S-language) and Spanish (V-language) speakers were asked to learn how to classify videos of bug-like creatures into categories with either a pseudo-verb, a pseudo-noun, or a numeral. While the creatures in each video differed in both their appearance (e.g., color, body shape, number of legs) and motion (i.e., path and manner) in relation to another stationary creature, participants had to decipher through trial and error that the creatures could only be categorized by either their path or manner. Kersten et al. (2010) found that English speakers were quicker than Spanish speakers to identify a creature's category when the relevant feature was manner but were equally as quick as Spanish speakers when the relevant feature was path. Interestingly, the category labels (i.e., pseudo-verbs, pseudo-nouns, or numerals) did not affect the rate at which participants were able to learn to assign creatures to path and manner categories.

This latter finding contradicts the predictions of Papafragou et al., (2002) who state that the prominence of the verb and its likelihood of encoding manner in English is likely to attune speakers to attend more to manner than to path, the latter of which is encoded outside the verb. Instead the results of Kersten et al. suggest that the syntactic class in which a motion feature is expressed is relatively unimportant. Yet, at the same time, the study by Naigles and Terrazas indicates that other linguistic cues, such as syntactic framing, may contribute to the conceptualization of a motion event especially when learning labels for such events, suggesting some linguistic features—but not all—may bias learners' interpretations of the meaning of novel words.

## 2.3 Gesturing about Motion in Typologically Distinct Languages

Language-specific patterns also become evident in gesture in one of two possible ways. One set of studies has shown that speakers use gestures to express information that is not present in speech (McNeill, 2000; McNeill & Duncan, 2000). For example, McNeill (2000) found that Spanish (V-language) speakers often use a manner gesture concurrently with phrases expressing path in their speech.

In contrast, another set of studies has shown that the information contained within the gestures is actually complementary to speech (Gullberg, Hendriks, & Hickmann, 2008; Kita & Özyürek, 2003; Kita et al., 2007; Özçalışkan, 2012; Özyürek & Kita, 1999; Özyürek et al., 2008; Özyürek, Kita, Allen, Furman, & Brown, 2005). For example, information that is omitted from speech (e.g., manner of motion) is often omitted from the speakers' gestures as well (Gullberg et al., 2008; Kita & Özyürek, 2003; Kita et al., 2007; Özyürek & Kita, 1999; Özyürek et al., 2008; Özyürek et al., 2005). Gestures also reflect syntactic features, such as the degree of clause coherence (Givón, 1989, 2001). For example, when path and manner are expressed in the same clause (as is common in the S-language English), gestures tend to conflate both elements. However, when the elements are expressed in separate clauses (as is common in the V-languages; e.g. Japanese and Turkish), the gestures tend to express either path or manner only (Kita & Özyürek, 2003; Kita et al., 2007; Özçalışkan, 2012; Özyürek & Kita, 1999; Özyürek et al., 2005). Furthermore, these language specific patterns are evident from age nine onwards (Özyürek et al., 2008). These typologically consistent findings support the idea that patterns in gesture are closely associated with patterns in speech in the expression of motion events in a particular language.

## 2.4 The Current Study

Prior cross-linguistic research has demonstrated that when verbalizations are present during the task, language may have a short-term effect on the perception of the different components of motion events: Similarity judgment tasks show clear differences between speakers of typologically distinct languages and in the performance of speakers in V-languages. However, these studies do not provide a consistent view on how the weak theory of linguistic relativity or the Thinking for Speaking hypothesis could be applied to the typology for motion events within speakers of S-languages. In this study, we focus on English—an S-language—and predict three possible trajectories in word learning in this language:

1. *Manner of motion bias*. Participants may show better learning of manner than path words. This could suggest a cognitive bias towards manner over path. Given that verbs in English tend to encode manner rather than path, this could also support the claim that the verb is an "informationally privileged element" (Papafragou et al., 2002), which has consequences for learning words for motion events.

2. *No bias*. There may be no difference in the learning of manner versus path words in English. This finding could suggest that because both features are encoded in speech, both must be equally represented in cognition. Alternatively this could indicate that if any biases exist, they do not affect motion word learning.

3. *Path of motion bias.* Participants may show a bias to learn path words more easily than manner words. This finding could suggest that what matters for motion word learning is the absolute rate of expressing path versus manner rather than the linguistic mechanism (verb vs. satellite) that is typically used to encode these features in English.

In order to further examine the effect of motion type on cognition in S-language speakers, we have performed a novel word (i.e., pseudo-word) learning study for different instances of path and manner variations. This paradigm has several advantages: First, the study implements the use of pseudo-words, which have been shown to be a sufficient way to evoke the verbal representations of motion events in participants without biasing them towards path or manner (Kersten et al., 2010; Naigles & Terrazas, 1998). Second, these pseudo-words have been designed with English phonemes using an ambiguous "inflectional marking" such that participants would not be able to identify the syntactic class of words. The use of such native or native-like phonemes has been shown to assist in word learning (e.g., Ellis & Beaton, 1993; Papagno, Valentine, & Baddeley, 1991; Service & Craik 1993), and the use of ambiguous syntactic class assignment enables the avoidance of biases based on lexicalization expectancies (Kersten, 1998).

Finally, motion type effects in word learning could reflect one or more underlying cognitive processes. In particular, working memory (e.g., Baddeley and Hitch, 1974) may affect word learning in two ways: through maintenance of visual representations (in our task, the image of the motion event to be learned)—i.e., the "visuo-spatial sketchpad"—or through verbal rehearsal (in our task, maintenance of novel words for each motion type)—i.e., the "phonological loop." In the present task, we can rule out differences in phonological working memory across conditions because participants received the same set of pseudo-words. Therefore, if working memory plays a role in our task, it is more likely to reflect the relative ease of encoding or maintaining path versus manner images (visuo-spatial sketchpad). These effects could in turn reflect either low-level differences in the perception of path and manner events or (more likely) higher-level difference in perception due to differences in top-down attention (expectancy). In particular, speakers of S-languages might expect that novel words for motion events will be more likely to encode either path or manner (see the "Thinking and Talking about Motion in Typologically Distinct Languages" section above) and may therefore attend more to certain features of the motion event being held in visuo-spatial working memory. This latter theory is even more likely to be the case given that attention is known to affect word learning across the lifespan—in infants and young children (e.g., Baldwin, Markman, Bill, Desjardins, & Irwin, 1996; Hirotani, Stets, Striano, & Friderici, 2009; Mathers, 2006; Redmond, 2004; Samuelson & Smith, 1998; Tomasello, 1995) as well as adults (e.g., Alt & Gutmann, 2009; Yu, Zhong, & Fricker, 2012).

Finally, given the importance of gestures in communication (e.g., Clark, 1996; Goldin-Meadow, 2004; Kendon, 2004; McNeill, 1985, 1992) and the typologically consistent information they express, we examined the effects of modality (i.e., speech-only, speech+gesture) on learning behavior. We predict that training with gestures will further help participants identify the relevant feature of the motion events thereby improving the participants' ability to learn the associated pseudo-word when those words are accompanied by gesture.

# 3 METHODS

## 3.1 Participants

The participants included 89 native adult English speakers ($M_{age}$ = 20.53, range = 18-46, 19 males), all college students in Atlanta. Six participants were excluded because of computer error during data collection resulting in a total of 83 participants ($M_{age}$ = 20.45, range = 18-46, 16 males). Of the remaining 83 participants, 22 were in the path and speech+gesture conditions ($M_{age}$ = 21.48, range = 18-46, 4 males), 20 were in the path and speech-only conditions ($M_{age}$ = 19.2, range = 18-22, 5 males) 22 were in the manner and speech+gesture conditions ($M_{age}$ = 21.32, range = 18-42, 4 males), and 20 were in the manner and speech-only conditions ($M_{age}$ = 19.65, range = 18-24, 3 males).

## 3.2 Stimuli

### *3.2.1 Motion animations*

Data were collected by using animations of a star-shaped character that moved in relation to a stationary spherical object (see Fig.1); these animations were adapted from Pulverman, Hirsh-Pasek, Golinkoff, Pruden, & Salkind (2008). The stimuli animations consisted of 12 path variations with the same manner (see Appendix A for all path variations) and 12 manner variations with the same path (see Appendix B for all manner variations). Each animation lasted two seconds and was repeated four times. All stimuli were created using Strata Design 3D CX 6 software.

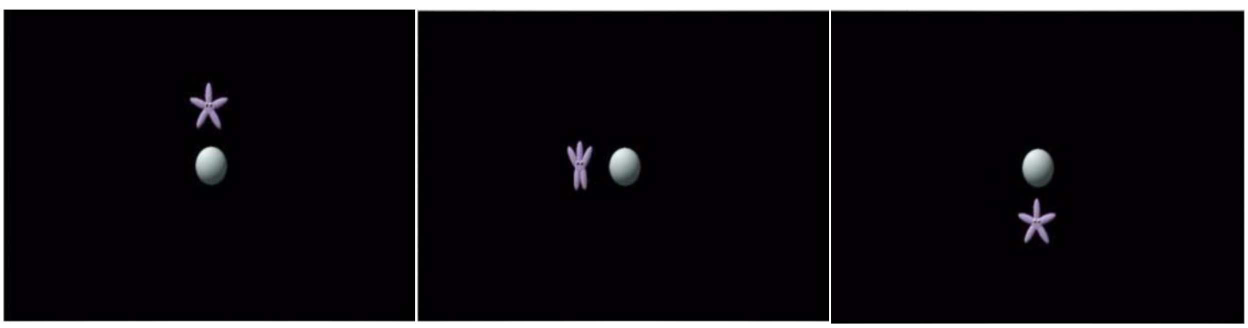


***Figure 1.*** *Three screen shots of a motion animation (i.e., path-Blomu).*

In order to control for the level of difficulty participants had in distinguishing pairs of animations within a motion type, data were collected from a separate group of 72 adult native English speakers ($M_{age}$ = 20.44; range = 18-39; 19 males). Fifteen participants were excluded due to computer error during data collection resulting in a total of 57 participants ($M_{age}$ = 20.11, range = 18-39, 15 males). Half of the participants (N = 28, $M_{age}$ = 20.43, range = 18-34, 8 males) viewed every possible pair of either path animations (i.e., path condition), and the other half (N = 29, $M_{age}$ = 20.43, range = 18-39, 7 males) viewed every possible pair of manner animations (i.e., manner condition); all stimulus pairs were displayed in a random order divided into four blocks.

