COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2014) 1–30

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12105

Investigating Constituent Order Change With Elicited Pantomime: A Functional Account of SVO Emergence

Matthew L. Hall,[a] Victor S. Ferreira,[b] Rachel I. Mayberry[c]

[a]*Department of Linguistics, University of Connecticut*
[b]*Department of Psychology, University of California, San Diego*
[c]*Department of Linguistics, University of California, San Diego*

Received 13 May 2012; received in revised form 13 May 2013; accepted 18 June 2013

Abstract

One of the most basic functions of human language is to convey who did what to whom. In the world's languages, the order of these three constituents (subject [S], verb [V], and object [O]) is uneven, with SOV and SVO being most common. Recent experiments using experimentally elicited pantomime provide a possible explanation of the prevalence of SOV, but extant explanations for the prevalence of SVO could benefit from further empirical support. Here, we test whether SVO might emerge because (a) SOV is not well suited for describing reversible events (a woman pushing a boy) and (b) pressures to be efficient and mention subjects before objects conspire to rule out many other alternatives. We tested this by asking participants to describe reversible and non-reversible events in pantomime, and we instructed some participants to be consistent in the form of their gestures and to teach them to the experimenter. These manipulations led to the emergence of SVO in speakers of both English (SVO) and Turkish (SOV).

Keywords: Constituent order; Pantomime; Sign language; Typology; Diachronic change; Lexicon; Language emergence; Functional explanations

1. Introduction

One of the hallmarks of human language is that it is organized at multiple levels. For example, all languages allow only certain sounds and combinations of sounds; our knowledge of language is what tells us that "hing" is a possible word in English, but "ngih" is not. Likewise, languages allow only certain combinations of words in sentences. For example, English allows us to say, "The woman pushed the box," but not typically "The

Correspondence should be sent to Matthew L. Hall, University of Connecticut, Department of Linguistics, Unit 1145, 365 Fairfield Way, Storrs, CT 06269-1145. E-mail: matthall.research@gmail.com

woman the box pushed." This level of organization is commonly called *constituent order*. The constituents in the preceding example can be described in both semantic and syntactic terms: "the woman" is the semantic agent (she does the action) and the syntactic subject; "the box" is the semantic patient (it receives the action) and the syntactic object; and "pushed" both identifies the semantic action and functions syntactically as the verb. Although our focus here is on the semantic categories of agent, patient, and action, we will describe these constituents with shorthand borrowed from the syntactic notions of subject (S), object (O), and verb (V).

English is an example of a language with fairly strict rules about constituent order. The grammatical example above uses SVO order; the ungrammatical example is SOV. And although it is possible to say, "The box was pushed by the woman" (OVS, in semantic terms), it is uncommon in writing, and even more rare in natural speech (e.g., Roland, Dick, & Elman, 2007). However, other languages have different rules about which orders are preferred, unusual, or ungrammatical. Turkish is an example of a language where the default order is SOV, while Hixkaryana (spoken in northern Brazil) is a language that primarily uses OVS. In fact, linguists have established that all six logically possible orders of S, V, and O are attested as basic orders in the languages of the world (e.g., Dryer, 2008). However, the distribution of these orders is extremely uneven, and these asymmetries have long been the subject of scrutiny in both linguistics and psychology for what they tell us about the way the human mind organizes the basic semantic elements of sentences.

One type of asymmetry concerns the distribution of constituent order across the languages of the world today, often called *synchronic variation*. Several large-scale surveys in recent decades (e.g., Dryer, 2008; Greenberg, 1963; Hawkins, 1983; Tomlin, 1986) have led to the widely accepted conclusion that the great majority of the world's languages belong to one of just two types: SOV or SVO, with SOV being slightly more common. VSO languages come in a distant third, with VOS, OVS, and OSV being extremely rare.

A second type of asymmetry concerns *diachronic variation*: predictable changes in constituent order over time. Just as not all constituent orders are equally common at any given point in time, it appears that languages do not change toward all orders equally often. Rather, there is a particularly intriguing pattern in which—once effects of geography and language contact are factored out—languages drift away from SOV, but not toward it; languages tend to drift toward SVO (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977).

Identifying the underlying sources of these asymmetries has motivated a great deal of linguistic and psycholinguistic research in recent decades. This work has revealed multiple sources of influence. For example, geographic proximity (*areal effects*) increases the probability of direct contact between languages, which is one way that constituent order can change (Dryer, 1992). A language's own history is another powerful predictor. Indeed, recent work using analytical methods borrowed from evolutionary biology has revealed that at least some long-standing claims about putative universals of language may in fact be attributable to the languages in question having a shared genealogy (Atkinson, 2011; Dunn, Greenhill, Levinson, & Gray, 2011; Levinson & Gray, 2012). In contrast, the primary focus of the present research is on a third potential source of language change: the nature of cognition itself.

Why this may be so can be illustrated by considering two thought experiments. First, consider whether and how language might first arise in a group of children who were raised under appropriate physical and social conditions but who lacked language input. Would they develop language at all and, if so, what would its structure be? Any systematicity in their behavior could be attributable to their shared cognitive preferences as humans. A second thought experiment would be to imagine a closed community of people who share a single language and who have no contact with speakers of any other language. If we could sample their language today and then again in 500 years, would there be any differences? If so, any differences we observe could not be due to areal or genealogical factors and would again be attributable to humans sharing certain cognitive preferences. Note that we use the term "cognitive" here in a very broad sense, encompassing communicative and social functions, since all complex human behavior is at some level a byproduct of the human mind. We intend it to stand in contrast to external influences on language change, such as contact with other languages. In all, it may be the case that over many successive iterations of production, comprehension, and acquisition, languages gravitate toward certain predictable states.

There are several approximations of experiments like these, and to the extent that diverse paradigms converge on similar answers, we can gain confidence in the accuracy of these approximations. One body of work has focused on the case of language emergence de novo, as in the first thought experiment above. This research considers how deaf children create inchoate communicative systems called *homesign*, and how the aggregation of multiple deaf homesigners (and sometimes their hearing interlocutors) can lead to the evolution of a full sign language. A robust finding in this literature is that when homesigners produce utterances with two arguments (e.g., subject and verb, or verb and object), they have a strong tendency to use SV and OV, rather than VS or VO (for a review, see Goldin-Meadow, 2003). This tendency is observed cross-culturally (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, Sancar, & Mylander, 2008) and cannot be attributed to the spoken or gestural input from hearing parents or the surrounding culture (Goldin-Meadow & Mylander, 1983). Furthermore, this pattern appears to carry over into the early stages of some young sign languages, where three-argument utterances (those with subject, verb, and object) are more common. Haviland (2011) reports preliminary evidence of SOV, SV, and OV order in a very young sign system used by a single family with three deaf siblings where the ambient spoken language, Tzotzil, uses VOS order. Sandler, Meir, Padden, and Aronoff (2005) have analyzed the constituent order of Al-Sayyid Bedouin Sign Language (ABSL), which has emerged over the past 75 years as a recessive gene for deafness has become relatively widespread within the village community. They have found that ABSL is a robust SOV language; that such consistency has emerged so quickly is all the more significant considering that this language has not yet converged on a unified phonological system (Aronoff, Meir, Padden, & Sandler, 2008). Nicaraguan Sign Language (ISN) emerged when large numbers of isolated homesigners were brought together in a school setting. Senghas, Coppola, Newport, and Supalla (1997) report that SOV, SV, and OV were commonly used by the first cohort of ISN signers, although OSV was also present. Thus, these findings are consistent with the idea that in the absence of linguistic input, something about human cognition privileges SOV order. However, no exhaustive typological work has yet established how stable this pattern is across all emerging sign languages, nor whether new sign languages vary in how long they retain SOV order as the language matures.

This latter question relates to the second thought experiment we proposed above: whether constituent order would change across time, even without outside influence from other languages. As mentioned previously, abundant research on spoken languages indicates that over long time scales, many languages do shift away from SOV and toward SVO, but not vice versa (Gell-Mann & Ruhlen, 2011; Givón, 1979; Li, 1977). Although all of the large-scale studies to date have focused exclusively on spoken languages, Fischer (1975) argues that the same SOV-to-SVO shift has also occurred in American Sign Language.

