THE UNIVERSITY OF CHICAGO

HIDDEN ICONICITY IN GESTURE, SIGN, AND SPOKEN LANGUAGE

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY
DE FU YAP

CHICAGO, ILLINOIS
JUNE 2018

TABLE OF CONTENTS

ABSTRACT

How are the forms of our communicative behaviors related to their meanings? Iconic form-meaning relationships are commonly found in people's co-speech gestures: Many gestures resemble the objects or events they refer to. Iconicity has also been observed to a lesser extent in signed languages, and to an even lesser extent in spoken languages, for which form-meaning relationships are assumed to be largely arbitrary. The assumption of arbitrariness has persisted, in part, because the search for iconic form-meaning relationships has focused primarily on signs with concrete (visible, tangible) referents: literal iconicity. Yet, via metaphor, it is also possible to create iconic form-meaning relationships between signs with abstract (invisible, intangible) referents: *metaphoric iconicity*. The studies reported here demonstrate previously unrecognized iconicity in gesture, sign, and spoken language. Study 1 shows iconic form-meaning relationships in a type of gesture long thought to be devoid of meaning, beat gestures. Blind, quantitative analyses of more than 5000 spontaneous gestures showed that when people told stories implying motion or extension in space their beat gestures reflected the directions implied by the accompanying speech, no matter whether space was used literally or metaphorically. Study 2 reveals metaphoric iconicity spread throughout the lexicons of three signed languages (American Sign Language, French Sign Language, and British Sign Language), and Study 3 shows a similar pattern of metaphoric iconicity in a spoken language, Mandarin Chinese. Together, these studies suggest that the extent of iconicity in human communicative signals has been underestimated. The hidden iconicity found in beat gestures reveals the automaticity with which people activate mental representations of physical space to scaffold even highly abstract thoughts. The hidden iconicity found in words shows that spatial metaphors constrain form-meaning relationships across the lexicons of signed and spoken languages.

ACKNOWLEDGMENTS

I want to thank my advisor, Daniel Casasanto, for constantly inspiring me to be a better scientist. He sets a high standard of excellence in my research, and has taught me how good experimental psychology is done. I am especially indebted to Daniel for all his contributions of time and ideas to make my graduate experiences productive and stimulating. I am grateful to have a role model to look up to, and also a friend to laugh with during the funny moments.

I also want to thank my thesis committee, Susan Goldin-Meadow, Howard Nusbaum, and Laura Staum Casasanto, for being a continual source of encouragement and support. I am grateful to Susan, Howard and Laura for being incredibly generous with their time, and for their insightful questions.

The members of the Experience and Cognition Lab have contributed immensely to my personal and professional growth, as we moved from New York to Chicago, and back to New York. Tom Gijssels, Geoffrey Brookshire, Benjamin Pitt, Che Lucero, Yagmur Deniz Kisa, Amanda Dobbyn and Amritpal Singh have been an invaluable source of friendships as well as good advice and collaboration. I am grateful for their generosity and willingness to share their time and remarkable minds with me as we would argue about ideas, methods, codes and data. I count myself extremely fortunate to have their company as we celebrate our academic triumphs, and also when I needed a listening ear. This thesis certainly would not have been possible without their generous and profound contributions.

I want to thank all the undergraduates who have contributed to my graduate experience. Sumer, Jiang Xue and Rachel were exceptional RAs who inspired me with their dedication and ingenuity. I also want to thank all the brilliant students I have taught during my graduate career. I am grateful for their astute and thought-provoking contributions to the class discussions. They have made my teaching experiences deeply enjoyable and inspired me to become a better teacher. In particular, I want to thank the class of *Psychology of Decision-making* for taking the class discussion to another level with their energetic enthusiasm when our professor visited the class. They asked me if they behaved well during the visit – they sure did!

I am eternally grateful to have the opportunity and good fortune to have spent my graduate years as part of the community at the New School for Social Research, Columbia University, University of Chicago and Cornell University. My time at these institutions was made enjoyable by the many friends and groups that became a part of my life. I am grateful for many memorable "firsts" we shared, including my first ice-skating experience, the first horrendous beating we received in a basketball game, and my first metal concert.

Finally, I want to thank my family for their unwavering love and encouragement. I am grateful to my parents, my brother and sister who supported me in all my pursuits. I remember fondly the delectable home-cooked food when I return home, and the lovely weekly conversations. I also want to thank my fiancee's family for treating me like their own, and for the many durian feasts during my vacations. Even though I was half the world away, I am grateful for the support and encouragement from my friends, and my colleagues from my teaching days. In particular, I want to thank Alvin, Dennis, Teck Tee, Randy and Mdm Lee for their love and faithful support. Most of all, I want to thank my loving and supportive fiancee, Lynn, for always believing in me. I owe a special debt to her, whose patience and belief in me allowed me to weather any uncertainty in my research. I am especially grateful to Lynn for making me smile like crazy so effortlessly, and for making my graduate experience more beautiful and enjoyable.

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Form and meaning in human communication

How are the forms of communicative signals related to their meanings? This dissertation explores form-meaning relationships in two of the most complex forms of human communicative behavior: gesture and language.

Since Saussure (1916/1959), the idea that words' forms are arbitrarily related to their meanings has been widely accepted. According to Saussure, the meaning of "tree" is unmotivated by the letters a-r-b-r-e in French, since in principle it can be represented by other letters in other languages, such as t-r-e-e in English.

The Saussurean doctrine of the "arbitrariness of the sign" has been based primarily on analyses of the obligatory features of the sounds of spoken words. However human communication involves a multi-modal composite signal: Speech and gestures (e.g., manual and verbal gestures) form an integrated communicative unit to convey information about the same referent (Clark, 2004; Levinson & Holler, 2014; Engle, 1998). For instance, speakers may produce co-speech gestures that mimic the shape of the object they are talking about (e.g., tracing a circle to mimic the shape of a round cake while talking about a cake). Speakers can also vary the speed of their utterances to convey information about the speed of the referent's motion, and their listeners rely on the verbal gesture to infer the speed of the motion even though it was not explicitly described in speech (Shintel, Nusbaum, & Okrent, 2006). The constituent parts of composite communicative signals are likely to differ in their degree of arbitrariness (Clark, 2004; Engle, 1998; Buchler, 2014). Therefore, an accurate assessment of the arbitrariness of the sign requires consideration of the different kinds of signs, and the different components of the *composite signs*, that humans use to communicate.

1.2 Known iconicity in co-speech gestures

Iconicity, according to Peirce, is a *perceptual resemblance* between the form of a sign and the form of the thing it signifies(Buchler, 2014). Iconic form-meaning relationships are common in the gestures we produce when we speak (McNeill, 1992).

Co-speech gestures are classified into four major categories (McNeill, 1992). Three of these gesture types are recognized to carry non-arbitrary form-meaning relationships. Most notably, *iconic gestures* depict some concrete aspect of the referents of the words they accompany (e.g., tracing an arc to show the shape of a rainbow). In a special class of iconics called *metaphoric gestures*, concrete objects or relationships are depicted with the hands in order to represent some aspect of an abstract idea (e.g., tracing an arc to indicate the *rise* and *fall* of a civilization; (Cienki, 1998; Sweetser, 1998).

Emblems are conventionalized signs that function much like words do. Some emblems have form-meaning relationships that are arbitrary (or at least obscure), but others depict concrete aspects of their referents. For instance, the *horns* emblem, which is formed by extending the index and pinky fingers while holding down the rest of the fingers, represents a pair of horns (to a heavy metal music enthusiast these are the horns of the devil, whereas to a fan of the University of Texas they are the horns of a longhorn steer (Casasanto, 2013)).

Deictic gestures are typically performed with a pointing finger to serve an indexical function (McNeill, 1992). Even though deictics are not iconic, they are still constrained by a strict relationship between form and meaning (Buchler, 2014). If you change the form by pointing in a different direction, you change the meaning of the gesture.

1.3 Known iconicity in signed and spoken languages

Although the presence of iconicity in sign was once controversial (Stokoe Jr, 2005), it is now well established that the forms of some signs are motivated iconically. However, for the most part, iconicity has been demonstrated for concrete nouns that are easy to depict with the hands (e.g., to make the sign for *phone* in American Sign Language (ASL), the signer mimics the shape of a telephone receiver) (Taub, 2001; Perniss, Thompson, & Vigliocco, 2010; Goldin-Meadow & Brentari, 2017).

In addition to iconically motivated words for concrete objects and actions, however, signed languages also contain some words for abstract ideas whose forms are motivated by *metaphoric iconicity.* In the domain of emotional valence, for example, Taub (2001) provides pairs of examples in ASL. The sign for "improve" moves upward, whereas the sign for "deteriorate" moves downward, consistent with metaphorical expressions in English and other languages linking positive ideas with upward motion and negative ideas with downward motion (Lakoff, 1980). Likewise, Taub (2001) noted that the sign for *sad* moves downward (two hands, palms facing chest moves down in a "V") and the sign for *happy* moves upwards (rubbing chest upwards). These examples provide evidence that the forms of signs for at least a few abstract concepts are motivated by the metaphorical link between valence and vertical space.

Iconicity in spoken languages is traditionally believed to be found only in limited pockets, such as ideophones (e.g., "tinkle" and "bang" sound like their referents; Nuckolls, 2004), and mimetics, in which words mimic other physical properties of their referents (e.g., "pota" in Japanese refers to a single event whereas "potapota" refers to repetition of the same event; Imai, Kita, Nagumo, and Okada, 2008). Linguists have also noted relationships between the way speakers articulate words related to size and the size of their vocal tracts when they pronounce them. Words referring to small sizes such as "teeny" in English and "petite" in French cause speakers to shorten their vocal tract, whereas words referring to large sizes such as "huge" and "grande" cause speakers to lengthen their vocal tracts. These kinds of iconic relationships rely on concrete qualities of the referent being echoed in the form of the word, so only certain kinds of meaning are eligible to participate in them (Ohala, Hinton, & Nichols, 1997). This constraint is believed to severely limit the extent of iconicity in speech (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015).

1.4 Hidden iconicity in gesture, sign and spoken language

Here we propose that the extent of iconicity has been underestimated, across forms of human communication. Study 1 shows iconic form-meaning relationships in a type of gesture long thought to be meaningless, *beat gestures*: simple repetitive flicks of the hand that are known to serve pragmatic functions but have been widely agreed to be without any semantics. Blind, quantitative analyses of more than 5000 spontaneous gestures showed that when people told stories implying motion or extension in space their beat gestures reflected the directions implied by the accompanying speech, no matter whether space was used literally or metaphorically. Study 2 suggests that the metaphoric form-meaning correspondences that have been documented in signed languages to date are only the tip of the (metaphorical) iceberg, and shows that metaphoric iconicity is spread throughout the lexicons of three signed languages (American Sign Language, French Sign Language, and British Sign Language). Study 3 shows a similar pattern of lexicon-wide metaphoric iconicity in a spoken language, Mandarin Chinese.

CHAPTER 2

BEAT GESTURES ENCODE SPATIAL SEMANTICS

People gesture when they talk, and often these hand movements convey meaning (Goldin-Meadow, 1999; Özyürek, 2014; Bavelas, Beavin, Chovil, Lawrie, & Wade, 1992; Alibali, Kita, & Young, 2000; Krauss, Morrel-Samuels, & Colasante, 1991; Kita et al., 2007; McNeill, Cassell, & McCullough, 1994) . Gestures can serve as depictions of objects or events, as pointers, or as symbolic stand-ins for words or phrases (McNeill, 1992). Beat gestures, however, do not carry information in any of these ways. Beats are a common type of gesture that is simple and repetitive, involving biphasic flicks of the hand or fingers (McNeill, 1992). Because their form is not obviously related to the meaning of the speech they accompany, beat gestures are widely believed to carry no semantic information [1] (McNeill, 1992; Tuite, 1993; Beattie & Coughlan, 1999; Alibali, Heath, & Myers, 2001; Leonard & Cummins, 2011; So, Sim Chen-Hui, & Low Wei-Shan, 2012; Goldin-Meadow & Alibali, 2013; Casasanto, 2013; Wang & Chu, 2013; Gullberg, 1998; Kelly, Manning, & Rodak, 2008; Tellier, 2009; Chu, Meyer, Foulkes, & Kita, 2014; Özyürek, 2014; Biau & Soto-Faraco, 2015; Holle et al., 2012). Instead, beats are believed to serve a pragmatic function, capturing attention (Dimitrova, Chu, Wang, Özyürek, & Hagoort, 2016; Krahmer & Swerts, 2007; McNeill, 1992; Hubbard, Wilson, Callan, & Dapretto, 2009) or adding emphasis to the spoken message (So et al., 2012; Biau & Soto-Faraco, 2013; Holle et al., 2012) without serving any semantic function.