After the presentation of each stimulus pair, participants were asked to rate on a seven point Likert scale how similar the two animations were to one another and then to indicate how confident they were in their rating also on a seven point Likert scale. The scales ranged between "Not at All Similar", "Not Very Similar", "Moderately Similar", and "Identical" for similarity ranking and "Not at All Confident", "Not Very Confident", "Moderately Confident", and "Very Confident" for the confidence ranking. The rankings on each scale were counterbalanced so that for half of the participants "Identical" and "Very Confident" were indicated by a 7 and for the other half by a 1. Responses were then rescaled for analysis so that a 7 would indicate "Identical" and "Very Confident" for all participants.

On average, the participants' mean similarity rating was 3.28 (SD = 0.71) for path pairs and 3.03 (SD = 0.74) for manner pairs ($t(55) = 1.30$, $p = .198$, ns); their average confidence rating was 6.49 (SD = 0.51) for path pairs and 6.50 (SD = 0.61) for manner pairs ($t(55) = -0.03$, $p = .977$, ns). These results thus suggest that participants in the path and manner conditions viewed the pairs of stimuli as being not very similar; they were also very confident in their responses. Overall, these results confirmed that participants did not view the stimuli of one motion type as being more difficult to distinguish (i.e., more similar) than the other.

#### *3.2.2 Instructional videos*

The word learning training was conducted through the use of instructional videos. Each instructional video was recorded using a Sony HDR-HC9 Handycam Camcorder and centered on the torso of a male instructor wearing a long sleeved black shirt against a dark backdrop. The actor recited the pseudo-word label for the video's associated motion animation twice. For half of the participants (speech-only condition), the actor remained stationary while reciting the pseudo-words (Fig. 2a); for the other half (speech+gesture condition), he accompanied the pseudo-word with an iconic gesture that depicted the manner or the path of the motion in the stimulus (Fig. 2b).

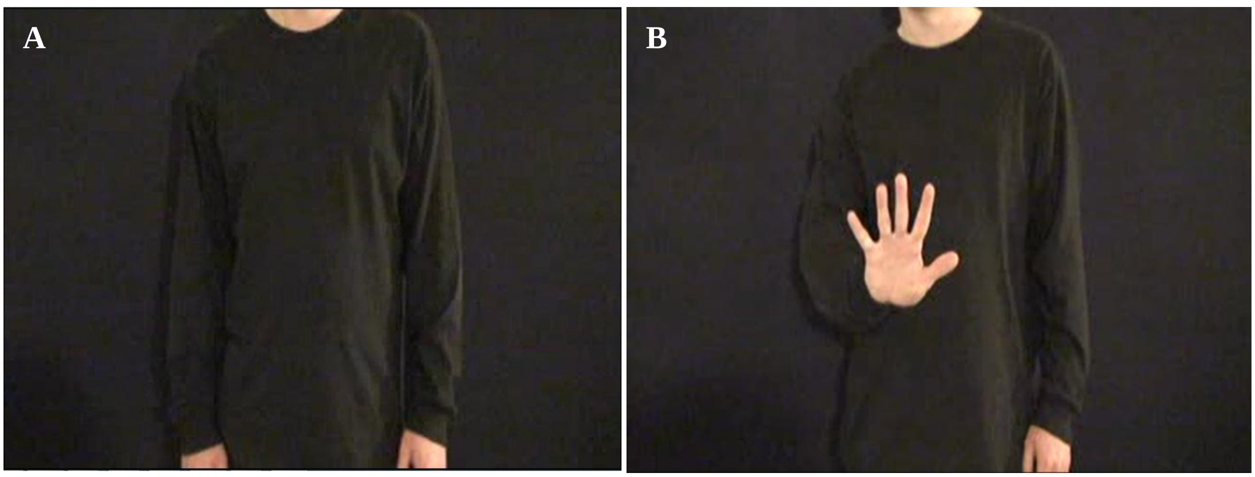



***Figure 2.*** *Screen shot of instructional videos for (a) the speech-only condition and (b) the speech+gesture condition.*

### *3.2.3 Pseudo-words*

Twelve pseudo-words were created for use in both the path and manner conditions. All pseudo-words were disyllabic and were matched in orthographic and phonological features. Additionally, all pseudo-words ended with the phoneme /u/. This was intended to mimic inflectional morphemes appended to nouns (i.e., plural /s/) or verbs (i.e., present tense /s/) for the fictional "new" language participants were learning. Pseudo-words were recorded during the instructional videos and converted in WAV audio files using Final Cut Pro for presentation during testing.

## 3.3 Procedure

### *3.3.1 Data collection*

Half of the participants were randomly assigned to the *manner condition* (different manners with the same path); the other half was assigned to the *path condition* (different paths with the same manner). Within each condition, half of the participants received the training in speech using pseudo-words only (*speech-only condition*) and the other half received the training in both speech and gesture using pseudo-words with iconic gestures (*speech+gesture condition*).

#### *3.3.1.1 Training*

Training sessions were arranged so that each participant first saw a motion animation followed by its corresponding instructional video (Figs. 3a and 3b). After the presentation of three sets, the participants were given a mini-test (Fig. 3c) in which two of the three previous animations were displayed side by side, accompanied by a pseudo-word that was the correct label for one of the animations. The keyboards had been marked with a red and yellow sticker on the "z" and "m" keys, respectively. Participants were asked to press the red key if they believed that the pseudo-word matched the animation on the left and the yellow key if they believed that the pseudo-word matched the animation on the right. These mini-tests were used to ensure that participants continued to attend to the stimuli throughout the training phase. In each condition, the training included four blocks, each of which consisted of three item pairs followed by a mini-test.

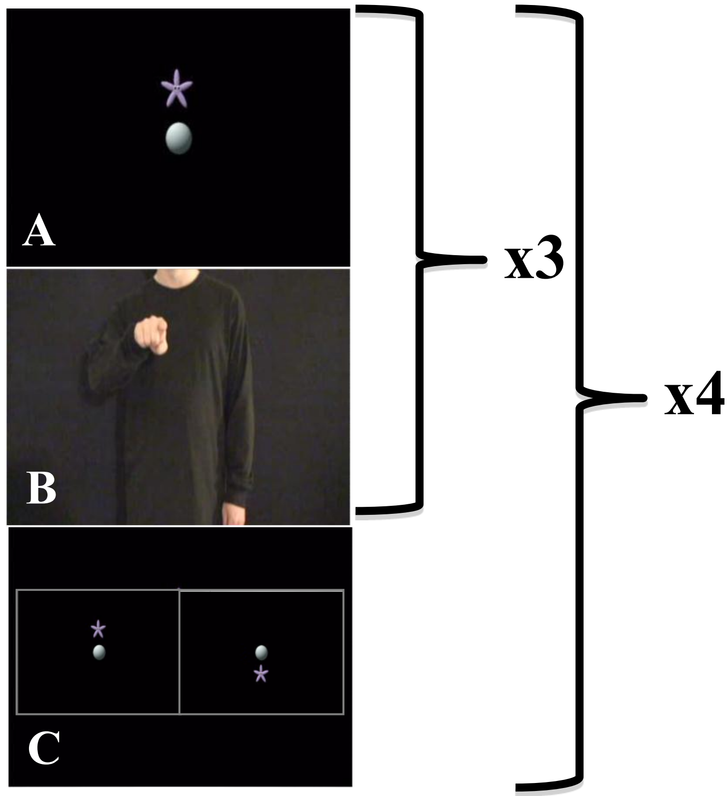



***Figure 3.*** *Order of presentation for training blocks.*[1]

*3.3.1.2 Block tests*

Upon completion of each of the four training blocks, participants were given a block test consisting of 12 trials that followed the same format as the mini-tests (but included testing of all 12 pseudo-words learned). Test trials were designed so that each pseudo-word was presented once and each motion animation was presented twice (once as the correct choice and once as the incorrect choice). Trials were given in random order, and accuracy and reaction times were recorded.

[1] Participants first saw (a) a motion animation followed by (b) its instructional video. This was repeated three times before they received (c) a mini-test, which included two of the previous three motion animations and one of the two animation's associated pseudo-word. These sets were repeated four times until all 12 variations had been displayed. The entire process was considered a single block, which was repeated four times.

### *3.3.2 Data Analysis*

We analyzed the data first by using a zero-order correlation matrix, which included the variables: *mean accuracy* (i.e., the proportion of participants who correctly matched the pseudo-word to animation), *mean similarity rating for item pairs*—based on ratings produced by an independent group of participants, *motion type* (i.e., path, manner), and *block* (i.e., one, two, three, four) to identify the relationship between all four variables.

Next we computed a hierarchical multiple regression across trials with mean accuracy as the dependent variable to determine the unique amount of variance in participants' word learning that was accounted for by motion type. In order to account for the effect of learning over time, block was entered as the independent variable in the first step. Next, to account for level of difficulty in discriminating stimuli, similarity rating was added to the model in the second step. Finally, motion type was added in the third step to examine the English speakers' abilities to learn new path and manner words above and beyond the effects of learning and discriminability between stimuli pairs. Modality was found *not* to account for a significant proportion of mean accuracy's variance and was therefore, removed from the analysis as an independent variable; accordingly all data in the hierarchical multiple regression analysis presented in the results section are collapsed across the speech-only and speech+gesture conditions.

## 4 RESULTS

As can be seen in Table 2, the mean accuracy was significantly correlated with block, $r$=.522, $p$<.001, showing an effect of learning. That is, participants continued to improve their accuracy in learning pseudo-words for motion events over the four training blocks. More specifically, for the path condition, the mean percentage of participants who answered each pair correctly in the first block was 65.45% (SD = 0.03), in the second block was 72.15% (SD = 0.03), in the third block was 78.05% (SD = 0.02), and in the fourth block was 82.72% (SD = 0.02). For the manner condition, the mean percentage of participants who answered each pair correctly in the first block was 57.74% (SD = 0.03), in the second block was 71.23% (SD = 0.02), in the third block was 74.60% (SD = 0.02), and in the fourth block was 79.76% (SD = 0.02). We found no other significant correlations between similarity, modality, or mean accuracy.