Although the phenomenon of constituent order change has been well described, assessing the role of cognition in shaping constituent order over such time scales is difficult. More compelling evidence would come from empirical demonstrations that constituent order preferences can vary under different circumstances even within a single individual and within a single testing session. Three recent studies have aimed to do just that by using a novel empirical technique for measuring constituent order preferences: elicited pantomime. In all three, naïve hearing participants of varying language backgrounds were shown pictures or video clips of simple events and were asked to describe those events in pantomime (i.e., gesture without speech). We discuss each study in turn.

Goldin-Meadow, So, Özyürek, and Mylander (2008) were the first to use elicited pantomime to assess constituent order preferences. They tested participants whose native language used SOV (Turkish), SVO (English, Spanish), or allowed both orders (Chinese) and found that when describing events in pantomime, speakers of all these language backgrounds showed a strong propensity to use orders that were consistent with SOV ordering (SOV, SV, and OV). In this respect, the data strongly resemble the constituent orders found in homesign and young sign languages, as discussed above. Goldin-Meadow, So et al. (2008) view these data as evidence that humans organize knowledge about events in a way that is most naturally linearized in SOV order, and they conclude that one reason SOV order is so prevalent among the world's languages is that it is compatible with mental representations of these events.

The elicited pantomime results account for the prevalence of SOV, but they do not explain why SVO is nearly as common or why languages commonly change from SOV to SVO but not vice versa. The second study, by Langus and Nespor (2010), investigated this issue. To explain the prevalence of SVO, Langus and Nespor suggest that there are two separate cognitive systems at work. One, the *conceptual* system, is responsible for non-linguistic processing and prefers SOV. The other, the *computational* system, is responsible for linguistic processing and prefers SVO. They replicated the basic SOV preference from Goldin-Meadow, So et al. (2008) and conducted a second experiment showing that pantomimed utterances did not have the characteristic structure that would be expected if the pantomimes were being generated by a linguistic (syntactic) system. In two additional experiments, they tested how quickly participants responded to various constituent orders when scenes were described with gestures (e.g., GIRL BALL THROW) versus synthesized speech that lacked prosody but had grammatical inflection (e.g., Italian and Turkish equivalents of "girl ball throws"). They found that for gestured stimuli, participants responded more quickly to OV-type orders, but when stimuli were synthesized words from spoken language, participants responded more quickly to VO-type orders. Importantly, this was true for speakers of both Italian (SVO) and Turkish (SOV). They concluded that non-lexical stimuli (e.g., gestures) engage the conceptual system, which prefers SOV, whereas words engage the computational system, which prefers SVO. They then suggested that the prevalence of both SOV and SVO in the world's languages results from the interaction of these two systems. Although they used gesture as a paradigmatic example of non-lexical communication, we know that some natural human languages (i.e., sign languages) are articulated in the manual modality. A remaining question, then, is how these systems would be engaged by communication in the manual modality that begins to take on language-like features, such as having a gestural lexicon and a communication partner who shares it. Langus and Nespor's study cannot answer this question, because the presence/absence of a lexicon is confounded with a difference between spoken and gestured modalities. Exploring this question is one aim of the present studies.

Langus and Nespor's (2010) account explains the extant data but assumes that these two systems are distinct. Langus and Nespor (2010, p. 291) explicitly state, "We rely on the proposal that the human faculty of language is modular and that it is possible to identify different cognitive systems responsible for specific linguistic tasks" (Chomsky, 2000; Fodor, 1983). According to this account, the SVO preference is specified as an inherent part of the human innate language faculty. This idea finds its strongest proponent in Kayne (1994), whose theoretical syntax approach analyzes all languages as underlyingly SVO. Absent from Kayne's proposal, however, is any consideration of why syntactic structure has this particular configuration rather than another one.

The goal of this study is to test whether a preference for SVO might be explicable in terms of cognitive-functional pressures that might bias certain constituent orders over others. That is, rather than positing a built-in preference for SVO, is it possible to identify specific factors that would cause a communication system to shift from SOV to SVO? The third study using elicited pantomime (Hall, Mayberry, & Ferreira, 2013) has taken a step in this direction. This study drew on three long-standing observations from linguistics. First, most SOV languages identify agents and patients overtly by using case marking (suffixes [or equivalent] that indicate a word's grammatical role), whereas this is less true of SVO languages (Greenberg, 1963). Second, overt object marking is especially common cross-linguistically when the object is a potential subject such as a human, known as *differential object marking* (Aissen, 2003; Bossong, 1991). Third, shifts from SOV to SVO are often triggered when a language loses its case marking system (Sapir, 1921; Sinnemäki, 2010; Vennemann, 1973). Taken together, these observations suggest that in the absence of case marking, SOV order may not be well suited for describing events in which both the agent and the patient are plausible agents (e.g., a boy kissing a woman). We refer to these as *reversible events*. Neither Goldin-Meadow, So et al. (2008) nor Langus and Nespor (2010) used reversible events in their experiments; almost all of their stimuli involved *non-reversible events*, where semantic role assignment was clear from the meaning of the gesture alone, regardless of constituent order (e.g., a boy kissing a box). To address this gap, Hall et al. (2013) asked native speakers of English (SVO) to describe both non-reversible and reversible events in pantomime. They found that although participants showed a robust preference for SOV for non-reversible events, they strongly avoided using SOV for reversible events. Instead, the most common solution that participants used was SVO, suggesting that perhaps SVO arises in the world's languages because it is well suited for describing both reversible and non-reversible events.

However, the participants in Hall et al. (2013) were all native speakers of English (SVO). In order for these results to be fully generalizable, we need to know whether SOV speakers also avoid using SOV to describe reversible events in pantomime. Preliminary support for this hypothesis comes from Meir, Lifshitz, Ilkbasaran, and Padden (2010), who report pilot data in which nine Turkish (SOV) speakers used SOV less often for reversible events (64%) than for non-reversible events (88%). These findings are echoed by Gibson et al.'s (2013) finding that bilingual speakers of Japanese (SOV) and English (SVO) also avoided using SOV for reversible events in pantomime. Although the available data indicate that SOV speakers are likely to avoid SOV for reversible events, it is not yet known whether they, like English speakers, would also favor SVO as a solution in a pantomime task.

In the experiments reported by Hall et al. (2012), participants produced a wide variety of non-SOV constituent orders to describe reversible events. Although SVO was the most common of these, it accounted for only 18–38% of reversible utterances across all experiments, leaving many other alternative orders besides SOV and SVO. Interestingly, many of these alternative orders are not commonly found in natural languages. In particular, these unusual orders tended to be inefficient or to mention the object before the subject (or both). These unusual orders are highly relevant to the issues explored here, and so we describe each type of order in turn.

All of the stimuli in Hall et al. (2013) contained one agent, one patient, and one action. Thus, a maximally efficient way to describe the events would be to use one of the six constituent orders consisting of three and only three elements (SVO, SOV, VSO, VOS, OVS, and OSV). Although participants were encouraged to include all three elements during practice trials, they were never explicitly instructed to produce three and only three gestures. This allows us to observe the natural variation in communicative efficiency under varying experimental conditions. In addition, some naturally emerging sign languages express single events with multiple utterances (Sandler et al., 2005; Senghas et al., 1997); if such patterns arose among our participants, we wanted to be able to capture them. Utterances that omit one or more elements (e.g., SV, VO, V, etc.) are underinformative when the to-be-described event is transitive. Those that repeat at least one element (e.g., SOSV, SOSOV, SOSVO, etc.) include redundant information. It is important to note that among the languages of the world, none are known to have a basic constituent order that is inefficient (either underinformative or repetitious). Although there is substantial variability in which constituent order a language will use most often, the order used is always one that is maximally efficient, under this definition. Accordingly, Slobin (1977) argues that the pressure for language to be efficient is one of the major forces that shapes linguistic structure. In this respect, it appears that there are pressures that constrain natural languages that did not similarly constrain elicited pantomime in previous experiments. Perhaps the inefficient orders we observed in pantomime would decrease under conditions that were closer to natural language, leaving more efficient orders like SVO and OSV.