We hypothesized that meaning may be hidden in a feature of beat gestures long believed to be meaningless: their direction. When people describe spatial scenarios, they build schematic models of space in their minds (Casasanto & Bottini, 2014). These spatial schemas lack many details of the spatial environment, but preserve features such as direction, orientation, and the relative positions of objects or events. Directional schemas can be used

1. Tuite (1993) proposed a category of "spatial beats" but these gestures were posited to help with discourse structuring (i.e., establishing opposition between two discourse entities), not to encode semantic content. Like other theorists, Tuite believed that "the beat gesture...is a kinetic realization of the underlying pulse, and *almost entirely lacking in semantic content derived from the speaker's internal representation*" (pg. 99, italics added).

to organize our thoughts not only about literal spatial scenarios (an object traveling along a path), but also about abstract scenarios that are spatialized metaphorically such as the temperature *going up* or prices *going down* (Lakoff, 1980).

Every beat gesture must have some direction, whether or not the gesturer intends to produce it along a certain trajectory (or intends to produce it, at all). Many beats are downward, but they occur in other directions, as well. Therefore, when a speaker produces a beat gesture while maintaining a spatial schema in mind, the direction of the gesture is necessarily either *congruent* or *incongruent* with the direction of the schema. We reasoned that, in order to avoid cognitive conflict, when speakers produce language that implies a literal or metaphorical spatial schema, any beat gestures they produce should tend follow the direction implied by the accompanying speech.

To test whether beat gestures reflect the spatial content of speakers' thoughts we videotaped pairs of participants as they took turns memorizing and retelling brief stories implying literal or metaphorical motion in one of four directions: up, down, left, or right. We varied the visibility of speakers' gestures to distinguish their cognitive and communicative functions, and to test how automatically speakers produced schema-congruent gestures. We varied the kind of language used in the stories to determine whether speakers would produce schema-congruent gestures only when they used spatial words to describe literal motion (my rocket *went higher*), or also when they used these spatial words metaphorically to describe abstract motion (my grades *went higher*), and even when speakers expressed the same metaphorically spatialized ideas without any spatial language (my grades *got better*).

Multiple factors constrain the directions of gestures, so we expected that the predicted speech-gesture relationships would be statistical, not absolute. We therefore collected a large corpus of stories containing more than 5000 spontaneous co-speech gestures. Speakers do not produce an equal number of gestures in each direction, so it is not possible to specify the chance level of gesturing in any one direction, *a priori*. We therefore developed permutation and bootstrapping tests to compare the observed rate of gestures in each direction to an empirically derived chance distribution [2].

2.1 Experiment 1a: Hidden iconicity in beat gestures

2.1.1 Methods

Participants

Twenty-eight Stanford University undergraduates (14 male) were recruited in pairs, and participated for course credit. The study was approved by Stanford University's institutional review board.

Materials

Brief stories (50-100 words) were constructed. Each story contained one of 3 types of language: Literal Spatial Language (LSL), Metaphorical Spatial Language (MSL), or Non-Spatial Language (NSL), and each story implied motion or extension in one of 4 directions (upward, downward, right, or left; see Appendix A).

LSL stories described physical objects and events oriented along either a horizontal or vertical axis, and directed either upward, downward, right, or left. MSL stories described non-spatial phenomena that are nevertheless commonly expressed using spatial metaphors implying both orientation and direction (e.g., the temperature *went higher*; the price *went lower*). NSL stories were identical to the MSL stories, except that all metaphorical spatial language was replaced with non-spatial language conveying nearly the same meaning (e.g., the temperature *got hotter*; the price *got cheaper*).

2. A preliminary report on these experiments appeared as Yap, D., Brookshire, G., & Casasanto, D. (2018). *Beat gestures encode spatial semantics.* Proceedings of the 40th Annual Conference of the Cognitive Science Society. Madison, WI

Procedure

Participants studied and retold stories one at a time to their partners. They sat on stools facing one another across a small table. Stories appeared on an iMac computer monitor positioned in front of each participant. They were given 60 seconds to study each story. When the story disappeared, they were instructed to try to retell it to their partner, as accurately as possible, and told their partner would be quizzed on the content of the stories that they heard at the end of the experiment. The post-experiment quiz was only intended to motivate participants to be attentive while telling stories and listening to them; it was not intended to diagnose performance. Stories were written in the second person (e.g. "You're testing some new model rockets"), but participants were asked to retell the stories in the first person (e.g. "I'm testing some new model rockets") as if relating their own experiences, to encourage them to adopt their own perspective when gesturing. After telling a warm-up story, each participant told 6 target stories: 2 LSL, 2 MSL, and 2 NSL. The order of stories was randomized, and content of stories was counterbalanced so that each pair of participants received only one version of each metaphorical story: either MSL or NSL (i.e., one pair of participants would receive the MSL story about "grades going higher", another pair of participants would receive the minimally paired NSL story about "grades getting better").

Participants were told that the experiment was "about storytelling". Although they were aware that they were being videotaped, they were not instructed to gesture, and did not know that their gestures were of interest. Each testing session lasted 20 to 30 minutes.

2.2 Experiment 1b: Gesturing for the blindfolded?

What function does the beat gesture serve? Did speakers produce beat gestures for an internal-cognitive function, or were they designed for communication? We adopted a standard method for probing the communicative function of gesture by making gestures invisible.

2.2.1 Methods

Participants

Twenty-eight Stanford University undergraduates (16 male) were recruited in pairs, and participated in exchange for course credit. Four participants' data were lost due to technical error, leaving 24 participants. The study was approved by Stanford University's institutional review board.

Materials and procedure

Materials and procedures were identical to those used in Experiment 1a, except that participants put on a blindfold before listening to the stories. They took off their blindfolds before they retold the next stories to their blindfolded partners.

2.3 Experiment 1c: Are beats recipient-designed?

A limitation of the standard method used in Experiment 1b, however, is that there are multiple interpretations of the persistence of gesture. Even if gesture's function were purely communicative, speakers might continue to gesture to blindfolded listeners out of a habit or *overlearned social behaviors* (e.g., gesturing while talking on the phone) (de Ruirer, 2000; Nass & Moon, 2000).

To resolve this ambiguity, we asked participants to tell stories while separated by a barrier that allowed partners to see only the tops of each other's faces. The barrier occluded the speakers' usual gesture space (McNeill, 1992), allowing us to test whether gestures were designed to be seen by the listener (Özyürek, 2002). If gestures were intended to be communicative, speakers had the option to raise them above the barrier, making them visible. Gesturing in the visible space was not physically challenging, but it required adapting one's usual gestural routines to accommodate the listener: a common process known as *recipient design* (Sacks, Schlegoff, & Jefferson, 1974).

2.3.1 Methods

Participants

Twenty-eight Stanford University undergraduates (11 male) were recruited in pairs, and participated in exchange for course credit. The study was approved by Stanford University's institutional review board.

Materials and procedure

Materials and procedures were identical to those used in Experiment 1a, with the following exception: Participants were partially separated by a barrier formed by a pair of dry-erase whiteboards, back-to back, which measured 3 ft. wide by 2 ft. high from the table's surface. Seated on stools, participants could see just the top of their partners' faces. The barrier occluded all but the extreme upper limit of their usual gesture space (McNeill, 1992). If speakers gestured in their usual space, nearly all of their gestures would be invisible to the listener, but it was entirely possible for speakers to make their gestures visible above the barrier. Listeners were instructed to jot down 3 keywords with a dry-erase marker after each story they heard, to remind them of the story's content.

Coding and analysis

Gestures and speech were coded separately by 3 independent coders, following a "blind-then-deaf" procedure (Casasanto & Jasmin, 2012). Speech was analysed by Coder 1, blind to the accompanying gestures. Gestures were analyzed by two independent coders (Coders 2 and 3) in three phases, the first two of which were done "deaf" to the accompanying speech.

Speech coding

Coder 1 transcribed the audio portion of the participants' stories, blind to the accompanying video. The transcripts were parsed into clauses. Each clause was classified as a directional clause if it contained language that implied literal or metaphorical motion, extent, or position along either the lateral or vertical axis. Phrases were classified as non-directional clauses if they did not imply any direction, literally or metaphorically, or if they implied a mixture of directions. Directional clauses were classified as LSL, MSL, or NSL.

Gesture direction coding

Coder 2 coded gestures' directions "deaf" to the accompanying speech. All gestures were parsed into gesture phrases (McNeill, 1992). Gesture strokes were then coded for to the hand(s) used to produce them (Left, Right, Bimanual), the axis along which the hand moved (Lateral, Sagittal, Vertical), and the direction (Upward, Downward, Leftward, Rightward, Away from the body, Toward the body, Back over the shoulder, Inward [both hands moving toward each other], or Outward [hands moving away from each other]; McNeill, 1992). For strokes with more than one directional component (e.g., down and left), Coder 2 recorded the dominant direction.

Gesture type coding

Coder 2 classified gesture types, in two stages. In the first stage, using only the video with no audio, each gesture was determined to be either a beat or a non-beat according to a strict application of McNeill's (1992) beat filter. Gestures were classified as beats if they had (i.) two movement phases, (ii.) a relaxed handshape, and (iii.) movement only within a single region of gesture-space. Importantly, gestures that included any features of other gesture types (e.g. deictics or iconics) were not classified as beats: *Only "pure" beats were included in our beat analyses.*

In the second stage, using both audio and video, the non-beat gestures were classified into the categories proposed by McNeill (1992): iconics, metaphorics, deictics, and emblems, based on the gestures' forms and on the co-occurring speech. Self-adaptive actions were excluded from all analyses.

Inter-coder reliability

Coder 3 recoded gestures from a random selection of 10% of spoken clauses (493 of 4932 total). Gestures were coded for (i.) direction and (ii.) beat/non-beat status using only the video with no audio. Inter-coder reliability was 91% for the coding of direction, and 87% for the coding of beats.

Coding of schema-congruity

If the direction of a gesture stroke matched the *a priori* prediction based on the overall direction implied by the story it was coded as schema-congruent (e.g., upward gesture during a story about a rocket going higher); otherwise it was coded as schema-incongruent. For analyses by language type (Fig 2.2A) congruity was determined with respect to the direction implied by each spoken clause, rather than the overall direction implied by the story.

Deriving permuted chance distributions

To test schema congruity across all four story directions we used permutation tests to determine whether congruity was greater than would be expected by chance. In each permutation ($k = 10^6$), we randomly shuffled the story direction within subjects. In the observed data, each gesture occurred during a story with a given direction. In the permuted data, the story directions were randomly shuffled, and congruity was computed over these randomized data. The permuted distributions are reported as Mean $\pm$ SD. P-values were calculated as the proportion of shuffled congruity values higher than the observed value. In all cases, the observed value was higher than all permuted values, and therefore $P < 1/k$.

Deriving bootstrapped chance distribution in each story direction

When computing congruity within a story direction we performed bootstrapping tests ($k = 10^6$) to test whether congruity for each story direction was greater than would be expected by chance. In each randomization, we sampled gestures with replacement within each subject but across all story directions, ensuring that each subject contributed the same number of gestures to the comparison distributions as to the empirical distribution. P-values were calculated as for permutations tests.

Between-condition statistical tests

To compare across experimental conditions (e.g. Directional vs. Non-directional clauses), we used mixed-effects binomial logistic regressions with the *lme4* package in R. Congruity was modeled with a fixed effect for the factor of interest, and maximal random effects for Subjects (random intercepts for between-subjects comparisons, and random intercepts and slopes for within-subjects comparisons). We tested for statistical significance by dropping the term of interest from the model, and comparing the two models using a likelihood ratio test.

2.3.2 Results

Participants produced a total of 5199 gestures (Fig. 2.1A). We identified beat gestures using a strict, conservative application of McNeill's *beat filter* (McNeill, 1992), a standard method for distinguishing beats from other gesture types. This analysis yielded 3946 beats (75.9% of all gestures). Of the beat gestures, 3113 (78.9%) were produced in the directions of interest: upward, downward, leftward, and rightward (Fig. 2.1B-C). We analyzed the probability that these gestures were produced in the same direction as the motion described in the stories.