***Table 2.*** *Correlation Matrix*

| | **Mean Accuracy** | **Similarity Rating** | **Motion Type** | **Modality** | **Block** |
|---|---|---|---|---|---|
| **Mean Accuracy** | ------ | -.108 | -.137 | -.014 | **.522**** |
| **Similarity Rating** | | ------ | -.037 | .000 | .046 |
| **Motion Type** | | | ------ | .000 | .000 |
| **Modality** | | | | ------ | .000 |
| **Block** | | | | | ------ |

*p<.05; **p<.001

We next looked at the unique amount of variance the different factors contributed to the participants' mean accuracy for word learning using a hierarchical multiple regression. Table 3 displays the unstandardized coefficients and standard errors for each step of the hierarchical multiple regression. The first model indicated that block was significantly associated with mean accuracy, accounting for 40.23% of the variance ($R^2$ = .402, $F$(1, 94) = 63.26, $p$ < .001). The simple main effects suggest that the percent of participants who were able to correctly match the pseudo-word to the stimuli increased by approximately 6.4% during each consecutive block. The increase in the amount of variance accounted for by the second model was also significant indicating that similarity rating accounted for an additional 2.59% of the variance ($R^2$ = .428,

$F(2, 93) = 34.82$, $p < .001$) in mean accuracy. The simple main effects suggest that, above and beyond the effect of block (i.e., learning), for every unit of increase in similarity between the two stimuli, the number of participants answering each item correctly decreased by 1.94%. Finally, the third model also indicated that motion type significantly increased the amount of variance accounted for by an additional 3.02% resulting in a final model which accounted for a total of 45.84% of the total variance ($R^2 = .458$, $F(3, 92) = 25.96$, $p < .001$). This last model shows that, above and beyond the effect of block and similarity rating (i.e., difficulty), path items were answered correctly at a rate of 3.90% more than manner items.

***Table 3.*** *Regression of Mean Accuracy on Block, Similarity Rating, and Motion Type*

| **Variable** | ***b* (*SE*)** | **95%*CI*** | ***t*** |
|---|---|---|---|
| **Step 1 ($R^2$ = .402)** | | | |
| Constant | .568 (.022) | .525, .612 | 25.95** |
| *Block* | *.064 (.008)* | *.048, .079* | *7.95*** |
| **Step 2 ($R^2$ = .428)** | | | |
| Constant | .618 (.032) | .554, .682 | 19.08** |
| Block | .064 (.008) | .049, .080 | 8.18** |
| *Similarity Rating* | *-.019 (.009)* | *-.037, -.001* | *-2.05** |
| **Step 3 ($R^2$ = .458)** | | | |
| Constant | .639 (.033) | .574, .705 | 19.33** |
| Block | .064 (.008) | .049, .080 | 8.36** |
| Similarity Rating | -.019 (.009) | -.037, -.002 | -2.18* |
| *Motion Type* | *-.039 (.017)* | *-.073, -.005* | *-2.27** |

*p<.05; **p<.001

Overall, our results show that, when the effects of learning and item pair similarity have been accounted for, participants were more able to learn words for path events than for manner events.

# 5 DISCUSSION

The aim of this study was to determine whether English speakers would show different patterns of word learning for words that refer to manner or path of motion and, if so, whether these differences could be explained by differences in attention to the two features. We considered three possibilities:

1. *Manner of motion bias*. Participants might show better learning of manner than path words, consistent with the theory that English speakers are more sensitive to manner than to path due to the greater likelihood of manner being encoded in the verb (e.g., Papafragou et al., 2002).

2. *No bias*. Participants might show no difference in learning manner versus path words, suggesting either that both features are encoded equally well following patterns of speech or, if there is a bias, that it is does not affect motion word learning.

3. *Path of motion bias*. Participants might show a bias to learn path compared to manner verbs, as path is a core component in the expression of motion events and is expressed at a higher rate than manner of motion in English.

The results showed that English speakers have more difficulty learning words for manner than words for path. This supports a *path of motion bias* in English and is consistent with the idea that features of motion that are verbally expressed more often (i.e., at higher frequency) within a language are more cognitively salient, regardless of its syntactic packaging (verb vs.

satellite). This finding also provides support for Talmy's (2000) claim that path—but not manner—is an obligatory component of motion. Thus, verbally encoded features are not all equally represented in cognition as the *no bias* hypothesis would suggest nor are features preferentially represented based on how they are lexicalized in speech as the *manner of motion bias* hypothesis would suggest.

Most of the previous cross-linguistic research on motion and its influence on cognition have assumed that S-language speakers should perform better or equally with manner stimuli than with path (i.e., *no bias* or *manner of motion bias* hypotheses). With this assumption in mind forced choice similarity judgment tasks have gained prominence in this field. However, as Kersten et al. (2010) points out, this paradigm requires participants to choose between path and manner such that the two motion components are actively competing for salience in the minds of the speakers. Recall that in this task a participant is presented a motion event and asked to choose which of two alternate events (i.e., a same-path and a same-manner) is most similar to the original. Therefore, if a participant is given 10 of these trials and chooses the same-path events for seven trials, it must follow that the participant would have chosen same-manner for three trials. In this way, the number of path choices is directly dependent on the number of manner choices and vice versa. Because of the *path of motion bias* in S-language speakers found in the present study, forcing speakers to choose between path and manner will result in an increased number of same-path variants choices, which necessarily leads to a decreased number of same-manner choices. Speakers' mental representations of these events, however, are unlikely to be so dichotomous. The overall salience of an event's path is independent from the salience of its manner; therefore, an S-language speaker may perceive the path of an event as more salient than its manner (resulting in them choosing the same-path variant) but still find the manner of the event more salient than a speaker of V-language would. Consequently, the similarity judgment paradigm produces a good measure of relative salience between path and manner *within* language types, but comparisons of manner performance *between* S- and V-languages should result in an attenuation of the manner effect. While studies using this paradigm have in general been successful in finding this effect when using a verbal manipulation, it is likely that, in actuality, the effect is much greater than their results would suggest. Furthermore, it is even possible that manipulations to this paradigm—perhaps including designs similar to our stimulus rating procedure—would result in the emergence of an effect without verbal manipulations, supporting the weak linguistic relativity hypothesis.

Furthermore, to the best of our knowledge, this is the first study of motion expression that has included in its analysis some measure of the range of variability within motion types. While our study found that on average there were no differences in English speakers' perception of similarity between pairs of path and manner items, within each motion type pair, similarity rating ranged between 1.50 and 6.43 for path and 1.33 and 5.36 for manner on a seven point scale, suggesting high variability among item pairs. This suggests the possibility that the particular paths and manners chosen as stimuli could create misleading effects if they were not equated in terms of similarity space. For example, in our experiment path stimuli whose motion remained close to the sphere passing either in front of or behind it (i.e., *bripu*, *chulsu*, *frengu*, and *plercu* in Appendix A) were rated as very similar to one another while manner stimuli that rotated along an axis (i.e., *bripu* and *frengu* in Appendix B) were rated as very distinct. If a study were to contain only stimuli of these types, the distinctiveness of the manners could have caused these stimuli to be more salient to the learner, overshadowing any other effect the study might reveal.

Moreover, the equality in ratings between pairs of stimuli found in our stimulus rating procedure may be related to findings from similarity judgment tasks without linguistic manipulations (e.g., Cardini, 2010; Gennari et al. 2002; Papafragou et al., 2002) discussed previously. As in these prior studies, we find evidence that language-specific effects based on motion expression may not extend to nonlinguistic cognition. In contrast to previous studies, our study did not force participants to choose between path and manner variants. Consequently, the results of each motion type are independent from one another, whereas in the previous similarity judgment tasks, the results of one motion type were necessarily the complement to the results of the other motion type (e.g., for every same-path choice the number of same-manner choices must be decreased by one). However, with this advantage comes a limitation: Because manner and path variants were tested using a between subject design, the two conditions cannot be directly compared to one another. Therefore, it is possible that despite both groups of stimuli receiving the same average rating, the stimuli in the two conditions were not evenly matched. In fact, in a very similar study performed by Czechowska and Ewert (2011), monolingual and bilingual speakers of two S-languages (Polish and English) rated pairs of manner stimuli as being more similar than pairs of path stimuli when tested in a within subjects design. Whether this effect is a pervasive one or an artifact of the particular stimuli that the authors chose is a matter for further research.

A second limitation of this study was the absence of a gesture effect on word learning. Research has shown that when gestures reinforce the content of speech, they often enhance the listener's comprehension of the conveyed information (Goldin-Meadow, Kim, & Singer, 1999); this was not the case with the current study. One possible reason why modality did not influence word learning could be the gestures themselves. A unique gesture was developed for each motion animation and designed to be maximally distinct from every other gesture. This may have resulted in gestures that seemed artificial or non-intuitive to the participants. An alternative possibility is that, because either the manner (in the path condition) or the path (in the manner condition) was always held constant, the relevant feature of each motion animation was always obvious to participant. Had the motion animations been more perceptually complex—containing variations in both path and manner—gestures may have been of more use to participant by calling attention to the focal component of the event when more than one was available to choose as the referent for the pseudo-word. This would be a more ecologically valid design, simply because real world referents for words do not typically occur without a variety of other distracting features to choose from.

Another limitation to our study is the fact that only one language, English, was tested. We predict that extending this paradigm to speakers of various V-languages will likely reveal a much stronger bias towards path over manner in word learning than observed among English speakers. The reason for this is twofold: First, if—as suggested by Talmy (1985)—path is an obligatory component of motion expression, biases in path word learning should be prevalent in both languages. Second, because manner is expressed at a much lower rate than path in these languages, speakers should be less able to learn manner words than path words.

More importantly, there is need for the paradigm to be extended to speakers of other S-languages. Slobin (2004, p. 8) suggests that variations in the rate of manner expression also exists within the category of S-languages depending on various morphosyntactic structures that make certain manner expressions more or less readily accessible to the speaker. Consequently, Slobin found that the descriptions of a particular scene by speakers of some S-languages such as Dutch and German expressed manner less frequently than English speakers while speakers of others such as Mandarin and Thai (also known as serial-verb languages) and Russian used manner more frequently. Examination of additional S-languages could, therefore, elucidate the cause behind the *path of motion bias* found in this experiment. If path is an obligatory component of all motion, we would expect the *path of motion bias* to prevail in these languages as well. Alternatively, if the effect is driven by the absolute rate of path expression relative to manner expression, we would expect speakers' abilities to learn path and manner words to be modulated by their rates of path and manner expression, respectively. One study by Czechowska & Ewert (2011) examined speakers of Polish and English—two S-languages that differ in manner production with Polish speakers producing manner at a greater rate than English speakers and English speakers producing path at a greater rate than Polish speakers. This study asked speakers of these two languages to rate on a five-point scale the similarity of pairs of manner and pairs of path stimuli in an experiment similar to our stimulus rating procedure. Contrary to our findings, Czechowska and Ewert found that speakers of both languages rated same-manner pictures as being more similar. However, they also found that, similar to the speakers' linguistic patterns, Polish speakers tended to rate same-manner pictures as being more similar than did English speakers and vice versa for English speakers on same-path pictures. These results would suggest that rate of expression rather that the obligatoriness of path was the main cause for our findings. In addition, because this study lacked a verbal manipulation, these results may provide support for the weak theory of linguistic relativity over the Thinking for Speaking hypothesis not seen in previous studies.