Another common strategy that participants used for describing reversible events in pantomime was to put the object before the subject (e.g., OSV, OSOV, OSVO, etc.). Once again, this tendency is extremely uncommon as a dominant order in natural language; some early researchers claimed that true O-before-S languages were unattested (Parker, 1980; Vennemann, 1973), but more recent work reveals that such languages do exist (Derbyshire & Pullum, 1981; Pullum, 1981). Still, they are confined mainly to the Amazon basin and surrounding areas, and orders with O-before-S (VOS, OVS, OSV) are unquestionably the rarest of the six efficient constituent orders.

The observation that subjects tend to precede objects is one of the most robust observations in language typology; in fact, it is Greenberg's first universal (Greenberg, 1963). The strength of this regularity has inspired many different researchers to propose explanations for the mechanism behind it. Some offer arguments that emphasize grammar-internal motivations, such as Pullum's (1977) proposal that grammatical categories are preferentially realized according to their position in the Keenan and Comrie (1977) accessibility hierarchy. Other grammar-internal explanations for subjects preceding objects include forward thematization (Dayley, 1985; Tomlin, 1986) and the verb–object bonding principle (Tomlin, 1986). A different view posits that subjects tend to precede objects because of the way that human language grew out of evolutionarily prior forms of communication, which first coded goals (typically objects), then began to specify topics (typically agents) before goals, and later still began to also encode actions after mentioning both topic and goal (Givón, 1979). Still other researchers offer accounts based on psychological processing, such as the idea that mentioning subjects before objects reduces memory demands on a parser (Gibson, 1998), increases parsing efficiency (Hawkins, 2002, 2007; Newmeyer, 2002), or maintains more uniform information density (Maurits, 2012).

Discriminating among these various accounts is not the goal of the present study; rather, for our purposes, it suffices to note that the vast majority of the world's languages use a basic constituent order in which the subject precedes the object. Therefore, the fact that O-before-S orders commonly appear when participants describe reversible events in pantomime (Hall et al., 2013; Meir et al., 2010) again suggests that some of the pressures that constrain natural languages did not constrain elicited pantomime under the conditions tested in these experiments. This, in turn, leads to the prediction that O-before-S orders should become less common under conditions that more closely approximate natural language.

In the present experiments, we evaluate the hypothesis that SVO emerges in natural languages in part because SVO allows participants to avoid SOV for reversible events, while still being efficient and keeping the subject before the object. This hypothesis is not novel; however, no previous studies have demonstrated that these forces are latently present and at work in the minds of naïve participants, even over very short time scales (i.e., within a single testing session). In particular, we explore the possibility that being efficient and ordering the subject before the object might become increasingly relevant constraints under more language-like circumstances, compared to previous studies of elicited pantomime, which have been fairly unconstrained. Therefore, we aimed to create a controlled laboratory situation where participants would describe both non-reversible and reversible events in pantomime, but where some participants would be in more language-like situations than others. We hypothesized that under the demands of more language-like situations, pantomime might start to exhibit more of the characteristics of natural languages, such as a decrease in inefficient and O-before-S orders, and increases in SVO, especially for reversible events.

To manipulate the extent to which the pantomime situation was language-like, we considered the ways in which new natural human languages emerge in the visual-manual modality, that is, homesign and new sign languages. One of the first features to emerge in these systems is a lexicon, in which a consistent reciprocal mapping develops between a gesture and a basic category of referents. For example, a long-hair gesture that originated as a way to describe a particular woman with long hair might be extended to refer to the general category "woman," regardless of the length of a woman's hair. We, therefore, chose to simulate this stage of language emergence by asking some participants to create and use a consistent gestural lexicon for describing transitive events in pantomime.

Another difference between previous experiments with pantomime and natural languages is that natural languages (including homesign) arise in the presence of other people. In theorizing about the various forces that shape language structure, Singleton, Goldin-Meadow, and McNeil (1995, p. 307) suggest that,

> Having people to share your communication system with also may be a key to… reaching a truly arbitrary, efficient linguistic system… . It is evident that by entering into and sharing a communication system with a willing partner, these kinds of evolutionary pressures would be introduced. Consequently, what might begin as a rudimentary communication system (e.g., iconic gestures created in the absence of speech) would evolve into a system that more closely approximates a full-fledged natural language.

Therefore, in the present study, the participants who received instruction to be consistent in the form of their gestures were further divided into two groups: shared and private.

In the shared condition, participants taught their gestures to the experimenter, who confirmed understanding of the form-meaning mapping of each gesture in isolation but did not otherwise use the gestures for any communicative purpose. We contrast the shared condition with the *baseline condition*, in which (as in previous experiments) participants were simply instructed to describe events in pantomime with no requirement to be consistent in the form of their gestures.

In the private condition, participants were instructed to create and use a consistent gestural lexicon but did not share it with the experimenter, who left the room during testing. We predicted the private condition to be intermediate between the baseline and shared conditions in terms of its language-likeness.

To summarize, this study builds on previous accounts of why SOV and SVO are both prevalent among the world's languages, with SOV common in a language's earlier stages while diachronic change favors SVO. To explain the prevalence of SOV, Goldin-Meadow, So et al. (2008) suggest that the way we organize our mental representations of event knowledge is especially compatible with SOV order. To explain the prevalence of SVO, Langus and Nespor (2010) suggest that, in addition to such conceptual representations, the human mind also houses a separate computational (linguistic) system, which prefers SVO. However, no explanation is offered for why that might be the case beyond an appeal to ideas about built-in preferences that are innately specified in the human language faculty (e.g., Chomsky, 2000). The present study aims to offer an explanation grounded in cognitive constraints on processing. Specifically, we note that there is strong evidence that participants avoid using SOV to describe reversible events, but that in previous studies, many people opted for alternative orders that are dispreferred in natural language, including those that are inefficient, and those that put object before subject. We hypothesize that, under circumstances that are a closer approximation of natural language situations, we will observe a decrease in constituent orders that are inefficient and that put object before subject. The interaction of these pressures would leave SVO as a preferred solution for describing reversible events, potentially explaining its prevalence both synchronically and diachronically. Therefore, to make the elicited pantomime task slightly more language-like, we instructed some participants to create and use a consistent gestural lexicon, and we contrast the constituent orders produced by participants in the consistent-gesture conditions with those produced by participants in a baseline condition, where no instruction was given about being consistent in the form of their gestures.

2. Experiment 1

2.1. Method

2.1.1. Participants

We tested 36 undergraduates at UC San Diego who reported being native monolingual speakers of English. All participants gave consent to participate and be video recorded as part of the study, and they received course credit for their participation.

2.1.2. Materials

Stimuli were 61 video clips of a human agent (man, woman, boy, girl) performing a transitive action (lift, kiss, pet, push) on a patient that was either non-human (ball, box, bike, car, cat, dog) or human (man, woman, boy, girl). Events with non-human patients are termed "non-reversible."[1] Events with human patients are termed "reversible"; in these cases, the agent and patient always differed in age, sex, or both. A full list of stimuli is given in the Appendix; these same stimuli were used in Hall et al. (2013). Stimuli were presented on a Macintosh laptop using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993).

2.1.3. Design and procedure

Participants were randomly assigned to one of three groups (baseline, private, or shared), yielding 12 participants per condition. In all cases, instructions were immediately followed by a brief exposure phase, where participants saw a preview of the kinds of things about which they would need to gesture. The exposure phase began with video sequences showing diverse instances of each action (e.g., several clips of lifting, several kissing, etc.). The videos were constructed so as to emphasize the common action shared by each short segment. After briefly exposing the participants to each of the four action elements, the nominal elements were presented as still images, one category at a time. For example, the first screen displayed still images of the various different balls that would appear in the test clips. The next screen displayed the various boxes, followed by bikes, cars, cats, dogs, and finally men, women, boys, and girls, in that order. This order of exposure (essentially VOS) was chosen so that it would not bias either SOV or SVO.