Do beat gestures encode spatial semantics?

Do beat gestures reflect spatial schemas in speakers' minds? To find out, we tested whether people produce beats in the same direction as the dominant spatial schema implied by each story. Because participants produced unequal numbers of gestures in each direction, we

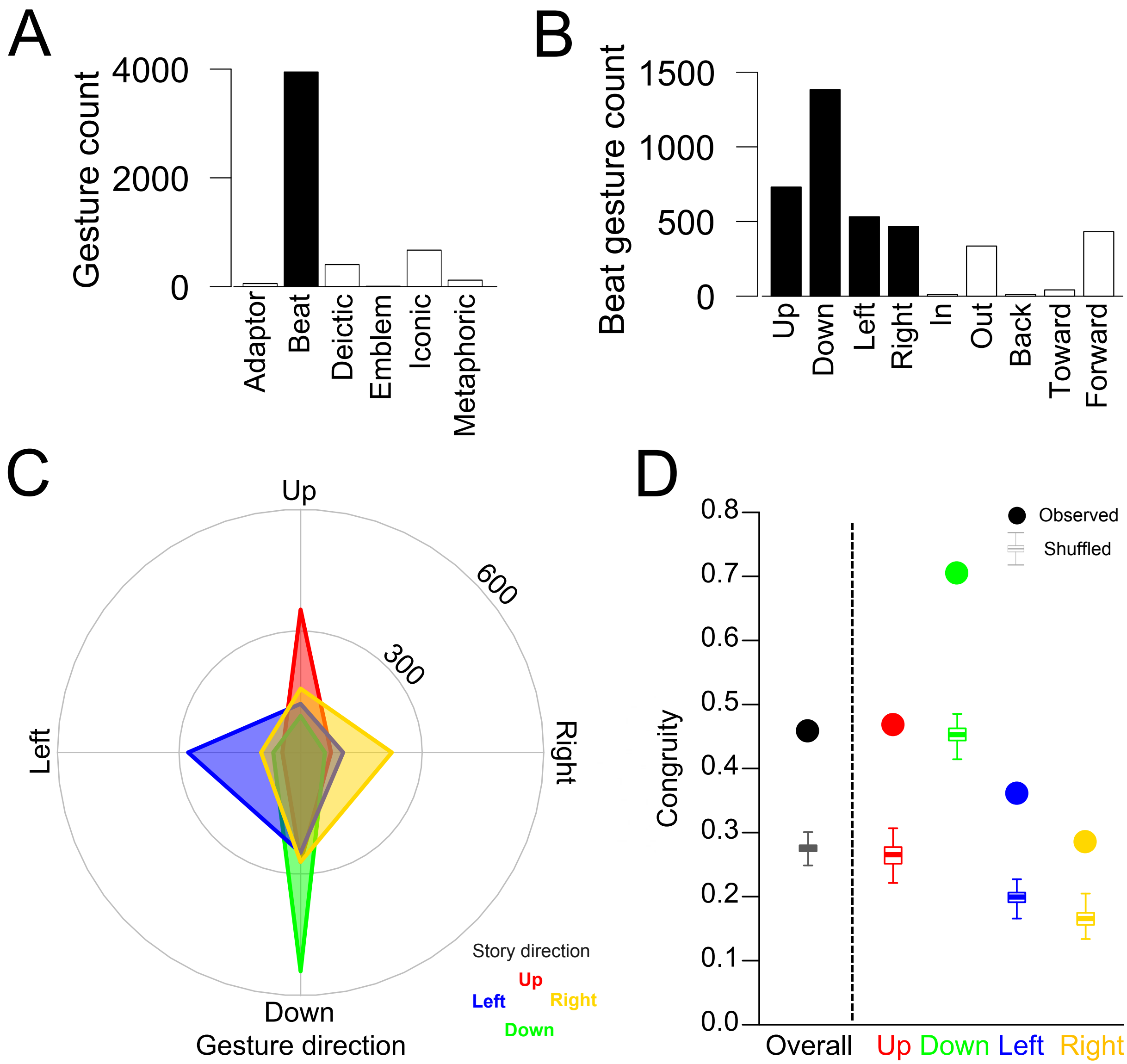

Figure 2.1: Beats were the most frequent type of gesture, and they were produced in the directions implied by the stories at rates far above chance (all Ps $< 10^{-6}$). *(A)* Types of gestures. Gestures ($N = 5199$) were classified as beats (black) or other standard gesture types (white) (McNeill, 1992). *(B)* Directions of beats ($N = 3946$). The majority of beats were produced in one of the four directions implied by the stories that participants told (black). Beats produced in the other directions (white) were not analyzed. *(C)* Number of beats produced in the four directions of interest ($N = 3113$) during stories implying motion up (red), down (green), left (blue), or right (yellow). Distance from the origin indicates the number of gestures observed in each direction. *(D)* Tests of congruity between the directions of beats and the directions implied by the accompanying speech, overall and in each story direction. The dots represent the observed rate of congruity. The boxplots represent the bootstrapped chance congruity distributions (Ks $= 10^6$). Boxes show the inner quartiles, and whiskers show the range of the randomized data. When the dot does not overlap with the boxplot, $P < 10^{-6}$.

cannot compare our data against any *a priori* chance value (Parrill & Stec, 2018). Instead, we used within-subjects permutation tests to determine whether gestures were congruent with the story direction more often than would be expected by chance.

Overall, participants gestured in the direction implied by the story 46.9% of the time, significantly more often than would be expected by chance (chance rate = $27.6 \pm 0.7\%, P < 10^{-6}$; Fig. 2.1D). Separate bootstrap tests revealed that participants gestured congruently with the dominant trajectory of the stories above chance for each of the four story directions (all Ps $< 10^{-6}$; Fig. 2.1D).

Do beats reflect both literal and metaphorical spatial schemas?

To the extent that people use spatial schemas to think about both concrete spatial scenarios and abstract scenarios that are spatialized metaphorically, this spatial information should be reflected in their beat gestures. We tested this prediction by examining gesture congruity rates across three types of language: Literal Spatial Language (LSL; "my rocket went higher"; $n = 493$ gestures), Metaphorical Spatial Language (MSL; "my grades went higher"; $n = 538$), and Non-Spatial Language that paraphrased the same metaphorically spatialized ideas (NSL; "my grades got better"; $n = 1063$).

Beat gestures were produced in the directions implied by the accompanying spoken clauses more often than would be expected by chance for all three types of language (LSL, MSL, and NSL; all Ps $< 10^{-6}$; Fig. 2.2A). The strength of the schema-congruity effect did not differ significantly across the three language types ($\chi^2(2) = .52, P = .77$). The directions of speakers' beats reflected the spatial schemas implied by their accompanying speech no matter whether the spatial content was literal or metaphorical, and no matter whether speakers were using spatial language on non-spatial paraphrases.

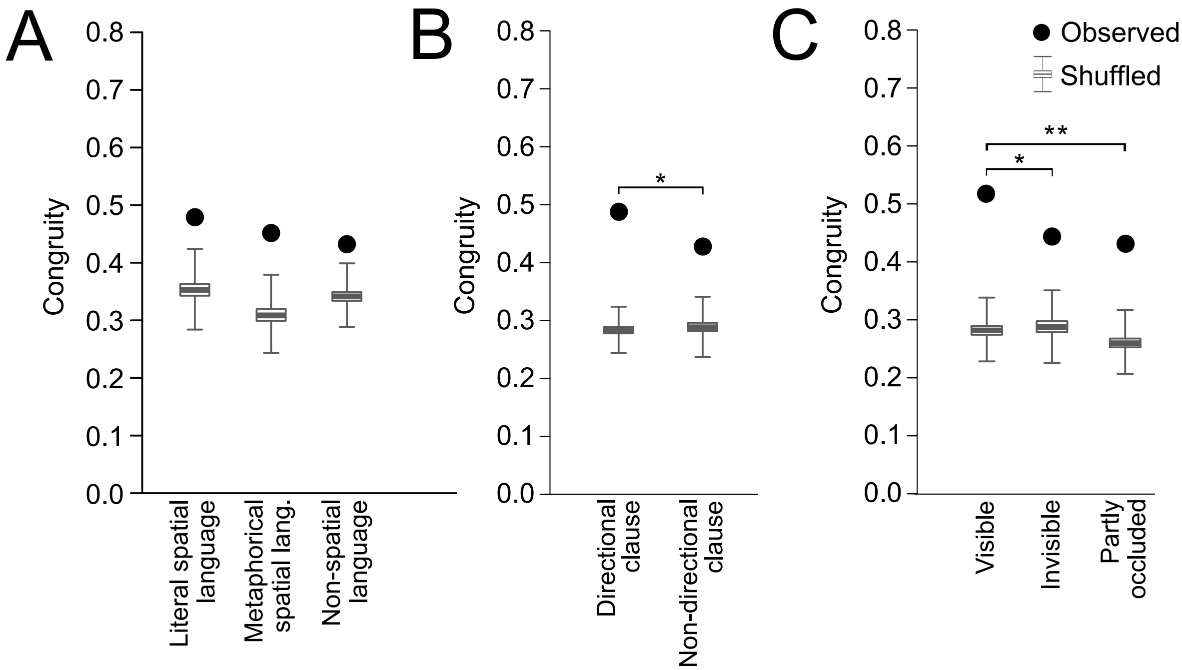

Figure 2.2: Beat gestures were congruent with the spatial schemas implied by the stories more often than would be expected by chance across all tasks and conditions (all Ps $< 10^{-6}$). *(A)* Rates of schema-congruity across language conditions. *(B)* Rates of schema-congruity for Directional Clauses (which implied motion or extension either up, down, left, or right) and during Non-Directional Clauses (with no clear spatial direction). *(C)* Rates of schema-congruity across visibility conditions. The dots represent the observed rate of congruity. The boxplots represent the permuted chance congruity distributions (Ks $= 10^6$). Boxes show the inner quartiles, and whiskers show the range of the randomized data. When the dot does not overlap with the boxplot, $P < 10^{-6}$ (* $P < .05$; ** $P < .01$).

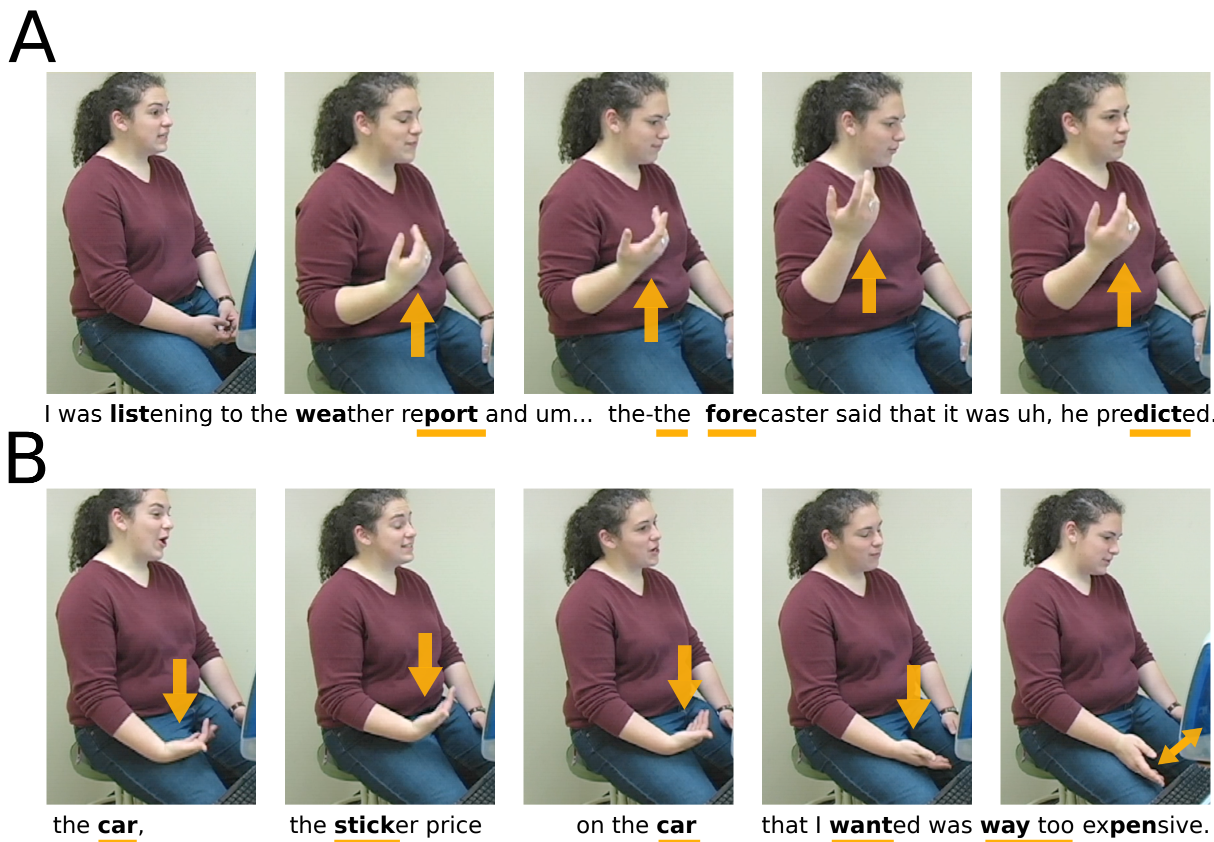

Figure 2.3: Beats tended to be co-timed with stressed syllables, and were produced in the directions implied by the stories even during non-directional clauses (Table 1). *(A)* Frames from a story implying upward motion, about rising temperatures. and *(B)* Frames from a story implying downward motion, about reducing the price of a car. Stressed syllables are in bold. Gestures were produced during the words or syllables underlined in orange.