In addition, more research could be conducted with bilinguals who speak both an S- and a V-language as well as bilinguals who speak two languages within either language type. If path is an obligatory component of motion, again we would expect to find Path of Motion biases in bilinguals across languages. However, if rate of expression is the driving force, we might expect to see more complicated patterns of results: First, it is possible that bilinguals' cognitions will reflect the patterns of linguistic expression in their first or more dominant language. A second possibility is that cognitions will follow the rates of expression for whichever language they are currently speaking. Third, bilinguals have been shown to express manner and path at rates that are intermediate between monolingual speakers of either of their languages, but rates of certain features may sometimes favor the linguistic patterns of one language over the other depending on the combination of languages (Emerson, Çörekli, & Özçalışkan, 2013; Hohenstein, Eisenberg, & Naigles, 2006; Wilson, 2005). Consequently, bilinguals' cognitions may reflect these more intermediate rates of expression of the language in which they are currently speaking. Finally, it is possible that exposure to multiple linguistic systems may heighten bilinguals awareness of those features which distinguish their two languages making them more attuned to these features. For example, Czechowska and Ewert (2011) also examined Polish to English bilinguals in their nonlinguistic similarity rating task. They found that as bilinguals' levels of English proficiency increased, so did their ratings of the same-path pictures (typical of English monolinguals). Interestingly, the most advanced English learners actually rated the same-path pictures more highly than even the English monolinguals. Unfortunately, no measures of path or manner expression were collected from either the bilinguals or monolinguals in this experiment so we cannot make and definite assertions about how these results compare to the speakers' linguistic expressions.

Finally, the general cognitive mechanisms underlying this effect must be further examined. While it is the opinion of the author that these effects may be explained by working memory and attention, it is not possible to ascertain this from the current paradigm. Future designs incorporating manipulations to these mechanisms as well as measures of individual differences are necessary to examine this possibility further.

Our findings also suggest some implications for future cross-linguistic research: (1) Studies should abstain from creating experimental designs in which participants are forced to choose between path and manner, ensuring that effects between languages are not actually the results of effects within the languages. (2) Studies should incorporate some measure of discriminability or salience for their stimuli, allowing for better control over the variance that exists between different paths and manners. (3) Further investigation on the role of gesture and other nonverbal cues for motion word learning must be conducted. (4) In order to gain a better understanding of the linguistic mechanisms behind this effect, it will be necessary to extend the current paradigm to speakers of other V- and S-languages as well as bilinguals. (5) Further investigation of the general cognitive mechanisms underlying the effects is also required.

In conclusion, our study showed that English speakers find new words referring to path of motion as more easy to learn than words referring to manner of motion—after controlling for the effects of learning and similarity of event pairs—regardless of the modality in which this learning took place. Contrary to the assumptions of former studies, this may suggest that English speakers are also more sensitive to path than manner of motion. Consequently, we recommend that future studies should account for this bias for path of motion even within S-languages when designing future cross-linguistic studies on motion events.

**REFERENCES**


Alt, M., & Gutmann, M. L. (2009). Fast mapping semantic features: Performance of adults with normal language, history of disorders of spoken and written language, and attention deficit hyperactivity disorder on a word-learning task. *Journal of Communication Disorders, 42*, 347-364. doi:10.1016/j.jcomdis.2009.03.004

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infant's reliance on a social criterion for establishing word-object relations. *Child Development, 67*, 3135-3153.

Cardini, F.-E. (2010). Evidence against Whorfian effets in motion conceptualisation. *Journal of Pragmatics, 42*, 1442-1459. doi:10.1016/j.pragma.2009.09.017

Chui, C.-Y., Leung, A. K.-Y., & Kwan, L. (2007). Language, cognition, and culture: Beyond the Whorfian Hypothesis. In S. Kitayama, & D. Cohen (Eds.), *Handbook of Cultural Psychology* (pp. 668-688). New York: The Guilford Press.

Clark, H. (1996). *Using language.* Cambridge: Cambridge University Press.

Czechowska, N., & Ewert, A. (2011). Perception of motion by. In V. Cook, & B. Bassetti (Eds.), *Language and bilingual cognition* (pp. 287-314). New York: Psychology Press.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Langauge Learning, 43*(4), 559-617. doi:10.1111/j.1467-1770.1993.tb00627.x

Emerson, S. N., Çörekli, D., & Özçalışkan, Ş. (2013). Ways of moving across spaces in a first and second language. *The 59th Annual Meeting of the Southeastern Psychological Association.* Atlanta, GA.

Feist, M. I. (2010). Motion through syntactic frames. *Cognition, 115*, 192-196. doi:10.1016/j.cognition.2009.11.011

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*, 49-79. doi:10.1016/S0010-0277(01)00166-4

Givón, T. (1989). *Mind, code and context: Essays in pragmatics.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Givón, T. (2001). Toward a neuro-cognitive interpretation of 'context'. *Pragmatics & Cognition, 9*(2), 175-201. doi:10.1075/pc.9.2.02giv

Goldin-Meadow, S. (2004). *Hearing gesture: How our hands help us think.* Cambridge: Havard University Press.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. *Journal of Educational Psychology, 91*(4), 720-730. doi:10.1037/0022-0663.91.4.720

Gullberg, M., Hendriks, H., & Hickmann, M. (2008). Learning to talk and gesture about motion in French. *First Language, 28*(2), 200-236. doi:10.1177/0142723707088074

Hirotani, M., Stets, M., Striano, T., & Friederici, A. D. (2009). Joint attention helps infants learn new words: Event-related potential evidence. *Developmental Neuroscience, 20*, 600-605. doi:10.1097/WNR.0b013e32832a0a7c

Hohenstein, J., Eisenberg, A., & Naigles, L. (2006). Is he floating across or crossing afloat? Cross-influence of L1 and L2 in Spanish-English bilingual adults. *Bilingualism: Language and Cogntion, 9*(3), 249-261. doi:10.1017/S1366728906002616

Kendon, A. (2004). *Gesture. Visible action as utterance.* Cambridge: Cambridge University Press.

Kersten, A. W. (1998). A division of labor between nouns and verbs in the representation of motion. *Journal of Experimental Psychology: General*, *127*(1), 34-54. doi:10.1037/0096-3445.127.1.34

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), 638-653. doi:10.1037/a0020507

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language, 48*, 16-32. doi:10.1016/S0749-596X(02)00505-3

Kita, S., Özyürek, A., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2007). Relations between syntactic encoding and co-speech gestures: Implications for a model of speech and gesture production. *Language and Cognitive Processes, 22*(8), 1212-1236. doi:10.1080/01690960701461426

Mathers, M. E. (2006). Aspects of language in children with ADHD: Applying functional analyses to explore language use. *Journal of Attention Disorders, 9*(3), 523-533. doi:10.1177/1087054705282437

McNeill, D. (1985). So you think gestures are nonverbal? *Psychological Review, 92,* 350-371. doi:10.1037/0033-295X.92.3.350

McNeill, D. (1992). *Hand and mind.* Chicago: University of Chicago Press.

McNeill, D. (2000). Analogic/analytic representations and cross-linguistic differences in thinking for speaking. *Cognitive Linguistics, 11*(1/2), 43-60. doi:10.1515/cogl.2001.010

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

Naigles, L. R., & Terrazas, P. (1998). Motion-verb generalizations in English and Spanish: Influences of language and syntax. *Psychological Science*, *9*(5), 363-369. doi:10.1111/1467-9280.00069

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), 521-549. doi:10.1080/016909698386429

Özçalışkan, Ş. (2012). When gesture does and does not follow speech in describing motion. In A. Biller, E. Chung, & A. Kimball (Ed.), *Supplement to the Proceedings of the 36th Boston University Conference on Language Development.* Retrieved from http://www.bu.edu/bucld/files/2012/07/ozcaliskan-36.pdf

Özçalışkan, Ş., & Slobin, D. I. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In A. Greenhill, H. Littlefield, & C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development: Vol. 2* (pp. 163-174). Somerville, MA: Cascadilla Press.

Özyürek, A., & Kita, S. (1999). Expressing manner and path in English and Turkish: Differences in speech, gesture, and conceptualization. In M. Hahn, & S. Stoness (Ed.), *The twenty first annual conference of the Cognitive Science Society* (pp. 507-512). Mahwah, NJ: Lawrence Erlbaum.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology, 44*(4), 1040-1054. doi:10.1037/0012-1649.44.4.1040

Özyürek, A., Kita, S., Allen, S., Furman, R., & Brown, A. (2005). How does linguistic framing of events influence co-speech gestures?: Insights from crosslinguistic variations and similarities. *Gesture, 5*(1/2), 219-240. doi:10.1075/gest.5.1.15ozy

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition, 84,* 189-219. doi:10.1016/S0010-0277(02)00046-X

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language, 30*, 331-347. doi:10.1016/0749-596X(91)90040-Q

Pulverman, R., Golinkoff, R. M., & Hirsh-Pasek, K. (2008). Infancts discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108*, 825-830. doi:10.1016/j.cognition.2008.04.009

Redmond, S. M. (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clinical Linguistics & Phonetics, 18*(2), 107-125. doi:10.1080/02699200310001611612

Roberson, D., Davies, I. R., Corbett, G. G., & Vandervyver, M. (2005). Free-sorting of colors across cultures: Are there universal grounds for grouping? *Journal of Cognition and Culture, 5.3*(4), 349-386. doi:10.1163/1568537057746485 36

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development, 69*(1), 94-104. doi:10.1111/j.1467-8624.1998.tb06136.x

Sapir, E. (1929). The status of linguistics as a science. *Language, 5*(4), 207-214. Retrieved from http://www.jstor.org/stable/409588

Service, E., & Craik, F. I. (1993). Differences between young and older adults in learning a foreign vocabulary. *Journal of Memory and Language, 32*, 608-623. doi:10.1006/jmla.1993.1031

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: Cambridge University Press.

Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist, & L. Verhoeven (Eds.), *Relating events in narrative: Vol. 2. Typological contextual perspectives* (pp. 219-257). Mahwah, NJ: Lawrence Erlbaum Associates.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and lexical description: Vol. 3. Grammatical categories and the lexicon* (pp. 57-149). Cambridge, England: Cambridge University Press.

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

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore, & P. Dunham (Eds.), *Joint attention: Its origin and role in development* (pp. 103-130). Hillsdale, NJ: Lawrence Erlbaum.

Whorf, B. L. (1956). Languages and Logic. In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (pp. 233-245). Cambridged, MA: MIT Press.

Wilson, N. L. (2005). *Conceptualizing motion events and metaphorical motion: Evidence from Spanish/English bilinguals.* (Unpublished doctoral dissertation). University of California, Santa Cruz.