The critical manipulation concerned the instructions that were given during the exposure phase. Participants in the baseline condition were told that they would see short video clips that they would later describe in gesture, and that the exposure phase was just a preview of the kinds of things they might see. For them, there was no task to perform during the exposure phase. Participants in the private group were given the same basic instructions but were also asked to use the exposure phase to think of a gesture that they would then use consistently throughout the session to refer to each action and each nominal element. However, they were not asked to produce those gestures overtly until the beginning of practice. The goal was to help the participants create a gestural lexicon that they would use consistently but that would not be shared with the experimenter, who left the room after the practice trials in this condition. Participants in the shared group received the same instructions as in the private group, except that they were asked to overtly teach their gestures to the experimenter, who both confirmed understanding, and remained visible in the room throughout the testing session.

Following the exposure phase, participants performed practice trials during which the experimenter encouraged them to include gestures for all three components of each event (agent, patient, action) but carefully avoided using language that would bias participants toward any single order. After the practice trials, participants were given 61 test trials. To facilitate comparison with previous studies that have used only non-reversible events (Goldin-Meadow, So et al., 2008; Langus & Nespor, 2010), participants began with 45 non-reversible events, followed by 16 reversible events. Hall et al. (2013) found no differences in constituent order when reversible trials came at the end of the testing session compared to when they came at the beginning or were mixed throughout. Within these blocks of non-reversible and reversible trials, we used a fixed random order with the constraint that consecutive trials shared no more than one element.

2.1.4. Coding and analysis

Subjects' gestures were transcribed by two trained coders, who identified each gesture's referent as subject, verb, object, or other. Multiple *consecutive* gestures for the same referent were considered as belonging to the same constituent. For example, the string "TALL LONG-HAIR WHEELS PEDALS HANDLEBARS PUSH" would be parsed as SOV (woman, bicycle, push). However, if the string were non-consecutive, a repetition would be noted. For example, "TALL LONG-HAIR WHEELS PEDALS PUSH HANDLEBARS" would be coded as SOVO (woman bicycle push bicycle). In the absence of any principled way to determine phrase or utterance boundaries in pantomime, everything that the participant produced was counted as a single utterance unless there was a pause of more than 2 s or the string was interrupted by the participant replaying the video. In these cases, the most complete string containing an action gesture was coded. Gestures classified as "other" were ignored in determining constituent order; these mainly contained information about the environmental surroundings of the scene. If it was not possible to determine the referent of a gesture, the trial was considered "ambiguous." If participants produced gestures that referred to more than one constituent at the same time, such that it was not possible to compute order of mention, it was considered "simultaneous." Trials coded as ambiguous or simultaneous were excluded from statistical analysis.

Coders were not blind to the experimental manipulations. However, they were blind to one another's ratings and agreed on 2,018 of the 2,196 utterances (91.9%). In cases of disagreement, a third rater coded the trial blind to the other raters' responses. If this resulted in two of three coders agreeing, the trial was included. If all three coders disagreed, the trial was excluded from statistical analysis (45 trials = 2%).

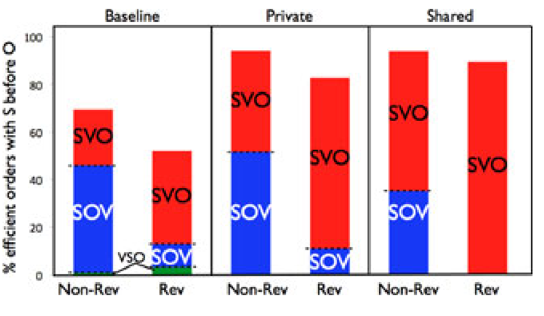

Fig. 1. Distribution of constituent orders in Experiment 1 (English speakers) that are both efficient and mention subject before object. Other constituent orders (not shown here) are given in Table 1.

Table 1
Distribution of constituent orders produced by participants in each condition in Experiment 1 (English speakers)

	Baseline		Private		Shared	
Type	Non-Rev	Rev	Non-Rev	Rev	Non-Rev	Rev
A-SOV	45	10	52	11	35	0
A-SVO	23	39	42	72	59	89.5
A-VSO	1	3	0	0	0	0
B	11	10	3	0	6	0
C	9	7	2	12.5	0	7
D	1.5	9	0	0.5	0	1.5
E	4	5	0.5	0	0	0
F	5.5	17	0.5	4	0	2
Total (%)	100	100	100	100	100	100

Note. Type A includes efficient orders with subject before object (SOV, SVO, VSO). Type B includes efficient orders with object before subject (OSV, OVS, VOS). Type C includes repetitious orders with subject before object (e.g., SOSV, SOSVO, etc.). Type D includes repetitious orders with object before subject (e.g., OSOV, OSVO, etc.). Type E includes underinformative orders (e.g., V, OV, etc.). Type F includes orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.

2.2. Results

2.2.1. Prevalence of SOV

Fig. 1 shows the relative prevalence of efficient orders with subject before object in each condition. The distribution of all orders is given in Table 1.

Only trials with SOV order were coded as consistent with SOV; all other trials were coded as inconsistent. We analyzed whether our manipulations of semantic reversibility (within subjects) or gestural consistency (between subjects) influenced the prevalence of SOV at both the group level and at the level of individual participants.

2.2.1.1. Group results: For each participant, we computed the proportion of SOV trials and submitted these proportions to a 2 × 3 mixed ANOVA with reversibility (non-reversible, reversible) as a within-subjects factor and group (baseline, private, shared) as a between-subjects factor. We found that group did not influence the prevalence of SOV [$F(2,33) = .79$, $p = .46$], but that participants used significantly less SOV for reversible trials [$F(1,33) = 29.66$, $p < .001$]. These effects did not interact [$F(2,33) = .07$, $p = .93$].

2.2.1.2. Individual results: It was not the case that every participant resembled the group average; rather, individual participants tended to have a particular order that they used more often than any other within each level of reversibility. We defined this as the participant's dominant order, and used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SOV-dominant. In the baseline group, 7/12 participants were SOV-dominant for non-reversible events but only 1/12 was SOV-dominant for reversible events ($p < .02$). In the private group, 6/12 participants were SOV-dominant for non-reversible events but only 1/12 was SOV-dominant for reversible events ($p < .04$). In the shared group, 4/12 participants were SOV-dominant for non-reversible events, but 0/12 were SOV-dominant for reversible events ($p < .05$).

We also used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SOV-dominant by comparing baseline to private and baseline to shared, at each level of reversibility. For non-reversible events, the number of SOV-dominant participants in the baseline group (7/12) was not significantly different from the number of SOV-dominant participants in the private (6/12, $p = .5$) or shared (4/12, $p = .2$) groups. The same was true for the reversible events: the number of SOV-dominant participants in the baseline group (1/12) was not significantly different from the number of SOV-dominant participants in the private (1/12) or shared (0/12) groups.

2.2.2. *Prevalence of SVO*

Following the above logic, only trials with SVO order were coded as consistent with SVO; all other orders were coded as inconsistent. Unless otherwise noted, the statistical methods below are identical to those mentioned above. It is important to note that these analyses of SVO are not strictly independent of the preceding analyses of SOV, because a trial cannot be both SOV and SVO. However, because the hypotheses that we aim to test make different predictions about the circumstances under which SOV and SVO should increase and decrease, it is important to examine both dependent measures. We note (a) that when the same manipulation affects SOV rates and SVO rates in opposite directions, this should only be taken as a single observation in support of the effect of that manipulation (because the non-independence could drive the second effect) and (b) that we do not compare rates of SOV and SVO directly, mitigating some concerns regarding their non-independence.