Table 2.1: Excerpt from *checking the weather* story.

Clause	Type	Transcript
1	Non-directional	Ok. So, it was **sum**mer
2	Non-directional	and I was **list**ening to the **wea**ther re**port**
3	Non-directional	and, um, the, the **fore**caster said
4	Non-directional	that it was, um,
5	Non-directional	he pre**dict**ed
6	Directional; up	it would be 99 degrees and the temperature was ris**ing**[.]

Clauses 1-5 are non-directional clauses, with no spatial content. Clause 6 is a directional clause, with (metaphorical) upward motion implied.

Do beats reflect story-level spatial content?

Although each of the stories that participants told implied a clear spatial trajectory, not every clause implied spatial motion or extension. For example, Table 1 shows the first six clauses a participant produced while telling a story about checking the weather forecast (see also Fig. 2.3A). This narrative includes some *directional clauses* containing words or phrases that imply upward motion (e.g., "the temperature was *rising*"; clause 6), and some *non-directional clauses* that do not imply any clear direction, literal or metaphorical (e.g., "I was listening to the weather report"; clause 2). Typically, meaningful gestures have been observed to co-occur with a word whose meaning they reflect, known as the gesture's *lexical affiliate* (Morrel-Samuels & Krauss, 1992). If beats are affiliated with words that imply a direction, then schema-congruent beats should be found primarily during directional clauses.

However, about one-third of the beats in our corpus were produced during non-directional clauses. Unexpectedly, these beats appeared to follow the spatial schema that structured the story, raising the possibility that gestures can be meaningfully affiliated with speech not only at the lexical level but also at the level of a whole story and its dominant spatial schema.

To determine whether beats were affiliated with speech at the lexical level, only, or also at the story level we compared the rate of schema congruity during directional clauses (n = 2094) and non-directional clauses (n = 1019). The rate of schema congruity was slightly higher for directional clauses than for non-directional clauses ($\chi^2(1) = 5.18, P = .02$; Fig. 2.2B), but it was far greater than chance for both clause types (Ps $< 10^{-6}$).

Could schema-congruent beats in non-directional clauses be "piggybacking" on beats produced during directional clauses (i.e., as anticipations or repetitions of lexically-affiliated beats)? This skeptical possibility would be supported if non-directional-clause beats only occurred immediately preceding or following directional-clause beats. However, this was not the case, as illustrated by the story excerpt in Table 2.1 (Fig. 2.3A). The speaker produced four schema-congruent (upward) beats during clauses 2, 3, and 5, prior to the first directional clause (clause 6). Rather than piggybacking on directional-clause beats, schema-congruent beats during non-directional clauses appear to be affiliated with the spatial schema that scaffolds speakers' mental representations of the whole story.

Further evidence that at least some beats are affiliated with speech at the story level comes from gestures during clauses containing *story-incongruent* language. Figure 2.3B presents an excerpt from a story about buying a car and seeking the cheapest price possible. In order to establish the story's overall downward trajectory (*cheaper* is metaphorically *lower*) the narrative begins with the information that the car's initial price was "way too *expensive*" (*more expensive* is metaphorically *higher*). If beats during this clause were locally affiliated with the direction-implying word "expensive" they should be upward; yet, the speaker produced four *downward* beats during this clause, congruent with the overall downward spatial schema that structures this story about a car's price getting gradually lowered.

These results suggest that at least some beats are affiliated with speech not at the level of the word, phrase, or clause, but rather at the story level; their directions reflect the spatial schema that scaffolds a speaker's mental representation of the whole story.

Are beat gestures for communicating or for thinking?

To what extent were beat gestures produced for communicative purposes, or for speaker-internal cognitive purposes? To address this question we manipulated the visibility of speakers' gestures to their partners. For each pair of participants, gestures were either *visible, invisible*, or *partly occluded*.

A standard method for probing the communicative function of gesture is to make gestures invisible (Alibali et al., 2001; Bavelas, Kenwood, Johnson, & Phillips, 2002); we did this by blindfolding the listeners. A limitation of this method, however, is that there are multiple interpretations of the persistence of gesture. Even if gesture's function were purely communicative, speakers might continue to gesture to blindfolded listeners out of a habit developed in ordinary communicative settings.

To resolve this ambiguity, we asked one group of participants to tell stories while separated by a barrier that allowed partners to see only the tops of each other's faces. The barrier occluded the speakers' usual gesture space (McNeill, 1992), allowing us to test whether gestures were designed to be seen by the listener (Özyürek, 2002). If gestures were intended to be communicative, speakers had the option to raise them above the barrier, making them visible. Gesturing in the visible space was not physically challenging, but it required adapting one's usual gestural routines to accommodate the listener: a common process known as *recipient design* (Sacks et al., 1974).

Participants produced beat gestures in all three visibility conditions, (Visible: $n = 1214$; Invisible: $n = 786$; Partly Occluded: $n = 1113$). The rate of schema-congruent gestures was far greater than chance in all conditions (all Ps $< 10^{-6}$), but also differed as a function of visibility (Overall: $\chi^2(2) = 10.28, P = .006$; Visible vs. Invisible: $\chi^2(1) = 4.72, P = .03$; Visible vs. Partly Occluded: $\chi^2(1) = 9.39, P = .002$; Fig. 2.2C). Crucially, when visibility was partly occluded speakers only raised their gestures above the barrier 1.6% of the time; 98.4% of their gestures were performed in the occluded space, where they were obviously invisible to their partners, suggesting that schema-congruent beats were produced, at least in part, to serve speaker-internal cognitive functions.

2.3.3 Discussion

Among the commonly recognized basic gesture types, only beat gestures are assumed to be "meaningless" (So et al., 2012), without "any semantic content" (Kelly et al., 2008), and with "no semantic connexion to the speech they accompany" (Tellier, 2009). Here we show that these assumptions are false. Beat gestures are meaningful insomuch as their directions carry information about the spatial content of the gesturer's words and thoughts.

Overall, beats were congruent with the direction implied by the accompanying speech at a rate far above chance. This robust congruity effect was found independently for stories implying motion or extension leftward, rightward, upward, and downward. Furthermore, the direction of beats was consistent with the spatial schemas scaffolding the stories regardless of whether space was used literally (e.g., "the rocket went up") or metaphorically (e.g., "my grades went up"), and even when metaphorically spatialized ideas were expressed without using any spatial language (e.g., "my grades got better"). Schema-congruent beats were produced at a rate far above chance whether or not they could be seen, suggesting that their production was highly automatic. When the gesture space was partly occluded speakers did not adapt their gestures to be visible to the listeners, suggesting that beats were produced, at least in part, for speaker-internal cognitive purposes.

Beat gestures are for thinking, not just communicating

Why do speakers produce beat gestures? According to standard analyses, the function of beats arises from their *timing*. Beats are synchronized with stressed words or syllables in order to mark emphasis or draw attention to the content of speech (McNeill, 2015). As expected, the beats in our corpus tended to follow stress patterns in speech. Seven of the eight schema-congruent beats in Fig. 2.3A-B, for example, occurred on or immediately following a stressed syllable.

Yet, the present data suggest that beats may also serve a function (or functions) by virtue of their *direction*. Speakers may beat in story-congruent directions in order to avoid the cognitive conflict that could arise from speech-gesture incongruity (e.g., conflict due to beating downward while talking and thinking about rockets going upward, or temperatures rising). Beats could also help to activate or maintain spatial schemas as an aid in recalling a story's details or structuring its narrative flow. On either of these accounts, beats would serve a speaker-internal, cognitive function (Chu & Kita, 2016). Alternatively, or in addition, beats could serve a communicative function, conveying a story's literal or metaphorical spatial content to listeners.

To investigate beats' cognitive and communicative functions we first tested whether speakers continue to produce beats when their listeners are blindfolded, and found that the rate of schema-congruent beats remained far above chance. Although this finding is consistent with beats serving a speaker-internal function, the blindfold data are open to an alternative interpretation. Unseen gestures could be produced because gesturing is an "overlearned social behavior" (Nass & Moon, 2000). Even if gesture's function were purely communicative, speakers could continue to gesture to blindfolded listeners out of habit.

To resolve this ambiguity, we tested whether beats show a hallmark of gestures that are intended to be communicative: recipient design (Bavelas et al., 2002). Speakers and listeners were separated by a barrier that occluded much of their usual gesture space. If gestures were designed to be communicative, speakers could easily raise them above the barrier; but they did not. Speakers continued to produce schema-congruent beats at a rate far above chance, but 98% of their beats were completely hidden from the listener. The finding that most beats were *not* recipient designed strengthens the argument that they were produced for speaker-internal purposes, and provides no evidence that they were intended to be communicative.

The only aspect of the data that supports a communicative role for beats comes from the finding that schema-congruity rates were highest when gestures were visible. Together, the results suggest that beat gestures during storytelling may serve both cognitive and communicative functions in the speaker's mind. It remains an open question whether listeners can gather information by following the directions of speakers' beats. If so, this information gathering is likely to be unconscious, given that even gesture researchers have been unaware that beats encode information about speakers' thoughts.

Rethinking how gestures are affiliated with speech

Speakers tended to produce beat gestures that were congruent with the dominant spatial schemas of the stories they told even during "non-directional" clauses with no literal or metaphorical spatial content (Fig. 2.3; Table 2.1). This pattern is difficult to explain on traditional models of speech and gesture, which assume that gestures are affiliated with speech locally, at the level of the word, phrase, or clause (Morrel-Samuels & Krauss, 1992). Here we observed a pattern of non-local affiliation: Beats matched the spatial schema that structured each story as a whole, even when there was no spatial content (literal or metaphorical) in the co-occurring word, phrase, clause, or sentence.

Contrary to the assumption that each co-speech gesture has a lexical affiliation with a particular spoken word (Morrel-Samuels & Krauss, 1992), the beats we observed during non-directional clauses (and perhaps other clauses, as well) may be best described as having a *conceptual affiliation* with the spatial schema that organized speakers' story-level representations of the stories (Casasanto & Boroditsky, 2008).

Beats provide a tool to detect automaticity of spatial thinking

Beat gestures can help to answer questions that have been debated broadly in the cognitive sciences concerning how automatically people activate literal and metaphorical spatial schemas to support thinking and language use. Numerous studies have used Stroop-like spatial congruity tasks to determine whether people activate schematic spatial representations when they understand words associated with particular locations in literal space (e.g., floor, ceiling) or metaphorical space (e.g., *happy* is up, *sad* is down;Casasanto, Brookshire, and Ivry, 2015). On the basis of null results, several studies have concluded that the activation of spatial schemas is "not automatic" (Lebois, Wilson-Mendenhall, & Barsalou, 2015; Santiago, Ouellet, Román, & Valenzuela, 2012). Yet, effect sizes in these reaction time experiments are typically small, and null results could reveal more about the sensitivity of these tasks than about the automaticity of the spatial schemas they are attempting to measure.