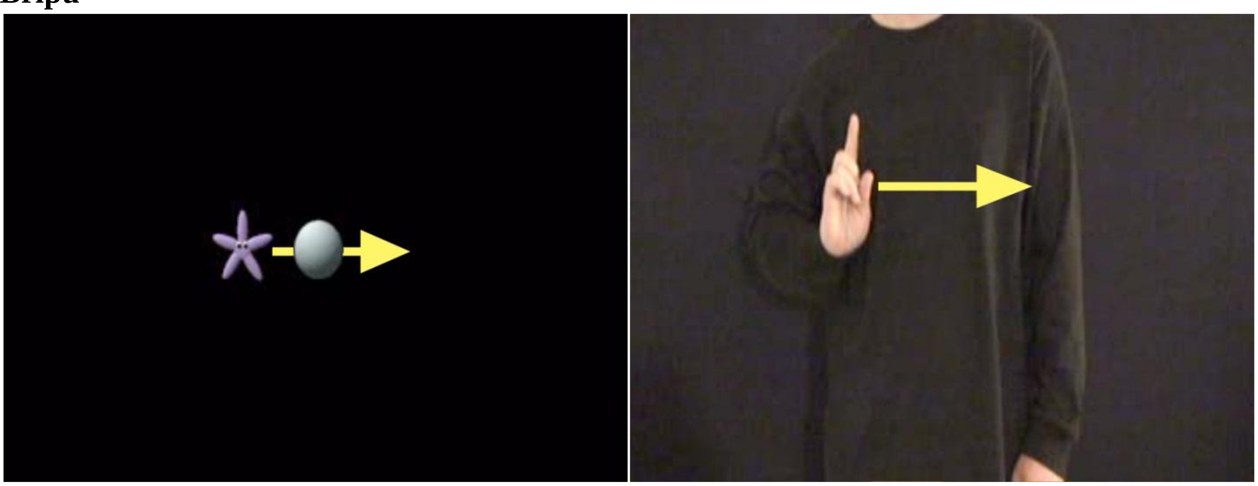
Yu, C., Zhong, Y., & Fricker, D. (2012). Selectie attention in cross-situational statistical learning: Evidence from eye tracking. *Frontiers in Psychology: Developmental Psychology, 3*, 1-16. doi:10.3389/fpsyg.2012.00148

# APPENDICES

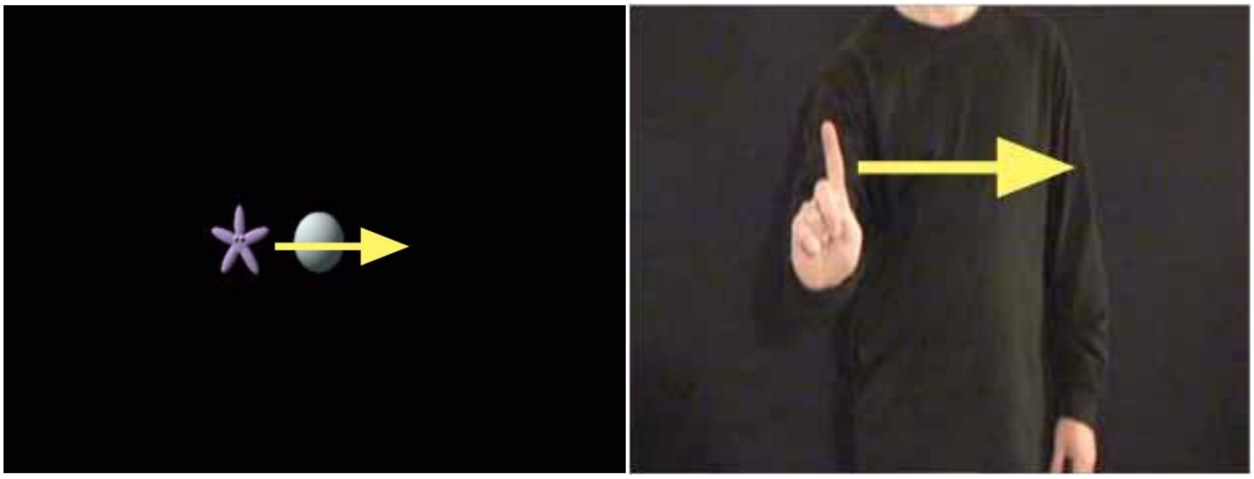
## Appendix A. Path Animations & Gesture Instructional Videos