2.2.2.1. Group results: The 2 × 3 ANOVA revealed that SVO was more prevalent in some groups than others [$F(2,33) = 3.90$, $p < .04$]. Planned contrasts found that SVO was more common in the shared group than in the baseline group [$F(1,33) = 7.68$, $p < .01$]. For the private group, the increase in SVO was not significantly different from baseline [$F(1,33) = 2.80$, $p = .10$]. SVO was also significantly more common for reversible trials than for non-reversible trials [$F(1,33) = 17.11$, $p < .001$], in complement to the results with SOV rates. There was no group × reversibility interaction [$F(2,33) = .62$, $p = .54$].

2.2.2.2. Individual results: We again used Fisher's exact test to determine whether reversibility influenced the probability that individual participants in each condition would be SVO-dominant. Although more participants in each group were SVO-dominant for reversible than for non-reversible events, these differences did not reach significance (baseline: 3/12 vs. 5/12, $p = .24$; private: 6/12 vs. 9/12, $p = .16$; shared: 7/12 vs. 11/12, $p = .07$).

Finally, we used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SVO-dominant by comparing baseline to private and baseline to shared, for each level of reversibility. For non-reversible events, the number of SVO-dominant participants in the baseline condition (3/12) was not significantly different from the number in private (6/12, $p = .16$) or shared (7/12, $p = .09$). For reversible events, the increase in the number of SVO-dominant participants did not reach statistical significance from baseline (5/12) to private (9/12, $p = .09$) but did reach significance from baseline (5/12) to shared (11/12, $p < .02$).

2.3. Discussion

Consistent with previous research (Gibson et al., 2013; Hall et al., 2013; Meir et al., 2010), participants avoided using SOV to describe reversible events. This pattern was robust at both the group and individual levels. However, we did not find evidence that the instruction to create and use a consistent gestural lexicon influenced how often participants used SOV descriptions (Although this latter result is a null effect, it stands in contrast to the effect that the same instruction had on the prevalence of SVO, discussed below.) Such findings are to be expected if indeed these two manipulations interact with different sets of cognitive pressures. If the instruction to use a consistent gestural lexicon incentivizes efficiency (i.e., including S, O, and V exactly once each) and information structure (i.e., keeping S before O), then SOV is a perfectly adequate order, and so is not predicted to decrease across groups. SOV becomes problematic only when the to-be-described events are reversible, at which point participants should avoid it.

We did, however, find that the prevalence of SVO was influenced by both reversibility and by the differential instructions given to each group. As in previous studies (Gibson et al., 2013; Hall et al., 2013), participants used more SVO for reversible events than for non-reversible events. This is consistent with the idea that using SVO is one way to avoid role conflict, minimize confusability, or both. A novel finding of the present study is that SVO became significantly more prevalent in the most language-like condition, that is, when participants were instructed to create and use a consistent gestural lexicon, and where that lexicon was also understood by a passive interlocutor. When participants had a gestural lexicon but no interlocutor, the prevalence of SVO was intermediate and not significantly different from either the baseline or shared conditions. Thus, we cannot yet dissociate the impact of the lexicon from that of the interlocutor. For reversible events, this effect is a straightforward consequence of the interaction of three cognitive pressures: If SOV is not a good option for describing reversible events (because of role conflict, confusability, or both), and if it is important to maximize efficiency and to keep the subject before the object, then SVO is the only order that satisfies those three constraints.

One unexpected finding, however, was that the instruction to create and use a consistent gestural lexicon increased SVO not only for reversible events but also for non-reversible events. Because SVO is also an efficient order with S before O, it should be preferred to orders like SOSOV, OSV, and VOS, which all occurred more in the baseline group than in the private and shared groups (see Table 1). The unexpected aspect of this finding was that SOV should have been just as good a solution on those grounds, and so we might have expected to see both SOV and SVO increase, but only SVO became more frequent across groups. There are three possible explanations for this finding. One is that, as a system becomes more language-like, it engages the computational system of syntax, predicted by Langus and Nespor (2010) to yield more SVO. Their account does not distinguish between reversible and non-reversible events, and so would predict an increase in SVO for both types of events, as we observed. From this perspective, the novel insight would be that this effect can be obtained even in pantomimic gesture. However, a second possibility is that some or potentially all of the increase in SVO across groups could come from another source: the participants' native language. It may be that the process of creating and using a gestural lexicon encourages participants to silently recode their gestures into words in their native language. That, in turn, could then bias the order in which participants gesture to more closely reflect the order of their native language: in this case, SVO. The third possibility is that both factors are involved to some extent.

Therefore, the data from Experiment 1 cannot determine the extent to which the increase in SVO across groups reflects a potentially universal cognitive pressure, a language-specific preference for SVO, or a combination of both. To explore this question in further detail, we replicated Experiment 1 with native speakers of Turkish, which uses SOV structure. Our hypothesis predicts that SVO should still emerge in reversible events when participants are instructed to create and use a gestural lexicon. If so, it cannot be attributed to influence from participants' native language, which would instead work against this finding. However, we might also find that SVO increases in both reversible and non-reversible events, which would support Langus and Nespor's hypothesis that SVO is a preferred order for language-like systems, but broaden the scope of that view to include non-linguistic gesture as well. Alternatively, we might find no evidence of SVO in Turkish speakers, which would suggest that the results of Experiment 1 were likely due to influence from English.

3. Experiment 2

3.1. Method

3.1.1. Participants

All testing was conducted in Turkey by a native Turkish speaker, mainly in Sariyer and Istanbul. Our goal was to find monolingual Turkish speakers who were relatively young and familiar with computers. Most people in this demographic have had some exposure to English during school but vary widely in their actual proficiency. Due to the practical realities of recruitment in Turkey, we needed a simple and quick measure, and chose to use a 0–5 self-report scale. Then, because different people might have different interpretations about what a "3" meant, we added the descriptions, reported in Table 2, as anchors. An ideal participant would have no contact with or knowledge of any SVO language and would therefore report a "0." Potential participants were excluded if an SVO language was spoken in their home. All but one of the participants were raised in a home where only Turkish was spoken; the one exception had one parent who spoke

Table 2
Language screening questions

5 = I am comfortable using this language in all situations
4 = I am mostly comfortable using this language, but I prefer another one in some situations
3 = I can use this language when necessary, but prefer not to
2 = I can usually make myself understood, but it is difficult
1 = I would only use this language if I had no other options
0 = I don't remember any of this language

Note. Potential participants rated themselves on every language they knew, using the scale above. Those reporting 3 or higher in an SVO language were excluded from participation.

Arabic (VSO) at home. (Two participants reported having one parent who was fluent in an SVO language [Albanian] but did not indicate that it was spoken in their home.) Roughly two-thirds of potential participants reported having some contact with English or another SVO language in school. Potential participants were excluded if they reported "3" or above in any SVO language. This left 33 participants, of whom 9 reported "0," 19 reported "1," and 5 reported "2." All participants gave consent to be videotaped as part of the study and were paid for their participation.

3.1.2. Materials

We used the same materials as in Experiment 1.

3.1.3. Design and procedure

The design and procedure were identical to Experiment 1, except that written and spoken instructions were delivered in Turkish.

3.1.4. Coding and analysis

Coding procedures were identical to Experiment 1. The first two coders agreed on 1,915/2,013 utterances (95.1%). After the third coder, only 27 trials (1.3% of the data) were excluded. Unless otherwise noted, the statistical methods were identical to those in Experiment 1.

3.2. Results

3.2.1. Prevalence of SOV

Fig. 2 shows the relative prevalence of efficient orders with subject before object in each condition. The distribution of all orders is given in Table 3. As in Experiment 1, the proportion of trials that had SOV order was analyzed at both the group and individual level.

3.2.1.1. Group results: The 2 × 3 ANOVA revealed a trend for SOV to be more common in some groups than others [$F(2,30) = 2.84$, $p = .07$]. Planned comparisons found that SOV was more common in the private group than in the baseline group [$F(1.30) = 4.49$,

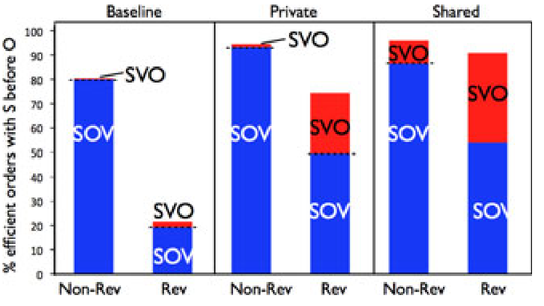

Fig. 2. Distribution of constituent orders in Experiment 2 (Turkish speakers) that are both efficient and mention subject before object. Other constituent orders (not shown here) are given in Table 3.