The circumstances under which our participants produced schema-congruent gestures suggest that the automaticity with which people activate spatial schemas to support literal and metaphorical thinking may have been underestimated by the methods used previously. Here, people activated story-congruent spatial schemas (i.) while producing beats, which they may not have been aware they were producing, (ii.) during directional clauses with no spatial language, (iii.) during non-directional clauses with no spatial significance whatsoever, and (iv.) when they knew their listener was blindfolded, and could not possibly interpret their gestures. Together, these circumstances suggest that story-congruent beats and the spatial representations they index in gesturers' minds are activated highly automatically. Beats may reveal spatial schemas in the mind that cannot be detected reliably with Stroop-like tasks, long held as the gold standard of automaticity.

2.3.4 Study 1 conclusions

According to decades of research, beat gestures are meaningless. Here we challenge this widely held belief in a quantitative analysis of over 3000 beats. Beats were the most frequent type of gesture in our corpus of over 5000 gestures. The directions of speakers' beats corresponded to the ideas that they expressed no matter whether they were using space literally or metaphorically, whether or not speakers used spatial language, and whether or not their gestures could be seen by the listener. Evidence that beat gestures are pervasive and meaningful motivates a reconsideration of beats in gesture research. Iconic relationships in emblems, iconic and metaphoric gestures, that have long been recognized by researchers are also expressed in beats.

Furthermore, the hidden systematicity of beat gestures was remarkably robust. Across three types of language (Fig. 2.2A) and three communicative contexts (Fig. 2.2C) the observed rate of gesture congruity was greater than all 1,000,000 of the chance congruity rates generated by permutation tests. Therefore, blind, hypothesis-driven, quantitative analyses of beat gestures provide a powerful tool for detecting the automaticity with which people activate spatial schemas to scaffold even highly abstract concepts.

CHAPTER 3

METAPHORIC ICONICITY IN SIGNED AND SPOKEN LANGUAGES

Discoveries about hidden iconicity in gesture may have implications for how we think about arbitrariness and iconicity in language. Even though gestures are an integral part of language (McNeill, 1992), they are thought to be different. Unlike gesture, sign language has phonology, morphology and a syntax, and aspects of words are assumed to be conventional (Stokoe Jr, 2005). Despite these differences, could the same principles of spatial thinking that motivate the forms of beat gestures also motivate aspects of words in sign language?

In metaphoric gestures, abstract ideas that we can never see or touch can nevertheless be represented with the hands via conceptual metaphor (Lakoff & Johnson, 2008). People often talk about abstract, non-spatial entities using spatial words (e.g., a long time, a high price, or a close friendship). Beyond talking in linguistic metaphors, there is a growing body of evidence suggesting that people also think in mental metaphors (Casasanto & Bottini, 2014): implicit associations between non-linguistic representations in abstract target domains and relatively concrete source domains like space, force, and motion (Lakoff & Johnson, 2008). Although target domains like time are nearly impossible to depict gesturally, per se, their source domains can often be depicted: A long time can be indicated by a long-distance sweep of the hand; a distant time can be represented by gesturing toward a far-away point in space (Cienki & Müller, 2008).

Although the evidence for metaphoric iconicity in gestures is strong, this type of iconicity is generally assumed not to extend to language. Even signed languages, which share a modality with hand gestures and therefore have the potential to express spatial iconicity, have been characterized as exhibiting largely arbitrary form-meaning mappings, in part for historical reasons having to do with establishing American Sign Language (ASL) as a full-fledged language, and not a simple system of pantomimes (Klima Edward & Bellugi, 1980).

Taub (2001) noted that signed languages' potential for iconicity is expanded by their ability to depict aspects of metaphoric source domains in sign, as in gesture. She and others have shown metaphoric iconicity in a number of ASL signs (Emmorey, 2001). In a multiple-choice test, non-signers were able to match the meanings of some metaphoric signs in ASL to their English glosses (O'Brien, 1999).

Taub (2001) reviewed a set of signs motivated by the metaphor Good is Up / Bad is Down, which spatializes emotional valence on a vertical continuum, and is evident in many spoken languages (e.g., feeling on top of the world or down in the dumps (Lakoff & Johnson, 2008). Taub (2001) describes several signs related to the notion of improvement or deterioration that make use of vertical motions to express positive or negative valence, consistent with the mental metaphor Good is Up.

We propose that rather than being a set of isolated cases, the examples of metaphoric iconicity in signed languages that have been described to date are only the tip of the iceberg. Metaphors for ubiquitous qualities such as positive and negative emotional valence may generate iconic relationships throughout the lexicon, making true arbitrariness of the sign vanishingly rare. In the domain of valence, ASL provides some metaphor-congruent examples: the sign for "bad" moves downward and the sign for "happy" moves upwards. However, there are also some signs that show the reverse mapping. For instance, the sign for "good" moves downward, and the sign for "insult" moves upward: it is not possible to infer from hand-picked examples like these whether spatial direction is correlated with the valence of words in the ASL lexicon, in general.

Spoken languages could also encode space-valence mappings in the forms of words. Since pitch is metaphorically mapped onto a vertical spatial continuum in many languages and cultures (Dolscheid, Shayan, Majid, & Casasanto, 2013), including in Mandarin Chinese, lexical tones in Mandarin could also be a source of metaphoric iconicity. Indeed, even though signed languages have more iconic form-meaning mappings than spoken languages, examples of metaphoric iconicity can be observed in the lexical tones of Mandarin. For instance, 能(capable; néng) is a positive word and 恨(hate; hèn) is a negative word, and they have a rising and falling tone, respectively. Like in ASL, there are also exceptions, such as 仇(hatred; chóu), a negative word with a rising tone, and 爱(love; ài) a positive word with a falling tone.

Because it is possible to find some examples that support our proposal and others that contradict it, we designed a quantitative study of corpora of signed languages (Study 2) and of Mandarin (Study 3) to determine whether there is any widespread systematic metaphoric iconicity in these languages. We hypothesized that vertical spatial metaphors for valence should be manifested in language through space in signed languages and through the spatialized dimension of pitch in spoken languages. We predicted that, on average, signs with upward "lexical movements" (Brentari & Padden, 2001) and Mandarin words with rising pitch contours should be the most positive in valence, consistent with the spatial metaphor Good is Up. By contrast, downward movements and pitch contours should be the most negative in valence (Bad is Down). Sign movements and pitch contours that do not move upward or downward should be intermediate in valence, on average. Valence ratings were taken from a corpus of English words (Bradley & Lang, 1999) that included some expressly evaluative words like "improve" but a great majority of non-evaluative words that range in valence from the strongly positive (e.g., leader, admired, adorable) to the strongly negative (e.g., blackmail, derelict, evil). These words have no spatial meanings, and do not need to be used in metaphorical constructions to convey positive or negative valence [1].

1. A preliminary report on these experiments appeared as Yap, D., Staum Casasanto, L., & Casasanto, D. (2014). *Metaphoric iconicity in signed and spoken languages.* Proceedings of the 36th Annual Conference of the Cognitive Science Society. Austin, TX, 1808-1813

3.1 Experiment 2a: Space and valence in ASL

3.1.1 Methods

Materials

We searched an online ASL dictionary (http://www.handspeak.com) for all 1034 of the words in the ANEW corpus (Affective Norms of English Words; Bradley and Lang, 1999): a set of words that were rated for valence on a 9-point scale by a large number of English speakers, and which have been used as stimuli in many experiments. We found 606 ANEW words that had clear translation equivalents in ASL. To ensure that the list of signs to be analyzed was constructed in an unbiased manner, translation equivalence was determined on the basis of the English glosses provided by the ASL dictionary; the experimenter was blind to the forms of the signs during list construction. The duration of each silent sign video was two seconds.

Sign analysis

The goal of the sign analysis was to determine the relationship between the vertical direction of the "lexical movement" phase of each sign and the valence of its English translation equivalent. The lexical movement phase of a sign is an invariant part of its phonology (Brentari & Padden, 2001), and part of the meaning-bearing portion of the sign. Like the stroke phase of a gesture, the movement phase can be identified on the basis of its form (McNeill, 1992; Kita, Gijn, & Hulst, 1998). A sign begins from a location and handshape that is a "hold" or starting position and entails a movement to a separate location or a change in the handshape. The directions of the preparation and retraction phases (i.e., "transitional movements") are generally not meaningful, and their directions were not analyzed.

All 606 signs were randomized and coded by one of the authors (D.Y.) who was nave to all signed languages. He was also blind to the signs' translation equivalents in English, and therefore to their meanings. The movement phase of each sign was coded for its vertical direction: Upward, Downward or Non-vertical. Signs with horizontal movement phases or "holds" were coded as Non-vertical signs. Some signs constitute a series of multiple movements and were coded as compound signs. Compound signs with movements in more than one direction were coded based on the direction that appeared to be dominant. For instance, the ASL sign for curtain consists of two movements that begin with a right closed fist moving to the right from the left closed fist, followed by a downward movement with open palms from both hands. It was coded as a Downward sign. The signs were then randomized again and 25% of the signs were selected for a second blind coding to determine the intra-rater reliability. The intra-rater agreement rate was 91% (139 out of 153 signs; $Kappa = .83, p = .001$).

3.1.2 Results and discussion

On average, signs with Upward movements ($n = 59$) were the most positive in valence, followed by Non-vertical signs ($n = 346$) and then by Downward signs ($n = 201$; Fig. 3.1). There was a significant relationship between the vertical direction of the signs and the valence of their ANEW translation equivalents, $F(2, 603) = 4.54, p = .01$. Upward signs were more positive than Downward signs, $p = .007$.

In summary, the vertical direction of the ASL signs predicted the valence ratings of their ANEW translation equivalents. Signs with upward movements were the most positive in valence, and signs with downward movements the most negative. Signs with non-vertical strokes were intermediate in valence.

These results support our hypothesis that the implicit mental metaphor Good is Up is manifested in the conventionalized forms of ASL words. Testing this hypothesis using English valence norms, rather than collecting new norms for these words in ASL, avoids circularity: Native ASL raters could be biased by the signs' movement directions, online, as they performed the ratings. Translation equivalence between ASL and English words is unlikely to be exact, but importantly, any noise introduced by inexact translations works

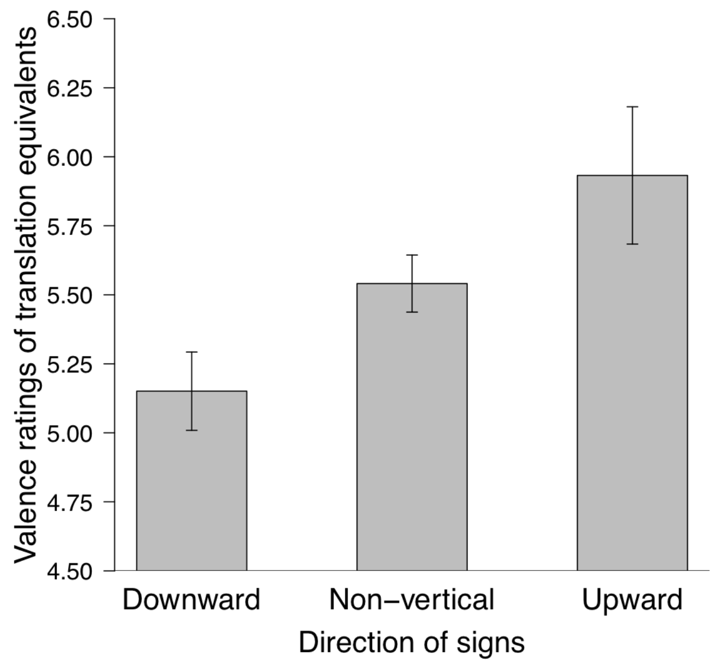

Figure 3.1: Valence ratings of the ANEW translation equivalents for ASL signs with downward (left), upward (right) and non-vertical (middle) strokes. The error bars show the standard error of the mean ($n = 606$).

against our hypothesis.

3.2 Experiment 2b: Space and valence in LSF

To generalize this novel result, we conducted the same analysis in French Sign Language (LSF).

3.2.1 Methods

Materials

The ANEW words were translated into French by a native speaker. We searched the LSF dictionary (http://www.lsfdicoinjsmetz.fr/index.php?page=motsalphalsf) for all of the ANEW words, and found 490 that had clear translation equivalents in LSF. Thirty words were translated twice into 30 nouns and 30 verbs because the ANEW corpus did not specify the word class. The duration of each silent sign video was three seconds.