**Blomu**

**Bripu**

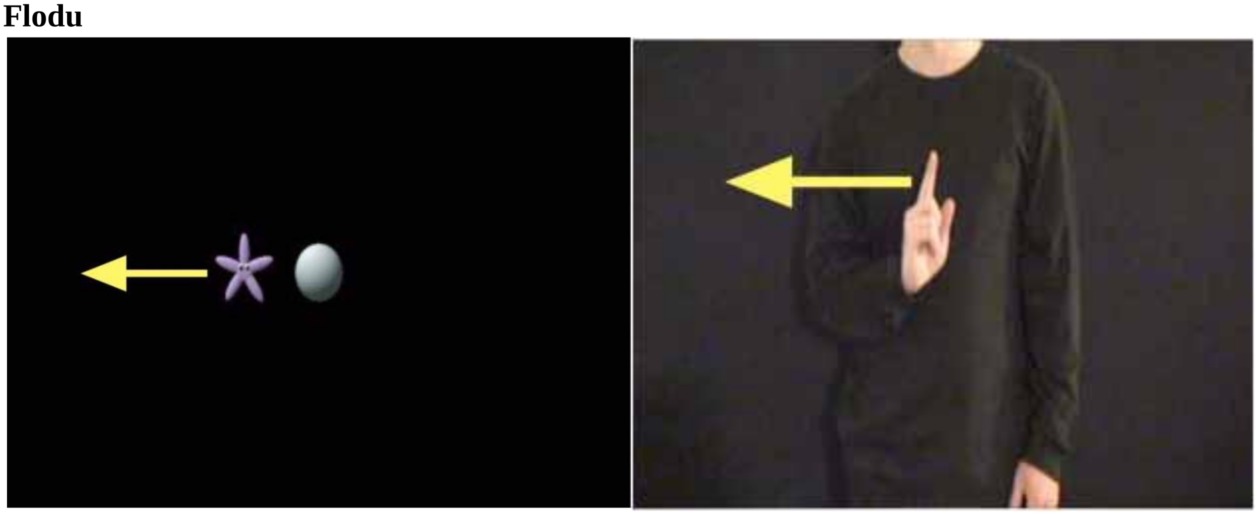


**Chulsu**



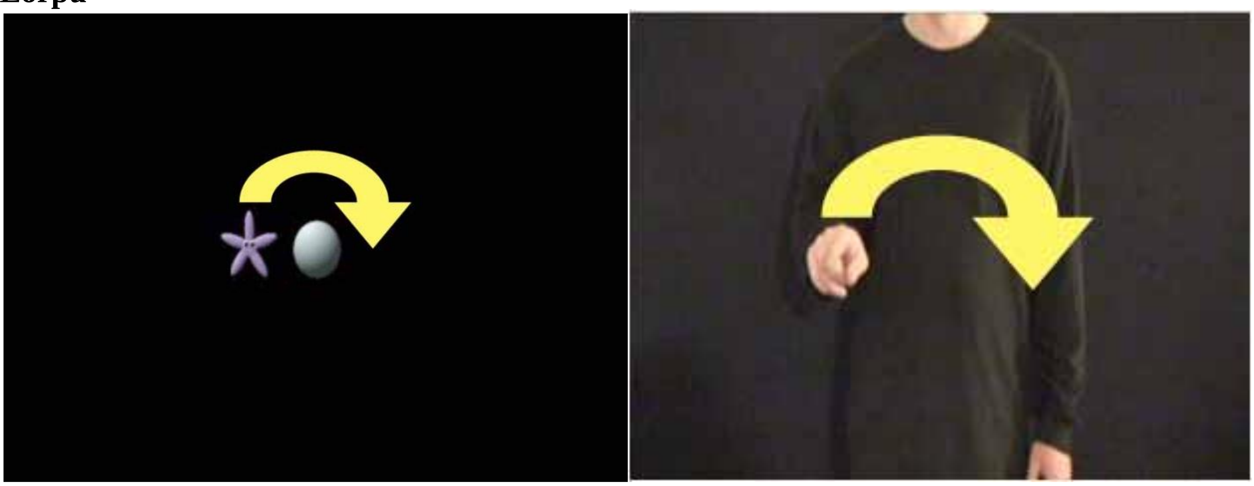
**Derlu**

**Flodu**

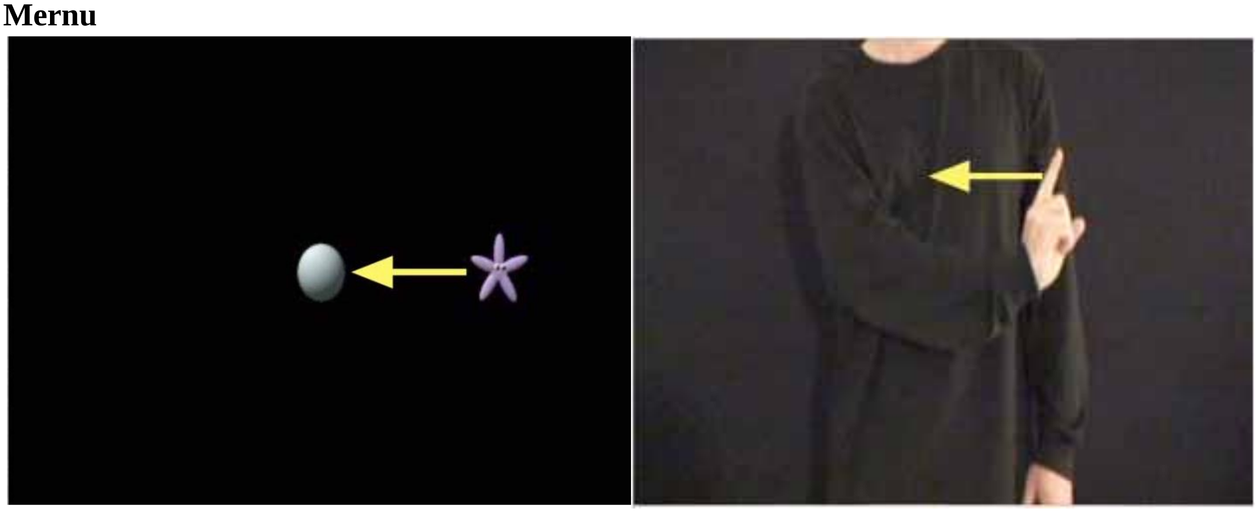


**Frengu**

**Lorpu**



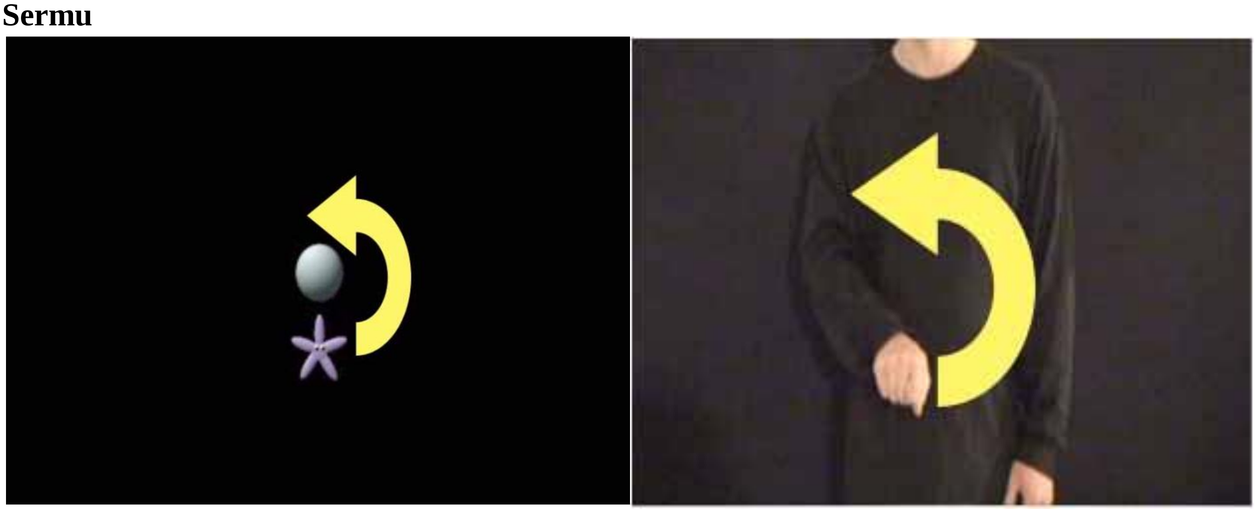
**Mernu**



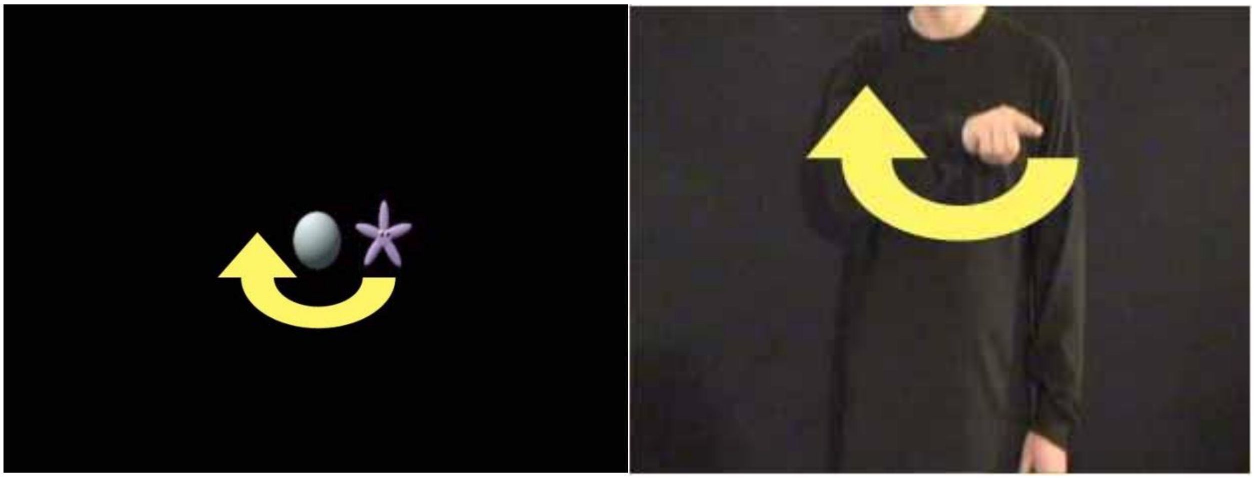