Table 3
Distribution of constituent orders produced by participants in each condition in Experiment 2 (Turkish speakers)

	Baseline		Private		Shared	
Type	Non-Rev	Rev	Non-Rev	Rev	Non-Rev	Rev
A-SOV	80	19.5	93.5	49	86.5	54
A-SVO	0.5	2	1	25	9.5	37
A-VSO	0	0	0	0	0	0
B	0.5	20.5	0.5	6	0.5	6.5
C	12.5	34.5	2	11	1	0.5
D	0.2	4	0.5	1	0	1
E	5.5	0	1.5	2	1.5	0.5
F	0.8	19.5	1	6	1	0.5
Total (%)	100	100	100	100	100	100

Note. Type A includes efficient orders with subject before object (SOV, SVO, VSO). Type B includes efficient orders with object before subject (OSV, OVS, VOS). Type C includes repetitious orders with subject before object (e.g., SOSV, SOSVO, etc.). Type D includes repetitious orders with object before subject (e.g., OSOV, OSVO, etc.). Type E includes underinformative orders (e.g., V, OV, etc.). Type F includes orders that were coded as ambiguous, simultaneous, or excluded due to lack of consensus.

$p < .05$], and that SOV was marginally more common in the shared group than in the baseline group [$F(1,30) = 4.02$, $p = .05$]. SOV was significantly less common on reversible events than on non-reversible events [$F(1,30) = 47.02$, $p < .001$]. There was no interaction between group and reversibility [$F(2,30) = 1.53$, $p = .23$].

3.2.1.2. Individual results: At the individual level, we used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SOV-dominant. In the baseline group, 10/11 participants were SOV-dominant for non-reversibles, whereas 0/11 were SOV-dominant for reversibles ($p < .001$). In the private group, 11/11 participants were SOV-dominant for non-reversibles, whereas 4/11 were SOV-dominant for reversibles ($p < .01$). In the shared group, 10/11 participants were SOV-dominant for non-reversibles, and 6/11 were SOV-dominant for reversibles (p = .07).

We also used Fisher's exact test to determine whether the between-subjects manipulation of gestural consistency influenced the probability of participants being SOV-dominant by comparing baseline to private and baseline to shared, for each level of reversibility. For non-reversible events, the number of SOV-dominant participants in the baseline group (10/11) was not significantly different from the number of SOV-dominant participants in the private (11/11) or shared (10/11) groups. For reversible events, the number of SOV-dominant participants in the baseline group (0/11) was significantly different from the number of SOV-dominant participants in the private group (4/11, $p < .05$) and the shared group (6/11, $p < .01$).

3.2.2. Prevalence of SVO

Following the above logic, the proportion of trials that had SVO order was analyzed at both the group and individual level. Group-level data are displayed in Fig. 2.

3.2.2.1. Group results: The 2 × 3 ANOVA revealed a trend for SVO to be more common in some groups than others, although the main effect of group did not reach significance [$F(2,33) = 2.54$, $p = .10$]. Planned comparisons found that the prevalence of SVO was not significantly different between the baseline and private groups [$F(1,30) = 1.47$, $p = .24$], but that SVO was significantly more common in the shared group than in the baseline group [$F(1,30) = 5.08$, $p < .04$]. SVO was also significantly more common in reversible events than in non-reversible events [$F(1,30) = 7.64$, $p < .01$], complementing the results with SOV orders. There was no interaction between group and reversibility [$F(2,30) = 1.56$, $p = .23$].

3.2.2.2. Individual results: At the individual level, we used Fisher's exact test to determine whether the reversibility manipulation influenced the probability of participants being SVO-dominant. In the baseline group, 0/11 participants were SVO-dominant for non-reversible events, and 0/11 were SVO-dominant for reversible events (p = 1). In the private group, 0/11 were SVO-dominant for non-reversibles, and 3/11 were SVO-dominant for reversibles (p = .11). In the shared group, 1/11 was SVO-dominant for non-reversibles, and 4/11 were SVO-dominant for reversibles (p = .14).

Likewise, we also used Fisher's exact test to determine whether the manipulation of group influenced the probability of participants being SVO-dominant by comparing baseline to private and baseline to shared, for each level of reversibility. For non-reversible events, there were no differences between the number of SVO-dominant participants in the baseline (0/11), private (0/11), and shared (1/11) groups (all $p > .5$). For reversible events, the difference between baseline (0/11) and private (3/11) did not reach significance (p = .11), but there was a significant difference between the number of SVO-dominant participants in the baseline (0/11) and shared (4/11) groups ($p < .05$).

3.2.3. Combined analysis of Experiments 1 and 2 (group level)

Finally, we combined the results from both experiments using a 2 × 2 × 3 ANOVA with reversibility (no, yes) as a within-subjects factor, and native language (English, Turkish) and group (baseline, private, shared) as between-subjects factors.

With percent SOV as the dependent measure, there was a main effect of native language [$F(1,63) = 37.29$, $p < .001$], indicating that Turkish speakers used more SOV than English speakers overall. There was also a main effect of reversibility [$F(1,63) = 75.07$, $p < .001$], indicating that SOV was less common for reversible events overall. Importantly, native language did not interact with reversibility [$F(1,63) = .82$, $p = .37$] or group [$F(2,63) = 2.01$, $p = .14$]. No other effects were significant (all Fs < 2, all ps > .25).

With percent SVO as the dependent measure, there was a main effect of native language [$F(1,63) = 29.77$, $p < .001$], indicating that English speakers used more SVO than Turkish speakers overall. There was also a main effect of reversibility [$F(1,63) = 23.59$, $p < .001$], indicating that SVO was more common for reversible events overall. Here, the main effect of group was significant [$F(2,63) = 6.07$, $p < .01$]. Planned comparisons revealed that SVO was more common in the shared group than in the baseline group [$F(1,63) = 12.04$, $p < .001$]. SVO was also more common in the private group than in the baseline group [$F(1,63) = 4.07$, $p < .05$]. Importantly, native language did not interact with group or reversibility (all Fs < 1), and no other effects were significant (all Fs < 2.1, all ps > .13).

3.3. Discussion

The data from Turkish speakers demonstrate that SVO begins to emerge for reversible events in the shared group and, to a lesser extent, for reversible events in the private group as well. Importantly, participants in the shared group were significantly more likely to use SVO to describe reversible events than participants in the baseline group. Participants in all groups avoided SOV, but many of the alternative orders employed by participants in the baseline group tended to put O before S (Table 3, Type B) or involved repetition (Table 3, Types C and D). Those tendencies decreased in the private and shared groups, with both SOV and SVO increasing instead. However, whereas the increase in SOV is potentially attributable to influence from the participants' native language, the increase in SVO is not. Instead, we propose that it emerges because it uniquely satisfies the constraints against using SOV for reversible events while still being efficient and keeping S before O.

As with English speakers, we did not find evidence that the instruction to create and use a consistent gestural lexicon led to reduced SOV. This is again consistent with the notion that SOV is an efficient order that keeps S before O. However, in contrast to English speakers, more participants in the private and shared groups were SOV-dominant for reversible events than in the baseline group. This pattern suggests that the instruction to create and use a consistent gestural lexicon may indeed have encouraged some participants to use verbal recoding, which in turn led to increased use of native-language order in the private and shared groups than the baseline group. (The lack of a corresponding increase in SOV among non-reversible events may be due to a ceiling effect.) Therefore, it seems likely that at least some of the increase in SVO that we observed in Experiment 1 might be attributable to influence from the participants' native language, rather than reflecting potentially universal principles of cognition. However, the crucial question of Experiment 2 is whether we have any evidence that SVO emerges as a response to our manipulations when it cannot be attributed to influence from the participants' native language. As we have noted above, SVO does emerge when Turkish speakers describe reversible events with a self-generated gestural lexicon, an effect that cannot be attributed to the speakers' native-language word order.