Sign Analysis

The signs were analyzed in the same way as in Experiment 2a. The intra-rater agreement was 92% ($Kappa = .86, p = .001$; 113 out of 123 signs).

3.2.2 Results and discussion

The relationship between sign movement direction and valence in LSF replicated the results in ASL (see Figure 3.2). Upward signs were the most positive ($n = 78$), followed by Non-vertical signs ($n = 277$) and Downward signs ($n = 135$). There was a significant relationship between the vertical direction of the signs and the valence of their ANEW translation equivalents, $F(2, 487) = 4.55, p = .01$. Upward signs were more positive than Downward, signs, $p = .003$.

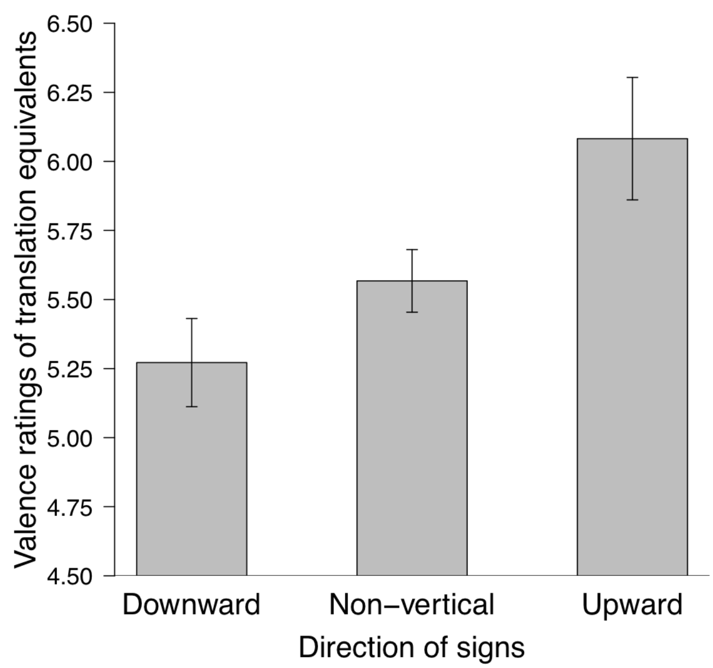

Figure 3.2: Valence ratings of the ANEW translation equivalents for LSF signs with downward (left), upward (right) and non-vertical (middle) strokes. The error bars show the standard error of the mean ($n = 490$).

Half of the 30 ANEW English words that were translated twice were coded with the same direction. After excluding 15 pairs of French signs with different directions and the duplicate signs with the same direction, there was a significant relationship between the vertical direction of the signs and the valence of their ANEW translation equivalents, $F(2, 487) = 4.55, p = .03$. Upward signs were more positive than Downward signs $p = .007$. The vertical direction of the LSF signs predicted valence ratings of their ANEW translation equivalents, replicating our findings in ASL.

3.3 Experiment 2c: Space and valence in BSL

Although ASL and LSF are not mutually intelligible, they are genetically related. We sought to generalize these findings further by testing our hypothesis in a third, genetically unrelated language, British Sign Language (BSL).

3.3.1 Methods

Materials

We searched the BSL dictionary (http://www.signstation.org/index.php/bsldictionary/desktop-dictionary) for all of the ANEW words (Bradley & Lang, 1999), and found 458 that had clear translation equivalents in BSL. The duration of each silent sign video was two seconds.

Sign analysis

The signs were analyzed in the same way as in Experiments 2a-b. The intra-rater agreement was 96% ($Kappa = .92, p < .001$; 110 out of 115 signs).

3.3.2 Results and discussion

The BSL results replicated the results of ASL and LSF (see Figure 3.3). Upward signs were consistently the most positive in valence ($n = 65$), followed by Non- vertical signs ($n = 281$) and Downward signs ($n = 112$). Overall, there was a marginally significant relationship between the vertical direction of the signs and the valence of their ANEW translation equivalents, $F(2, 455) = 2.33, p = .099$. Upward signs were significantly more positive than Downward signs, $p = .03$.

Analysis of the BSL corpus showed a similar relationship between space and valence as shown for ASL and LSF, in a genetically unrelated sign language.

3.4 Experiment 3: Space via pitch and valence in Mandarin

Although iconicity in signed languages has been discussed in the linguistic literature, iconicity in spoken languages is considered to be rare (Perniss et al., 2010). Mandarin Chinese is a tonal language with four basic lexical tones: (1) a high, level tone, (2) a rising tone, (3) a low falling tone (but with a rising tail in single characters) and (4) a high falling tone. Figure 3.4 (based on Speer et al., 1989) shows the four tones with the corresponding tone contours.

In some multisyllabic words, a group of characters (e.g., 子) assume the neutral tone. The pitch value of a neutral tone is influenced by its preceding tone. For instance, the neutral tone following the four lexical tones are [55-2], [35-3], [21-4] and [51-1] respectively (Duanmu, 2007).

3.4.1 Method

Materials

The entire ANEW corpus of 1034 words was translated into Chinese characters and their respective pinyin using the Goggle Translator (http://translate.google.com/♯en/zh-CN/). A native Mandarin speaker then reviewed the list and edited 176 words.

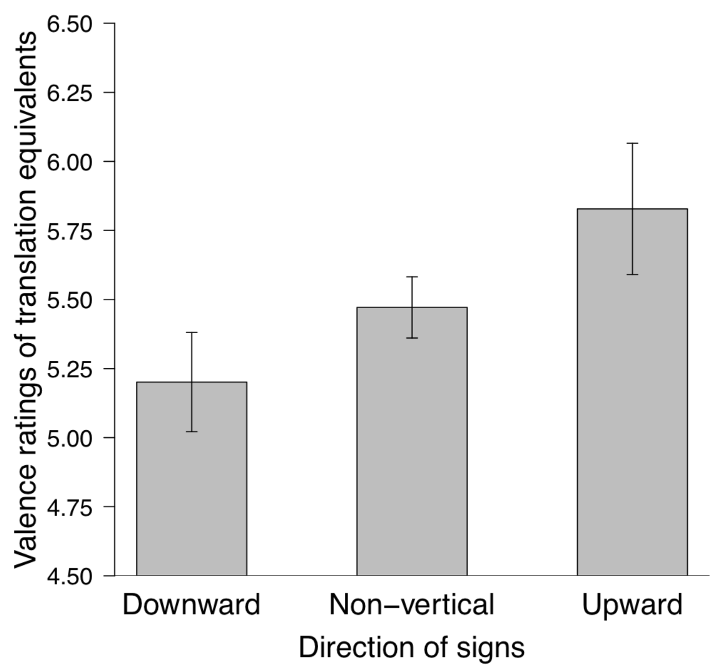

Figure 3.3: Valence ratings of the ANEW translation equivalents for BSL signs with downward (left), upward (right) and non-vertical (middle) strokes. The error bars show the standard error of the mean ($n = 458$).

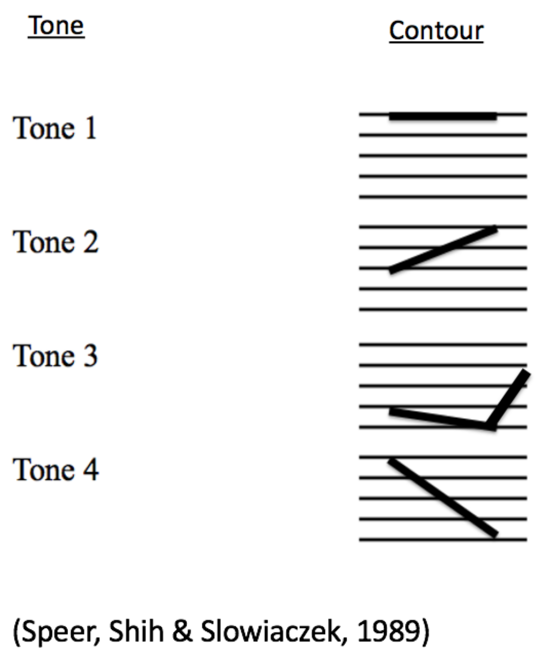

Figure 3.4: Pitch Contours of the four lexical tones in Mandarin Chinese. Each lexical tone is schematized along a vertical axis of five units (Speer et al., 1989).

Tone analysis

The four lexical tones were classified into three vertical pitch movement categories according to their respective pitch contours as shown in Figure 3.4. Tone 1 is defined as a Level pitch contour because its pitch value remains at 5. Tone 2 (rising tone) is defined as an Upward pitch contour that moves up two pitch values (from 3 to 5). Tones 3 and 4 fall one pitch value (from 2 to 1) and 4 pitch values (from 5 to 1) respectively and thus are defined as Downward pitch contours.

The pitch analysis of multisyllabic words considers the entire word as a continuous pitch contour because we want the level of analysis to be as similar to the natural speech stream as possible. Therefore, the overall pitch movement is defined as the sum of all the individual tone's vertical pitch movements. The pitch transition of two tones is also included in the calculation of the overall pitch movement. A resulting positive sum constitutes an Upward pitch contour, and a negative sum a Downward pitch contour. A sum of 0 constitutes a Level pitch contour. For example, a 2-character word, 食品(tone 2 and tone 3) as shown in Figure 3.4, would result in the value of (5 3) 3 + (1 2) = -2 which classifies it as a continuous Downward pitch contour.

3.4.2 Results and discussion

The vertical pitch contours of the Mandarin pinyin were analyzed at three levels: (1) the entire corpus, (2) monosyllabic words, and (3) multisyllabic words. Entire corpus Valence ratings for the Mandarin Chinese words' ANEW translation equivalents were highest for words with Upward pitch contours ($n = 226$), followed by Level ($n = 227$) and Downward ($n = 581$) pitch contours respectively. The vertical pitch contours predicted the valence of the words' ANEW translation equivalents, $F(2, 1031) = 5.52, p = .004$. Words with Upward pitch contours were more positive in valence than words with Downward pitch contours, $p = .001$ (see Figure 3.5).

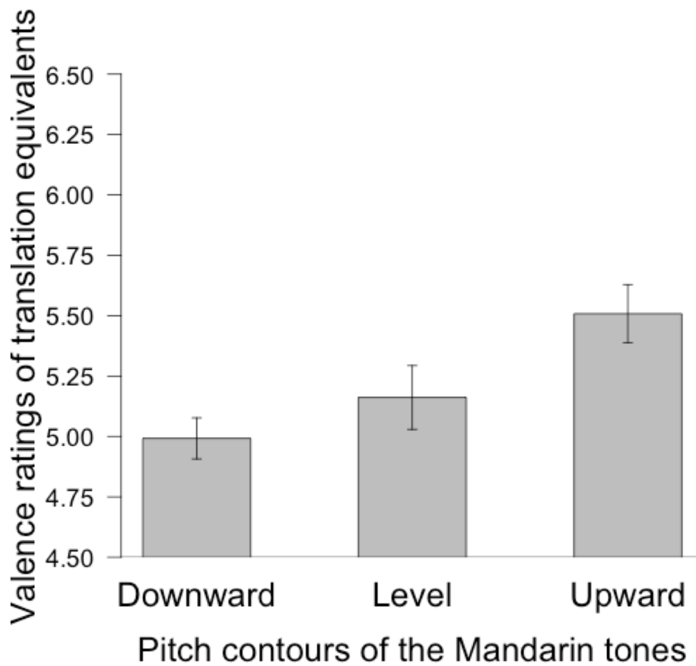

Figure 3.5: Valence ratings of the ANEW translation equivalents for Mandarin tones with downward (left), upward (right) and level (middle) pitch contours. The error bars show the standard error of the mean ($n = 1034$).

Monosyllabic words

The analysis of 133 monosyllabic (single character) words replicated results in the whole corpus. Valence ratings in ANEW are highest in Upward pitch contours ($n = 70$; $5.50 \pm .21$), followed by level pitch contours ($n = 31$; $4.59 \pm .29$) and downward pitch contours ($n = 32$; $4.24 \pm .33$). The vertical pitch contours predicted the valence of the words' ANEW translation equivalents, $F(2, 130) = 6.82, p = .002$. Words with Upward pitch contours were more positive in valence than words with Downward pitch contours, $p = .001$.