**Norcu**

**Plercu**

**Sermu**

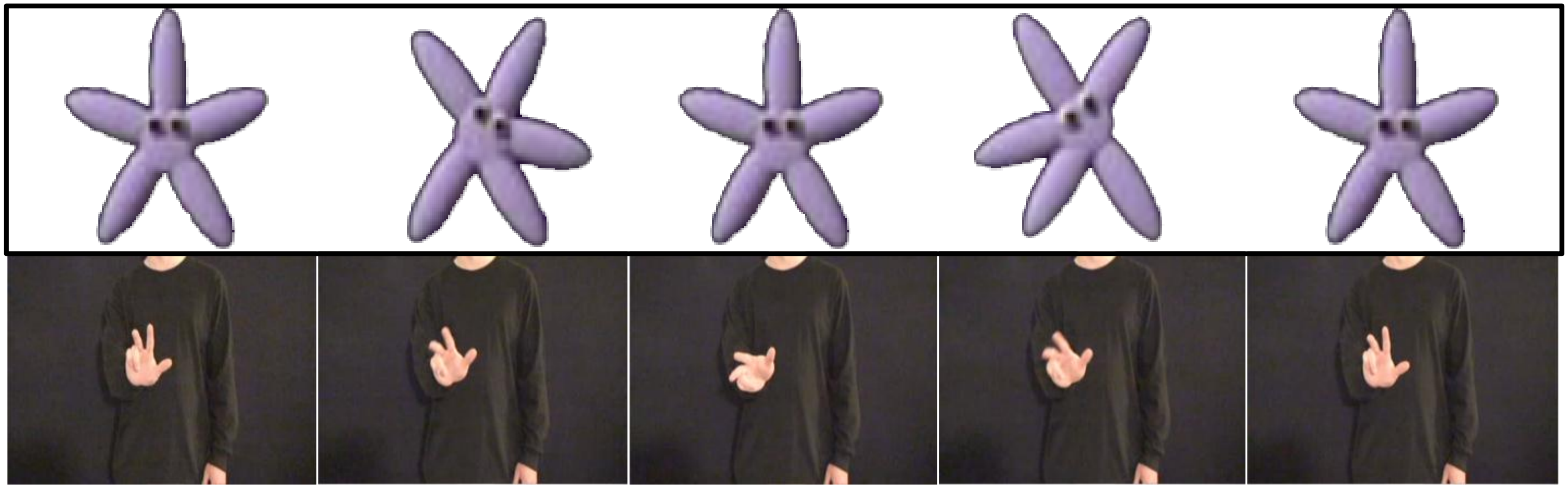


**Trospu**

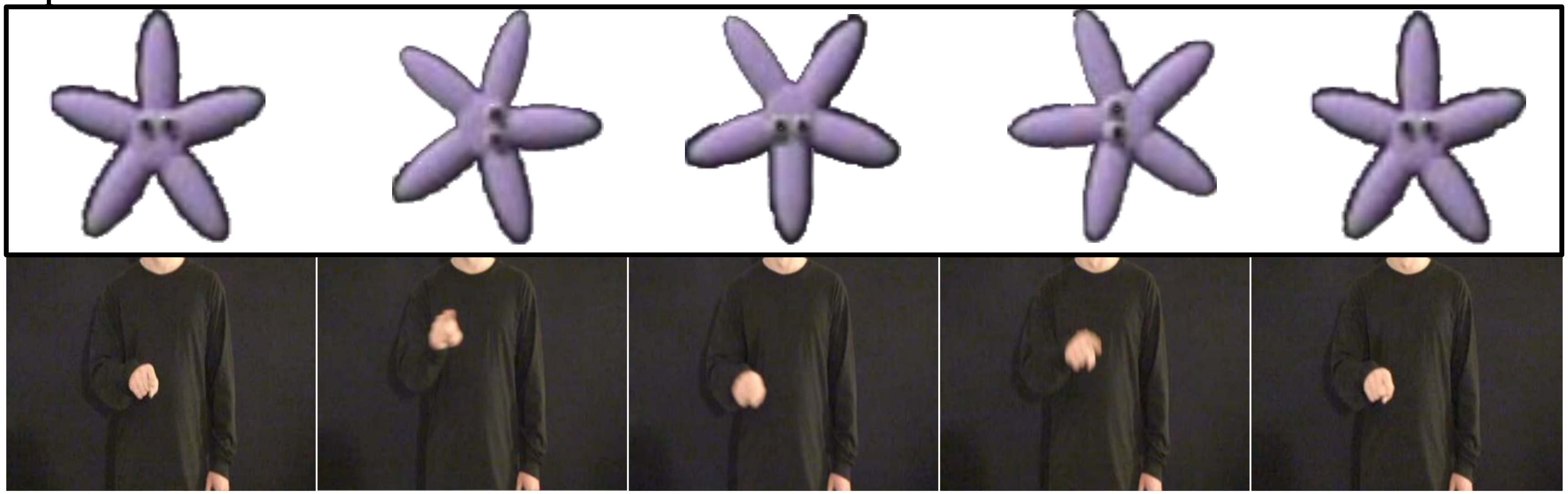


**Appendix B. Manner Animations & Gesture Instructional Videos**

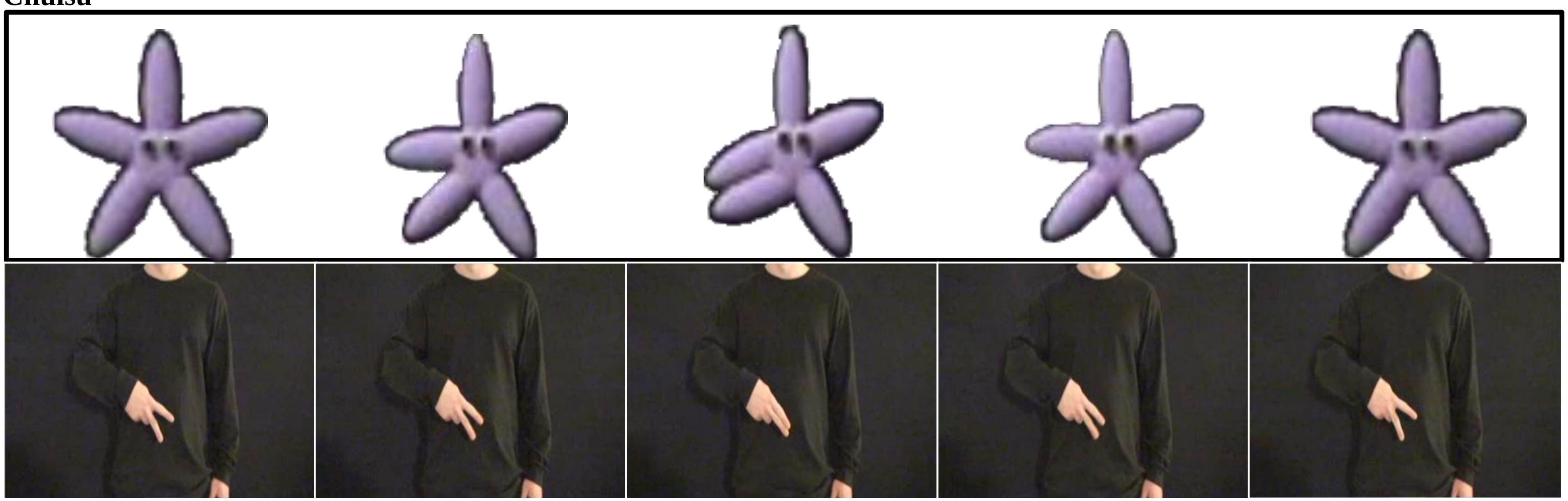
**Blomu**



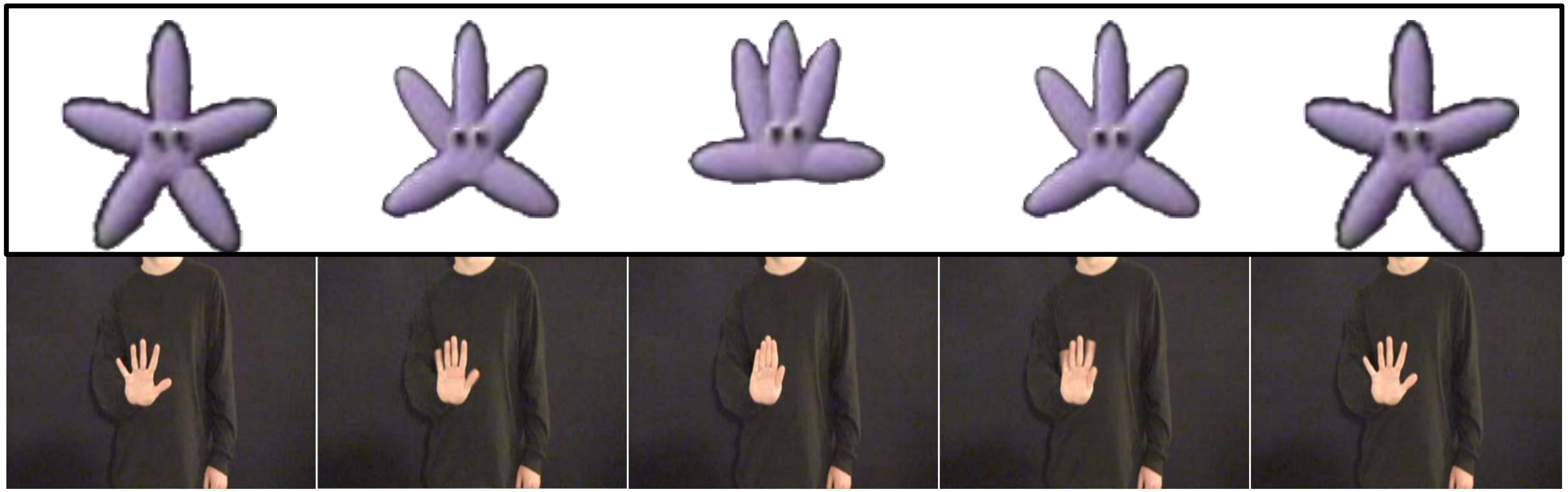
**Bripu**



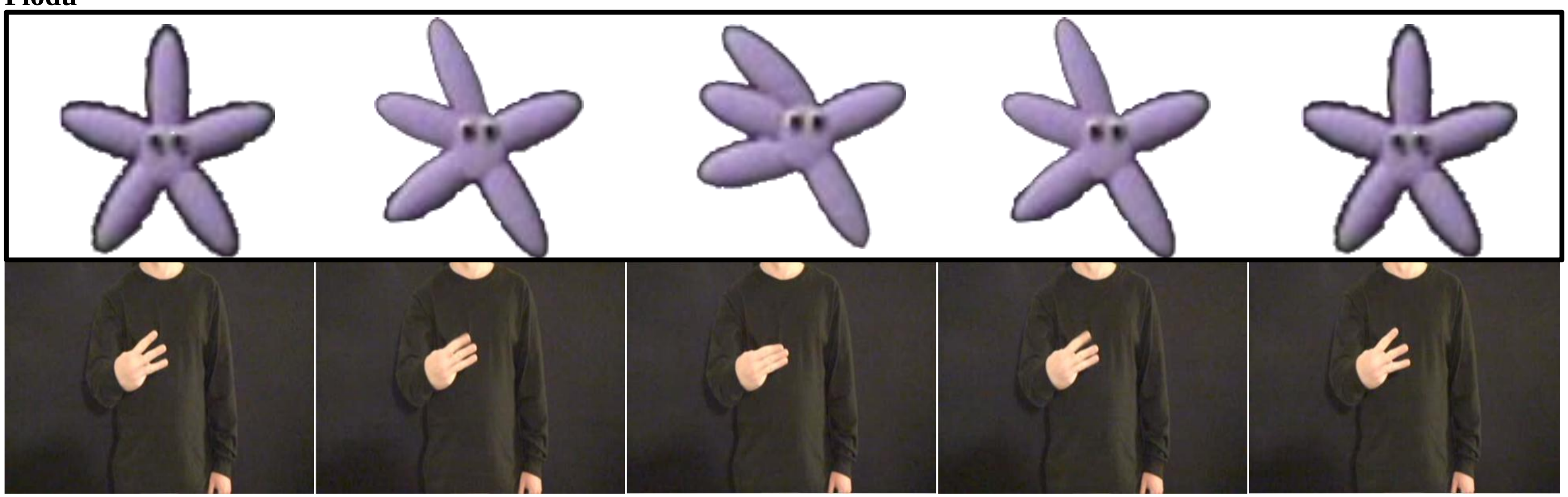
**Chulsu**



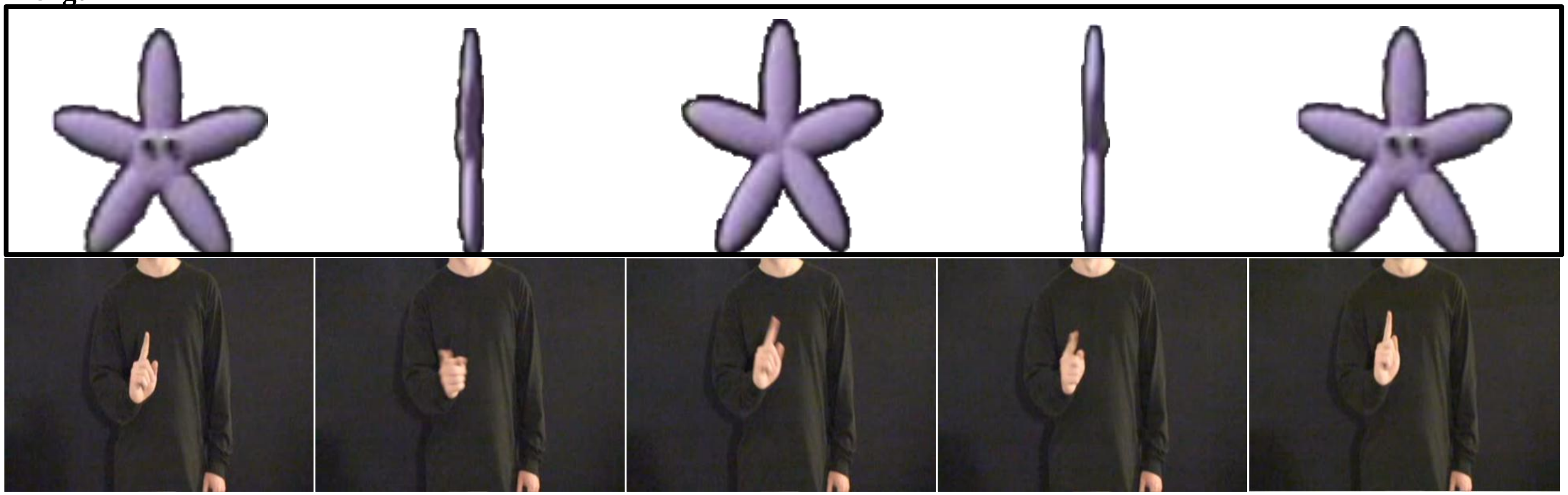