One final aspect of the present data deserves comment. We found that native Turkish speakers avoided using SOV descriptions for reversible events, which replicates a pattern described by Hall et al. (2013). The present observation is especially noteworthy because SOV is the characteristic order of Turkish participants' native language for both reversible and non-reversible events. Therefore, the pressure that drove these participants to avoid SOV must have been strong enough to outweigh the natural tendency to describe events by using the structure of one's native language. Similar findings in SOV speakers have also been observed by Gibson et al. (2013), who tested Japanese-English and Korean-English bilinguals, and by Meir et al. (2010), who reported preliminary data from nine Turkish monolinguals.

4. General discussion

The experiments presented here show two main points. First, we demonstrated that even native speakers of an SOV language (Turkish) avoid using SOV to describe reversible events in pantomime. This is consistent with earlier results from English speakers (Gibson et al., 2013; Hall et al., 2013), as well as preliminary data from nine Turkish monolinguals (Meir et al., 2010) and from Japanese-English bilinguals (Gibson et al., 2013). Despite giving contrasting explanations for why people avoid SOV for reversible events, these authors all agree that there is some functional motivation behind this behavior and suggest that whatever the cause might be, the same functional motivation likely also applies to natural language.

Second, the present experiments show that SVO may arise in part because it is an efficient way to describe reversible events while still keeping subjects before objects. In previous studies, participants often used constituent orders that were inefficient (either underinformative or repetitious) or placed objects before subjects; this happened especially often for reversible events. We hypothesized that these inefficient and O-before-S orders were relatively common primarily due to the absence of other pressures that act on natural language. To test this hypothesis, we manipulated two aspects of the pantomime task. First, since a lexicon is one of the earliest language structures to emerge in new languages, we instructed some participants to create and use a gestural lexicon. Second, because natural languages arise in the context of human relationships, we instructed half of these participants to teach their gestures to the experimenter (the shared condition), while the other half performed the task alone (the private condition). We compared the constituent orders produced by the participants in each of these conditions against those produced by participants in the baseline condition, who received no special instructions (as in previous experiments). We found that both English and Turkish speakers were more likely to use SVO to describe reversible events in the shared condition than in the baseline condition, which was again largely characterized by constituent orders that were either inefficient (Types C and E of Tables 1 and 3) or mentioned objects before subjects (Type B of Tables 1 and 3), or both (Type D). Although the results from the English speakers could reflect covert influence of L1, the same cannot be true of Turkish speakers. Therefore, we take these results as evidence that at least part of the reason that SVO emerges in the world's languages is because it allows language users to satisfy the three constraints, being efficient, keeping subjects before objects, and avoiding SOV for reversible events.

The present results raise the question of whether the emergence of SVO could be attributable to the presence of an interlocutor, rather than to the use of a consistent gestural lexicon. This question arises because the frequency of SVO in the private condition (consistent lexicon, no interlocutor) did not differ significantly from baseline except in the combined analysis. In contrast, SVO was significantly more frequent than baseline whenever an interlocutor was present. Three interpretations of this pattern are possible. First, it could be that the emergence of SVO is entirely due to gestural consistency, but that having an interlocutor present increases the extent to which people are consistent in the form of their gestures. Although it was not possible to quantify this pattern, coders did anecdotally report that participants in the private condition seemed qualitatively less consistent in the form of their gestures than participants in the shared condition. A second possibility is that the emergence of SVO is due entirely to the presence of an interlocutor, with gestural consistency having no impact. We did not include a condition in which an interlocutor was present but participants were not asked to create and use a consistent gestural lexicon; therefore, we cannot yet rule out this explanation. The third possibility, of course, is that both gestural consistency and the presence of an interlocutor contribute to the emergence of SVO. These three possibilities could be discriminated with a $2 \times 2 \times 2$ experiment that factorially manipulates gesture consistency (baseline/consistent), the presence of an interlocutor (present/absent), and native-language word order (SVO/SOV). For now, however, we can safely conclude that SVO emerged most strongly when both gestural consistency and a passive interlocutor were involved, although we cannot delineate the independent contribution of each factor.

We return, then, to Langus and Nespor's (2010) proposal that the constituent order distribution in the world's languages stems from two systems: a conceptual system that prefers SOV, and a computational system that prefers SVO. Under a strict interpretation of this account, our data suggest that it is incorrect, or at least incomplete. First, there is now strong evidence that elicited pantomime does not always prefer SOV: for reversible events, participants clearly avoid SOV. Therefore, insofar as pantomime engages the conceptual system alone, this system is sensitive to more than simply the order of agents, patients, and actions. Second, the present study has also established that it is possible to account for the emergence of SVO even in the absence of the posited computational (syntactic) system. We agree with Langus and Nespor that participants' behavior in elicited pantomime tasks is not being governed by bona fide syntactic processing, and yet we have demonstrated that merely instructing participants to be consistent in the form of their gestures and to share them with the experimenter was sufficient to increase the frequency of participants using SVO to describe reversible events. Because participants were never exposed to other people's pantomimes, this result cannot be due to a process like creolization, which is known to give rise to SVO. Because we observed an increase in SVO even among Turkish (SOV) speakers, the effect likewise cannot not be attributed merely to covert influence from the participants' native language. Thus, it would seem that at least some aspects of a strict interpretation of Langus and Nespor's model require modification.

Relaxing some of the assumptions of Langus and Nespor (2010) allows the model to capture the data with minimal modification. In particular, the model is largely successful if the assumption that the two systems are strictly segregated is dropped, and that a push for SVO is triggered by exposure to linguistic input during infancy. In that case, the model would essentially propose that, over time, the various constraints that languages face tend to be best satisfied by SVO order. Here we have considered only three potential constraints, but there are likely many others as well. For example, there may also exist a cognitive preference for mentioning the subject before the verb, as suggested by Givón (1979). If so, that could explain why our data contain so few instances of VSO, which is the only other efficient non-SOV order that keeps subjects before objects. And indeed, perhaps it is not a coincidence that VSO is the third most common order across spoken languages.

There may be additional reasons to avoid orders with both nominal arguments on the same side of the verb. For example, Gibson et al. (2013) propose that languages evolve away from SOV toward SVO in part because orders with subject and object on opposite sides of the verb are more resistant to information loss during communication. Although there are some drawbacks to this particular account (see Hall et al., 2013), it nevertheless illustrates the principle that language structure is likely to be influenced by functional pressures. Here we have demonstrated the impact of some basic pressures that are likely to be active early on in the process of a system becoming organized into language (evolving a lexicon and having communicative partners).

Given the above evidence that constituent order is sensitive to these pressures in a laboratory context, it is important to ask whether similar patterns are attested in other situations of language creation and evolution over varying time scales. We consider this question in four contexts: (a) how homesign systems change as children grow up; (b) how constituent order becomes organized in emerging sign languages across a few generations; (c) the evolution of pidgins into creoles; and (d) how established languages change over long time scales.

Homesign refers to idiosyncratic systems (or sometimes family-lects) invented by deaf children who cannot hear speech and have no exposure to a conventionalized sign language. The constituent order of homesign systems has been reported to be robustly SV and OV, across different children and across multiple cultures (Goldin-Meadow & Mylander, 1998; Goldin-Meadow, Özyürek, et al., 2008). However, this research does not typically distinguish between reversible and non-reversible events, making it difficult to know whether their behavior is consistent with the patterns observed in elicited pantomime. Moreover, much of the linguistic research on homesign focuses on young children, who typically refer to people and things by pointing (Goldin-Meadow & Feldman, 1977). These pointing gestures are very different from the more embodied kinds of gestures that our adult participants typically produced and may only represent an initial stage of homesign development. Coppola (2002) followed three adult homesigners for several years and found that their systems often involved more embodied representations of people and enacted representations of action, just as we observed in elicited pantomime. At one point, she asked these adult homesigners to describe both reversible and non-reversible events and found that, in all three cases, semantic reversibility led to changes in constituent order. Two of the homesigners primarily SOV for non-reversible events but avoided it for reversible events, with one preferring OSV and the other preferring SVO. This pattern is strikingly similar to what we observed in elicited pantomime. The third homesigner showed the opposite pattern, preferring SVO for non-reversible events and SOV for reversible events. However, unlike the other two, his use of constituent order varied over the years, making it difficult to draw conclusions from any one sample of his homesign. At the very least, these data strongly suggest that adult homesigners do use different constituent orders to describe reversible and non-reversible events.