Multisyllabic words

The analysis of 861 multisyllabic (more than one character) words based on the overall pitch contours also provided converging results. Valence ratings in ANEW are highest in Upward pitch contours ($n = 156$; $5.51 \pm .15$), followed by Level pitch contours ($n = 196$; $5.25 \pm .15$) and Downward pitch contours ($n = 549$; $5.04 \pm .09$). The vertical pitch contours predicted the valence of the words' ANEW translation equivalents, $F(2, 898) = 3.62, p = .03$. Words with Upward pitch contours were more positive in valence than words with Downward pitch contours, $p = .009$.

Pitch contour was a significant predictor of the words' ANEW translation equivalents. Words with rising tones were the most positive, words with falling tones the most negative, and words with level tones intermediate in valence. The Mandarin Chinese corpus replicated the findings of the three signed languages and demonstrated the same vertical spatial metaphors for valence.

3.4.3 Study 2-3 General Discussion

Here we demonstrate a previously undiscovered relationship between form and meaning, in three signed languages and a spoken language. The vertical direction of signs predicted the valence ratings of their ANEW translation equivalents for all three signed languages. On average, signs with upward lexical movements were the most positive in valence, and signs with downward movements the most negative. Signs with non-vertical movements were intermediate in valence. Likewise, in Mandarin Chinese, a tonal spoken language, words with upward pitch contours were more positive in valence than words with downward pitch contours, and words with level pitch contours were intermediate.

Why is this particular non-arbitrariness preserved in the lexicon, across signed and spoken languages? One possible reason is that metaphoric iconicity makes words easier to learn. Activating mental metaphors via simple motor actions can improve word learning. In one study (Casasanto & de Bruin, in revision), students learned the definitions of positive and negative words better after moving vocabulary flash cards in a vertical direction consistent with the Good is Up metaphor. The same principle could facilitate the learning of metaphor-congruent words in signed languages and tone languages.

If metaphoric iconicity improves word learning, why isn't non-arbitrariness more pervasive in languages? That is, why don't all positive words have upwards movements or tones, and all negative words downward movements / tones? One possible explanation is that there may be many weak iconic (and other) constraints on word forms. Thus, the spatial metaphors described in this study (Good is Up and Bad is Down) could be the source of one such constraint, but they operate in the context of many others.

Another reason that iconicity in language might be limited: perhaps both arbitrary and non-arbitrary mappings have roles to play in language. Computational analyses and findings from an artificial language learning study demonstrated that both arbitrariness and non-arbitrariness facilitate word learning via complementary functions (Monaghan, Christiansen, & Fitneva, 2011). Specifically, non-arbitrariness facilitates the generalization of words to semantic categories while arbitrariness facilitates the mapping of words to specific meanings.

Form-meaning relationships are not as arbitrary as was once assumed. Beyond special cases like onomatopoeia, implicit metaphorical mappings may provide opportunities for multiple kinds of non-arbitrariness, throughout the lexicons of signed and spoken languages.

CHAPTER 4

CONCLUSIONS

Together, the studies presented here provide evidence for previously hidden iconicity in gesture, sign and spoken language. Study 1 shows that beat gestures reflect the spatial content of speakers' thoughts, no matter whether space is being used literally or metaphorically, whether or not speakers are using spatial language, and whether or not their gestures can be seen by the listener. Studies 2-3 show that metaphoric iconicity is encoded in obligatory aspects of the forms of words in three signed languages and a spoken language.

These studies challenge three beliefs that are deeply entrenched in theories of human communication: (i.) that beat gestures are meaningless, and only serve pragmatic functions; (ii.) that iconicity primarily constrains the forms of words referring to concrete objects, actions, or physical attributes; (iii.) that iconic form-meaning relationships in language are limited to small clusters of words.

Study 1 shows that, in addition to the pragmatic functions that beats serve by virtue of their timing, beats also encode semantic content by virtue of their direction. Beats constituted the majority of gestures in our corpus, suggesting that by overlooking the spatial semantics of beats, gesture researchers have been unaware of much of the meaning that is encoded in co-speech gestures. Studies 2 and 3 show that, via implicit mental metaphors, spatial iconicity can constrain form-meaning relationships for words with concrete and abstract meanings, alike, and that iconic relationships can be found throughout the lexicons of signed and spoken languages.

These findings suggest a revision to current theories that distinguish two different types of non-arbitrariness in language: iconicity and systematicity. Whereas iconicity is a physical resemblance between the form of an individual sign and the form of its referent, systematicity is a statistical relationships between forms and meanings (Farmer, Christiansen, & Monaghan, 2006; Dingemanse et al., 2015). For example, the lexical categories of nouns and verbs in English were found to have two different clusters of phonological properties (Farmer et al., 2006). Whereas iconic relationships are thought to be limited in scope, systematic relationships are by definition found throughout a lexicon. Whereas the same iconic relationships are posited to arise across many languages, systematic relationships are thought to be language-specific. Whereas the phonological features that give rise to iconicity are often easy to identify, the phonological features underlying systematicity are thought to be obscure, therefore systematic relationships are discoverable only through corpus analyses.

The form-meaning relationships revealed in studies 2 and 3, however, break down these divides between iconicity and systematicity, and demonstrate a previously undocumented form of non-arbitrariness: *systematic iconicity.* Form-meaning relationships in ASL, LSF, BSL, and Mandarin are iconic, but they are not limited in scope: Since every content word has some valence, *every content word* is eligible to participate in a systematic mapping between valence and height (in space or pitch), within and across lexicons. Since this height-valence mapping is only one of many factors that constrain form-meaning relationships, the extent to which this iconic mapping motivates the forms of words can only be discovered through the kind of hypothesis-driven corpus analyses we introduce here.

The discovery of systematic iconicity in gesture, sign, and spoken language – iconicity that is visible in corpus analyses but invisible to the naked eye – gives rise to a radical possibility. Languages may be shaped by a lattice of weak iconic constraints. The space-valence mapping shown here may be one "thread" among many in a network of iconic threads (among other kinds of threads) which, together, cause each word to occupy a particular location within a high-dimensional space of phonology and semantics. On this proposal, form-meaning relationships *appear* largely arbitrary because the threads of this lattice are weak and numerous enough for the effects of any single thread to go unnoticed – unless its effects are tested for explicitly, via quantitative analyses.

The studies presented here tested for the effects of only one thread, the space-valence mapping, but we found evidence that this thread constrains form-meaning relationships across all five of the corpora we tested (beats, ASL, LSF, BSL, and Mandarin), and across all three communicative forms (gesture, sign, speech). In ongoing studies, we have discovered another thread that appears to exert a similar influence on form-meaning relationships across signed and spoken languages: the space-arousal mapping. Like emotional valence, emotional arousal is often described using vertical spatial metaphors (more aroused is up, less aroused is down), and we have found that words' arousal ratings significantly predict the directions of lexical movements in ASL, LSF, and BSL, and the directions of pitch contours for tones in Mandarin. Space-valence and space-arousal mappings are only two among hundreds of mental metaphors that may be shaping our thoughts, and may be constraining form-meaning relationships in our communicative signals via systematic iconicity. If so, then true arbitrariness of the sign may be vanishingly rare.

Exploring this hypothetical lattice of weak iconic constraints (and other non-arbitrary constraints) via methods like we introduce here is a first step toward discovering how and why these hidden form-meaning relationships are preserved in human communicative behavior – and how they have shaped the evolution of languages, and perhaps of the language faculty itself.

REFERENCES

Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: Some gestures are meant to be seen. *Journal of Memory and Language*, *44*(2), 169–188.

Alibali, M. W., Kita, S., & Young, A. J. (2000). Gesture and the process of speech production: We think, therefore we gesture. *Language and cognitive processes*, *15*(6), 593–613.

Bavelas, J., Beavin, Chovil, N., Lawrie, D. A., & Wade, A. (1992). Interactive gestures. *Discourse processes*, *15*(4), 469–489.

Bavelas, J., Kenwood, C., Johnson, T., & Phillips, B. (2002). An experimental study of when and how speakers use gestures to communicate. *Gesture*, *2*(1), 1–17.

Beattie, G., & Coughlan, J. (1999). An experimental investigation of the role of iconic gestures in lexical access using the tip-of-the-tongue phenomenon. *British Journal of Psychology*, *90*(1), 35–56.

Biau, E., & Soto-Faraco, S. (2013). Beat gestures modulate auditory integration in speech perception. *Brain and language*, *124*(2), 143–152.

Biau, E., & Soto-Faraco, S. (2015). Synchronization by the hand: the sight of gestures modulates low-frequency activity in brain responses to continuous speech. *Frontiers in human neuroscience*, *9*.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for english words (anew): Instruction manual and affective ratings* (Tech. Rep.). Citeseer.

Brentari, D., & Padden, C. (2001). Native and foreign vocabulary in american sign language: A lexicon with multiple origins. *Foreign vocabulary in sign languages: A cross-linguistic investigation of word formation*, 87–119.

Buchler, J. (2014). *The philosophy of peirce: Selected writings.* Routledge.

Casasanto, D. (2013). Gesture and language processing. *Encyclopedia of the Mind*, 372–374.

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time.

Cognition, *106*(2), 579–593.

Casasanto, D., & Bottini, R. (2014). Spatial language and abstract concepts. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(2), 139–149.

Casasanto, D., Brookshire, G., & Ivry, R. (2015). Meaning is not a reflex: Context dependence of spatial congruity effects. *Cognitive science*, *39*(8), 1979–1986.

Casasanto, D., & de Bruin, A. (in revision). Metaphors we learn by: Directed motor action improves word learning.

Casasanto, D., & Jasmin, K. (2012). The hands of time: Temporal gestures in english speakers.

Chu, M., & Kita, S. (2016). Co-thought and co-speech gestures are generated by the same action generation process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(2), 257.

Chu, M., Meyer, A., Foulkes, L., & Kita, S. (2014). Individual differences in frequency and saliency of speech-accompanying gestures: the role of cognitive abilities and empathy. *Journal of Experimental Psychology: General*, *143*(2), 694.

Cienki, A. (1998). Metaphoric gestures and some of their relations to verbal metaphoric expressions. *Discourse and cognition: Bridging the gap*, 189–204.

Cienki, A., & Müller, C. (2008). *Metaphor and gesture* (Vol. 3). John Benjamins Publishing.

Clark, H. H. (2004). Pragmatics of language performance. *The handbook of pragmatics*, 365–382.

de Ruirer, J. P. (2000). 14 the production of gesture and speech. *Language and gesture*, *2*, 284.

Dimitrova, D., Chu, M., Wang, L., Özyürek, A., & Hagoort, P. (2016). Beat that word: How listeners integrate beat gesture and focus in multimodal speech discourse. *Journal of cognitive neuroscience*.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science*, *24*(5), 613–621.

Duanmu, S. (2007). *The phonology of standard chinese*. Oxford University Press.

Emmorey, K. (2001). *Language, cognition, and the brain: Insights from sign language research*. Psychology Press.

Engle, R. A. (1998). Not channels but composite signals: Speech, gesture, diagrams and object demonstrations are integrated in multimodal explanations. In *Proceedings of the twentieth annual conference of the cognitive science society* (pp. 321–326).

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*(32), 12203–12208.

Goldin-Meadow, S. (1999). The role of gesture in communication and thinking. *Trends in cognitive sciences*, *3*(11), 419–429.

Goldin-Meadow, S., & Alibali, M. W. (2013). Gesture's role in speaking, learning, and creating language. *Annual review of psychology*, *64*, 257–283.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, *40*.

Gullberg, M. (1998). *Gesture as a communication strategy in second language discourse: A study of learners of french and swedish* (Vol. 35). Lund University Press.

Holle, H., Obermeier, C., Schmidt-Kassow, M., Friederici, A. D., Ward, J., & Gunter, T. C. (2012). Gesture facilitates the syntactic analysis of speech. *Frontiers in psychology*, *3*.

Hubbard, A. L., Wilson, S. M., Callan, D. E., & Dapretto, M. (2009). Giving speech a hand: Gesture modulates activity in auditory cortex during speech perception. *Human brain mapping*, *30*(3), 1028–1037.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition*, *109*(1), 54–65.

Kelly, S. D., Manning, S. M., & Rodak, S. (2008). Gesture gives a hand to language and learning: Perspectives from cognitive neuroscience, developmental psychology and education. *Language and Linguistics Compass*, *2*(4), 569–588.