**Derlu**



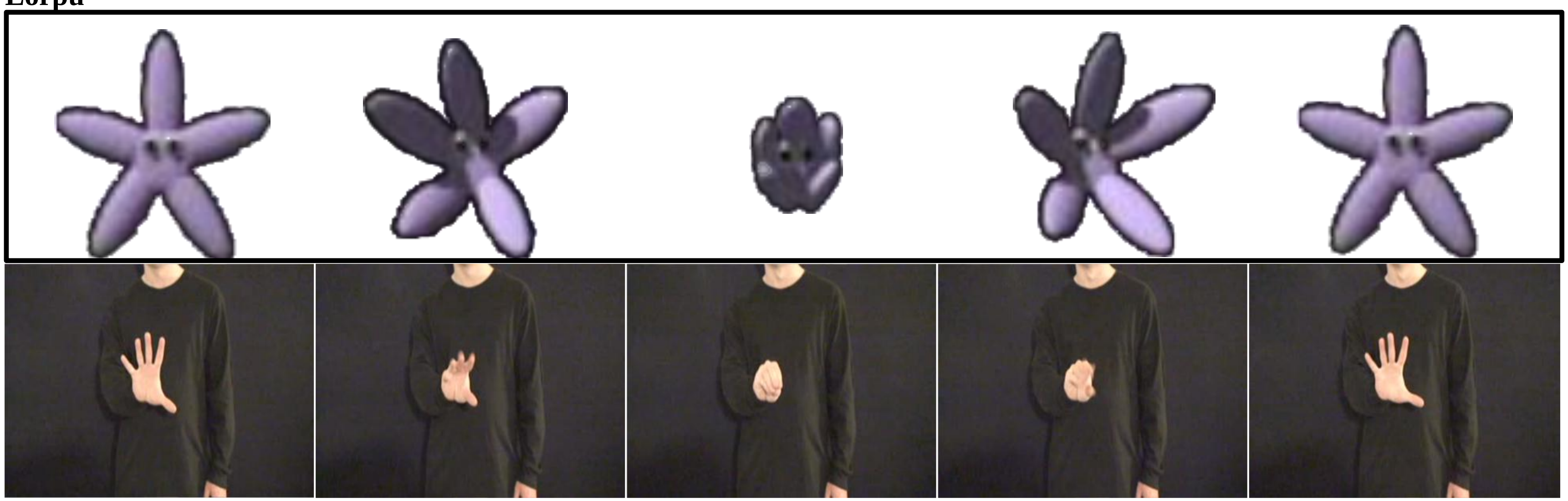
**Flodu**



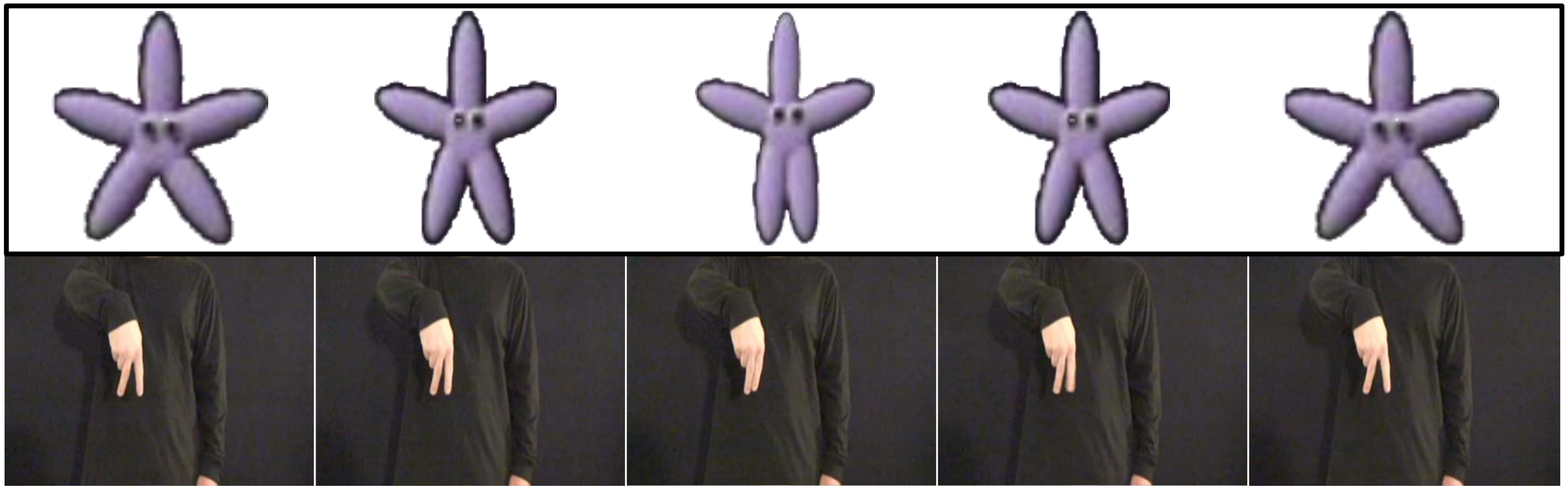
**Frengu**



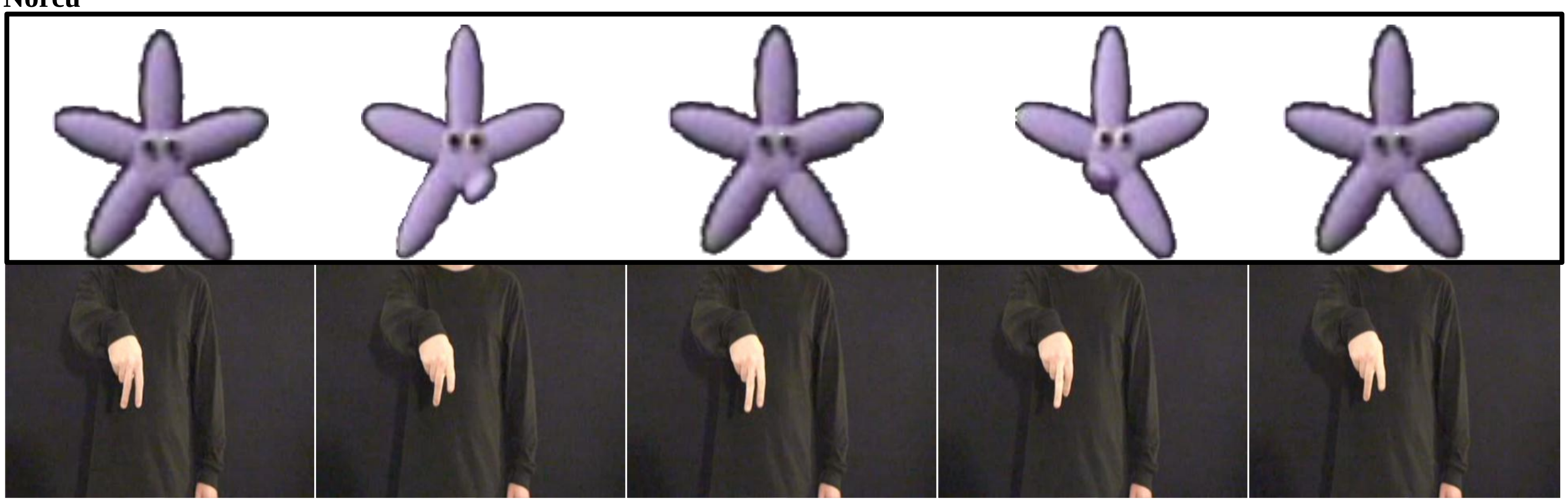
**Lorpu**



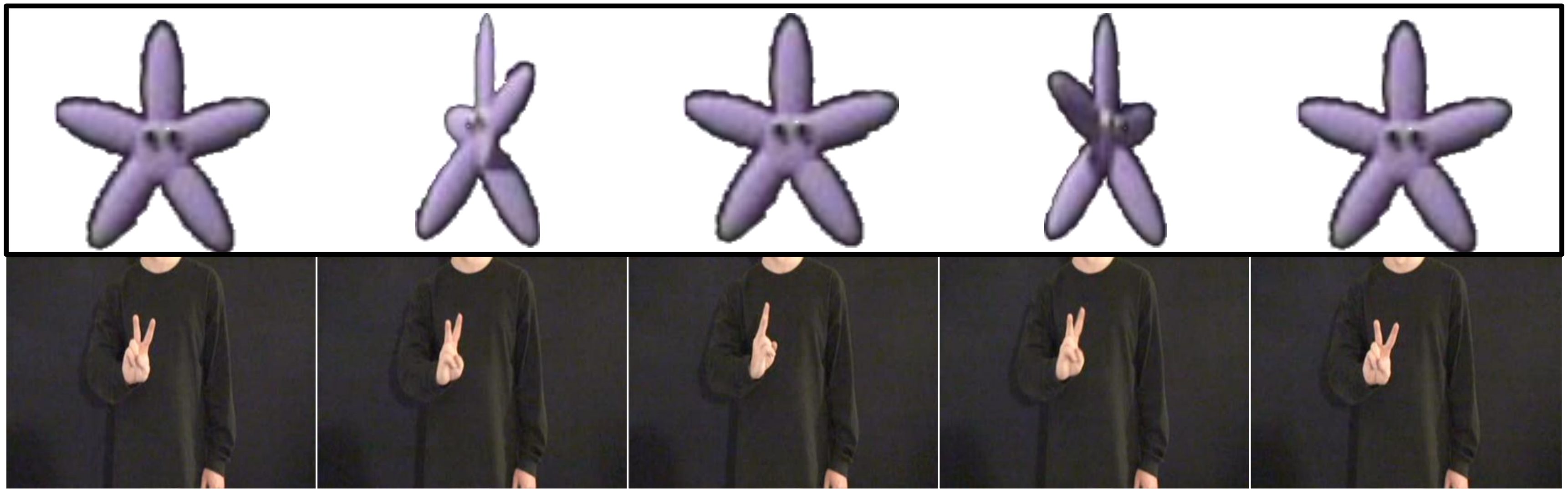
**Mernu**



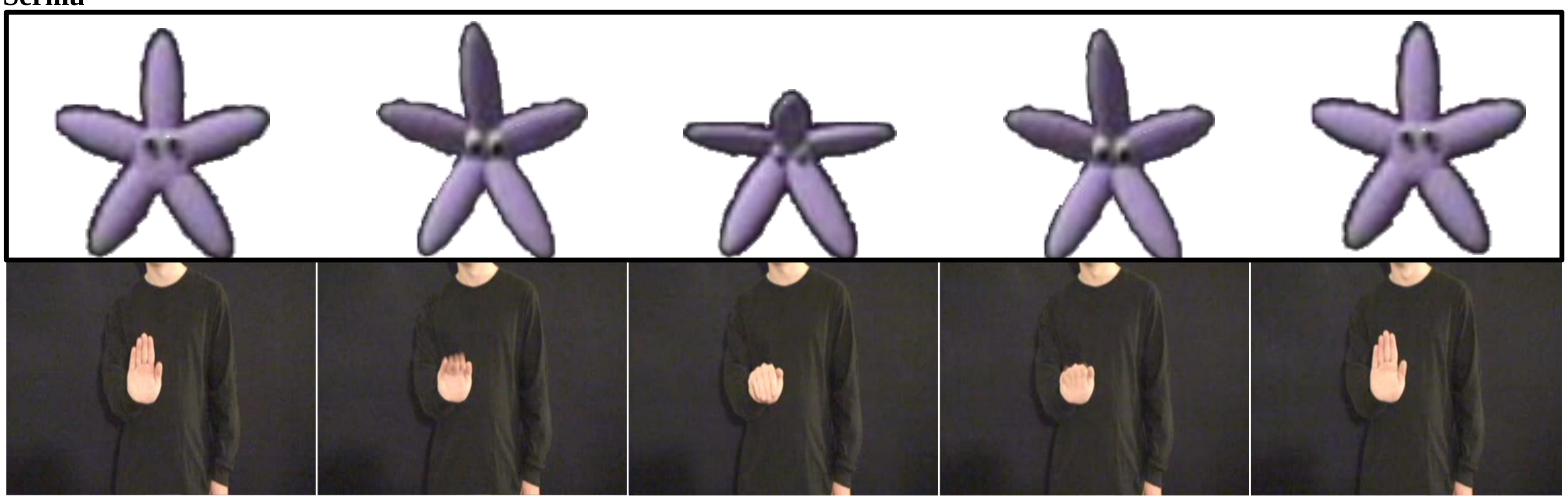
**Norcu**



**Plercu**



**Sermu**



**Trospu**