Kita, S., Gijn, I. v., & Hulst, H. v. d. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In *Gesture and sign language in human-computer interaction: International gesture workshop, bielefeld, germany, september 1997. proceedings* (p. 23).

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.

Klima Edward, S., & Bellugi, U. (1980). *The signs of language.* Harvard University Press.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, *57*(3), 396–414.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of personality and social psychology*, *61*(5), 743.

Lakoff, G. (1980). *Metaphors we live by/george lakoff, mark johnson.* Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (2008). *Metaphors we live by.* University of Chicago press.

Lebois, L. A., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). Are automatic conceptual cores the gold standard of semantic processing? the context-dependence of spatial meaning in grounded congruency effects. *Cognitive Science*, *39*(8), 1764–1801.

Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, *26*(10), 1457–1471.

Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Phil. Trans. R. Soc. B*, *369*(1651), 20130302.

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

McNeill, D. (2015). *Why we gesture: The surprising role of hand movements in communication*. Cambridge University Press.

McNeill, D., Cassell, J., & McCullough, K.-E. (1994). Communicative effects of speech-mismatched gestures. *Research on language and social interaction*, *27*(3), 223–237.

Monaghan, P., Christiansen, M. H., & Fitneva, S. A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, *140*(3), 325.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(3), 615.

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of social issues*, *56*(1), 81–103.

Nuckolls, J. B. (2004). Language and nature in sound alignment. *Hearing cultures: essays on sound, listening and modernity*, 65–86.

O'Brien, J. (1999). Metaphoricity in the signs of american sign language. *Metaphor and Symbol*, *14*(3), 159–177.

Ohala, J. J., Hinton, L., & Nichols, J. (1997). Sound symbolism. In *Proc. 4th seoul international conference on linguistics [sicol]* (pp. 98–103).

Özyürek, A. (2002). Do speakers design their cospeech gestures for their addressees? the effects of addressee location on representational gestures. *Journal of Memory and*

Language, *46*(4), 688–704.

Özyürek, A. (2014). Hearing and seeing meaning in speech and gesture: insights from brain and behaviour. *Phil. Trans. R. Soc. B*, *369*(1651), 20130296.

Parrill, F., & Stec, K. (2018). Gestures of the abstract. *Pragmatics & Cognition*, *24*(1), 33–61.

Perniss, P., Thompson, R., & Vigliocco, G. (2010). Iconicity as a general property of language: evidence from spoken and signed languages. *Frontiers in psychology*, *1*, 227.

Sacks, H., Schlegoff, I., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language*, *50*, 696–735.

Santiago, J., Ouellet, M., Román, A., & Valenzuela, J. (2012). Attentional factors in conceptual congruency. *Cognitive Science*, *36*(6), 1051–1077.

Saussure, F. d. (1959). Course in general linguistics (w. baskin, trans.). *New York: Philosophical Library*.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, *55*(2), 167–177.

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665–681.

Stokoe Jr, W. C. (2005). Sign language structure: An outline of the visual communication systems of the american deaf. *Journal of deaf studies and deaf education*, *10*(1), 3–37.

Sweetser, E. (1998). Regular metaphoricity in gesture: Bodily-based models of speech interaction. In *Actes du 16e congrès international des linguistes.*

Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in american sign language.* Cambridge University Press.

Tellier, M. (2009). *The development of gesture.* Routledge.

Tuite, K. (1993). The production of gesture. *Semiotica*, *93*(1-2), 83–106.

Wang, L., & Chu, M. (2013). The role of beat gesture and pitch accent in semantic processing: an erp study. *Neuropsychologia*, *51*(13), 2847–2855.

APPENDIX A

STUDY 1 STORIES

Upward stories

LSL 1: Testing some model rockets

You're testing some new model rockets next to an abandoned factory. You light the first rocket's fuse. As it starts rising you hear a loud swoosh, but the rocket doesn't even rise above the first floor of the building. Your second rocket has a more powerful engine. It soars to the highest story of the building before it disappears from view.

MSL 1: Checking the weather

It's summer and you are listening to the weather report on the radio. The announcer says it's warm, and that the temperature is rising. He predicts that it will be 99 degrees today, but says the temperature could even rise above 105 degrees. This would be the highest temperature on record for this date.

NSL 1: Checking the weather

It's summer and you are listening to the weather report on the radio. The announcer says it's warm, and that the temperature is getting warmer. He predicts that it will be 99 degrees today, but says the temperature could even exceed 105 degrees. This would be the hottest temperature on record for this date.

MSL 2: Studying for the exam

You are studying for your final exam in the most difficult class you have ever taken. Your midterm grade was terrible – the lowest in the class, and you're determined to raise your grade on the final. If you can bring your score up by one letter grade, you'll qualify for the honor roll. If you can raise it by two letter grades, you'll receive high honors. And if you can boost your grade all the way up to an A+, you'll qualify for the Cum Laude Society, the highest honor in your school.

NSL 2: Studying for the exam

You are studying for your final exam in the most difficult class you have ever taken. Your midterm grade was terrible – the worst in the class, and you're determined to improve your grade on the final. If you can improve your score by one letter grade, you'll qualify for the honor roll. If you can improve it by two letter grades, you'll receive special commendation. And if you can transform your grade to an A+, you'll qualify for the Cum Laude Society, the most prestigious honor in your school.

Downward stories

LSL 1: Bungee jumping contest

You're a sports photographer, covering an international Bungee Jumping contest. The first contestant stands on the roof of the highest skyscraper in San Francisco. He starts to fall, and the crowd gasps. As he drops to the 10th floor, his bungee cord tightens and springs him back to the top. When the second contestant falls, he looks like he's going to plummet all the way to the bottom of the building, but his bungee cord catches him in the nick of time.

LSL 2: Scuba diving

You're scuba diving near a deep canyon. You plunge into the water and see the seafloor beneath you. You are on sandy shoal, but ahead of you is gorge where an ancient pirate ship has sunken. You descend into the gorge, and the water grows darker as you swim deeper and deeper. You reach an undersea ledge, but realize the floor of the gorge is even farther down. A glowing fish swims past your feet and darts downward toward the bottom of the gorge, where the wrecked ship must lie.

MSL 1: Buying a car

You are buying a used car, and trying to get it for the lowest price possible. The sticker price on the car that you want is far too high, so the salesman says he'll lower it a little bit. When you hesitate, he agrees to lower the price even farther. You still appear undecided, but the salesman swears this is the lowest price he can offer, and that his manager will fire him if he drops the price any lower.

NSL 1: Buying a car

You are buying a used car, and trying to get it for the cheapest price possible. The sticker price on the car that you want is much too expensive, so the salesman says he'll reduce it some. When you hesitate, he agrees to reduce the price even more. You still appear undecided, but the salesman swears this is the cheapest price he can offer, and that his manager will fire him if he cuts the price any more.

MSL 2: Basketball ranking

You're a star player on your college's basketball team. Your team is the highest ranked in the nation, but you've just been injured, and you worry that your team's ranking may begin to fall. If you lose to your arch-rival, you'll drop from the top ranked spot. If your injury causes a losing streak, your ranking could fall to the middle of the pack. And if another star player gets hurt, you're afraid your standing may plummet to the bottom of the entire division.

NSL 2: Basketball ranking

You're a star player on your college's basketball team. Your team is the best ranked in the nation, but you've just been injured, and you worry that your team's ranking may begin to get worse. If you lose to your arch-rival, you'll lose your excellent ranking. If your injury causes a losing streak, you could find yourself in the middle of the pack. And if another star player gets hurt, you're afraid your standing may suddenly be the worst in the entire division.

Leftward stories

LSL 1: The tightrope artist

You're watching a tightrope artist getting ready to walk across a highwire suspended between two towers. From the rightmost tower, she steps onto the long wire. You see that her balance is perfect for the first few yards, but farther along her she wobbles. She recovers and starts taking long strides toward a ring of fire that's suspended before her, encircling the highwire. You see her leap through the ring of fire, and approach the platform that lies before her on the leftmost tower, where her coach is waiting.

LSL 2: Stuck at the railroad crossing

You are stuck in your car at a railroad crossing. The locomotive engine is stopped in front of you, and train cars extend behind it to your left as far as you can see. The first car behind the engine has pink graffiti painted on the side. The next car is a livestock carrier filled with chickens. Farther to your left you see a hopper car overflowing with gravel. Down the line form that, you see a flatbed car carrying a backhoe. Looking as far as you can to the left, you see a red caboose.

MSL 1: Learning your family history

You are learning about your family's history. First, you find a picture of your great grandmother, taken long ago. Next, you discover a letter from her mother, who lived even farther in the past. You research your ancestry to the time of the earliest settlers, and wonder what life must have been like so long ago, before the invention of the computer, or the airplane, or the automobile, even before the time when houses had electric lights.

NSL 1: Learning your family history

You are learning about your family's history. First, you find a picture of your great grandmother, taken many years ago. Next, you discover a letter from her mother, who lived even earlier. You research your ancestry to the time of the earliest settlers, and wonder what life must have been like in this earlier age, prior to the invention of the computer, or the airplane, or the automobile, even prior to the time when houses had electric lights.

MSL 2: History of farming

You're writing a report on how farming methods changed over the course of the 20th century. Today farming is mostly automated, but long ago methods were much more primitive. Modern farmers use diesel-powered tractors for plowing their fields, but only a short time ago, half a century back, many farmers were still using horses. And only a generation before that, most farmers didn't even have horses and had to plow to their fields by hand. Looking back, you wonder how the farmer's lifestyle at the beginning of the 20th century differed from that of farmers today.

NSL 2: History of farming

You're writing a report on how farming methods changed during the 20th century. Today farming is mostly automated, but many years ago methods were much more primitive. Modern farmers use diesel-powered tractors for plowing their fields, but only a brief time ago, half a century earlier, many farmers were still using horses. And only a generation earlier than that, most farmers didn't even have horses and had to plow to their fields by hand. Thinking about the past, you wonder how the farmer's lifestyle at the beginning of the 20th century differed from that of farmers today.

Rightward stories

LSL 1: Watching a parade

You're watching a parade, and you see your best friend approaching from the left, riding on a float that celebrates the Roman Empire. His job is to operate the catapult anchored to the back of the float. As the float rolls by, you see the catapult project a styrofoam rock onto the float just ahead of them, which has a Hawaiian theme. Your friend loads the catapult again, but this time you see it projects a giant water balloon at the hula dancers on the float ahead.

LSL 2: Examining blueprints

You're examining the blueprints for a new Student Activities building on campus. The rooms area arranged along a single long corridor. On the far left you see a foyer that contains a fountain. Next to the foyer is a food court. To the right of the food court is a movie theater, and next to that is a bowling alley. On the right of the bowling alley there's a swimming pool, and all the way at the far right end of the corridor there's a fitness center with treadmills and weight lifting equipment.

MSL 1: Thinking about your future

You're about to graduate from college, and you're wondering what lies ahead. You're looking forward to starting your first real job, and you're thinking about other milestones in your life. First, maybe you'll get married, then have children, then maybe you'll even become a grandparent. Then some day, in the future, your children or your grandchildren might be graduating from college themselves, and trying to imagine what lies ahead in their lives, just as you are now.

NSL 1: Thinking about your future

You're about to graduate from college, and you're wondering what the future holds. You're anticipating starting your first real job, and you're thinking about other important events in your life. First, maybe you'll get married, then have children, then maybe you'll even become a grandparent. Then some day, eventually, your children or your grandchildren might be graduating from college themselves, and trying to imagine what will happen in their lives, just as you are now.

MSL 2: Future of robotics

You're discussing the past and future of robotics. Several decades back, novelists envisioned that by the start of the 21st century, robots would live among us and would look human. In reality, today's robots are found mostly on assembly lines, and look nothing like us. Scientists project that in the near future robots will replace soldiers on the battlefield. Eventually, by the start of the 22nd century, scientists project that nanotechnology will produce robots small enough to travel in our bloodstream. These nanobots will cure diseases in the future that are incurable today.

NSL 2: Future of robotics

You're discussing the past and future of robotics. Several decades ago, novelists envisioned that by the start of the 21st century, robots would live among us and would look human. In reality, today's robots are found mostly on assembly lines, and look nothing like us. Scientists expect that soon robots will replace soldiers on the battlefield. Eventually, by the start of the 22nd century, scientists expect that nanotechnology will produce robots small enough to travel in our bloodstream. These nanobots will cure diseases someday that are incurable today.