These observations lead naturally to the question of whether natural spoken languages exhibit differential word orders for reversible and non-reversible events. Such patterns are often overlooked by formal analyses of language grammars, which tend to focus on what is possible in a language, rather than what is most common. However, both typological and psycholinguistic research finds that animacy can influence constituent order in natural spoken languages. For example, many languages in which word order is typically flexible (e.g., Russian, Japanese, Korean, Hindi) have fewer permissible options when both agent and patient are human: a phenomenon known as word-order freezing (for a review, see Lee, 2001). In addition, animacy still affects constituent order choice even when a grammar permits multiple options (Branigan, Pickering, & Tanaka, 2008). It is therefore, possible that as a language undergoes a change in constituent order, it passes through a phase where speakers tend to use SVO to describe reversible events while still preferring SOV for non-reversible events. We presently lack direct evidence for this process; however, phenomena such as these have received little study, relative to the large body of work classifying languages according to their dominant order. We hope that future work in language description will give greater consideration to the potential impact of animacy on constituent order, thereby providing evidence that either supports or refutes the present hypothesis.

Next, we consider how constituent order changes when new sign languages emerge. Senghas et al. (1997) studied the constituent orders used by the first- and second-cohort signers of Nicaraguan Sign Language, and again found that reversibility strongly influenced constituent order. To describe non-reversible events, the signers used a mixture of (S)OV orders along with some OSV. For reversible events, however, Senghas et al. did not observe even a single instance of SOV among either the first- or second-cohort signers. Sandler et al. (2005) report that the nascent ABSL community has quickly converged on SOV as a basic word order, but they do not report whether reversibility impacted word order. The above findings lead to the prediction that ABSL signers would avoid SOV for reversible events, but at present it remains an empirical question. Several recent investigations of unrelated sign languages have found that SOV is more common for non-reversible events than for reversible events, while SVO shows the opposite pattern. These include Turkish Sign Language (Sevinç, 2006), Russian Sign Language (Kimmelman, 2012), and Kenyan Sign Language (Morgan, 2012).

The preceding evidence makes it clear that the pressure to avoid SOV for reversible events is attested in naturalistic contexts. But these systems also have consistent lexicons and interlocutors. Although some of these new languages have indeed begun to use SVO to describe reversible events, we might wonder why this shift is not more widespread, both within and across languages. There are two factors that are likely responsible for this: the availability of other devices for marking semantic relations, and the interaction of competing constraints from production, comprehension, and acquisition. The potential impact of these factors can be illustrated by considering the evolution of pidgins into creoles, and the way that spoken languages change across long time scales.

One of the most robust patterns in language evolution is that when a creole emerges out of a pidgin, it will be SVO (Bakker, 2008; McWhorter, 2001). This pattern holds even if the pidgin's source languages were all SOV (Kouwenberg, 1992). The natural question is why these systems arrive at SVO so quickly while emerging sign languages do not. A likely explanation is that sign languages, unlike spoken languages, can exploit physical space to convey who did what to whom: a function that is commonly performed by case marking in spoken languages. When case marking gradually erodes from spoken languages, it commonly triggers a gradual shift to SVO order (Sinnemäki, 2010), including a stage in which case marking is more likely to be retained for reversible events: a pattern known as differential object marking (Aissen, 2003, Bossong, 1991). It is unsurprising, then, that in pidgins, where case marking is quickly dropped, a shift to SVO happens just as quickly. The ability of sign languages to exploit space for grammatical marking may slow this global shift to SVO, resulting in a system that uses SOV for most events but avoids it for reversible events (see Gibson et al., 2013, for a similar proposal). To test this hypothesis, future experiments could forbid the use of space to test whether SVO emerges as a preferred constituent order when physical space is not available. In other words, participants would be instructed to produce all of their gestures in the space immediately in front of the torso, without having the option to locate some nominal gestures on either side, with action gestures going from one side to the other. This manipulation would place the entire interpretive burden on linear order and could very well influence constituent order preferences in production. Such a study could be a helpful complement to Langus and Nespor's (2010) fourth experiment, which is a non-spatial comprehension task in the auditory modality, using words from spoken language.

The second difference between our laboratory contexts and attested patterns of natural language evolution and change in the world is that our experiments were explicitly designed to probe the cognitive preferences that influence communication from a production standpoint only. However, natural systems are doubtless influenced by constraints on comprehension and acquisition as well, which could slow the pace of change. For example, it could be that the orders that are easiest to produce are different from the orders that are easiest to understand (Hall et al., 2013, report evidence suggesting that this may indeed be the case). This would then require systems to evolve some sort of compromise between these competing pressures. From an acquisition standpoint, successful acquisition implies the faithful reproduction of the patterns in the input. Therefore, at least some aspects of the acquisition process actively work against other pressures that might pressure a learner to inject change into the system. This hypothesis could be tested by having participants act not just as producers in the lab but also as learners and comprehenders. We would then expect to see attenuation of the effects in our present data.

The differences discussed above highlight the potential role of modality (i.e., vocal or manual) in influencing constituent order preferences. This is especially critical since our two main sources of information about word order (studies of homesign/emerging sign languages and large-scale typological surveys of spoken languages) focus on different modalities. It will likely be important to characterize which pressures are invariant across modalities, and which appear to be stronger constraints on one modality than another. Fedzechkina, Jaeger, and Newport (2012) have recently supplied evidence from an artificial (spoken) language learning paradigm showing that adult language learners are sensitive to the relationship between semantic reversibility and case marking. As we observed in pantomime, participants in their study avoided using unmarked SOV (and OSV) utterances to describe reversible events in the spoken modality. Although their participants did not use SVO as an alternative, this may be because (a) case marking was provided as an alternative cue to argument structure and (b) both SOV and OSV order were explicitly modeled for the participants, whereas SVO was not. Future work could address whether a shift to SVO might also happen in the spoken modality if either or both of those cues were absent.

5. Conclusions

In sum, we suggest that the distribution of constituent orders across the world's languages, both synchronically and diachronically, likely reflects cognitive pressures that operate on language. Although attributing some of these pressures to an innate human language faculty may be descriptively successful, the current research sought more functional explanations. The present experiments illustrate how systematic patterns of language structure, such as the emergence of SVO, can be explained by general cognitive principles. These results add to accumulating evidence that elicited pantomime is a helpful empirical tool for studying cognitive influences on language structure, especially with regard to testing hypotheses derived from observations of homesign and emerging sign languages. By the same token, the generalizability of these results to spoken languages should also be considered by comparing results from this paradigm to those obtained from experiments with spoken language as well as computational simulations.

Acknowledgments

The authors gratefully acknowledge members of the Laboratory for Multimodal Language Development, the Language Production Laboratory, the Sign Language Research Group, and the Center for Research in Language at UC San Diego for helpful feedback. We are especially grateful for the partnership of Aylin Kuntay and Ecem Çoban at Koç Üniversitesi. This research would not have been possible without assistance with data collection and coding from Caitlin Scott, Shereen Cohen, Isabel Sharman, Kenny Oyama, Alex Kuo, Elena Churilov, Jayd Blankenship, Kristi Cheng, Lena Sun, and Youngju (Danbi) Ahn. Funding for this research was provided by NIH grant HD051030, NIH grant 5T32DC000041-19, and by the Rita L. Atkinson Graduate Fellowship.

Note

1. We acknowledge that some animals are capable of performing some actions; however, in the context of the experiment, animals were never shown as agents. An analysis with a three-way division of objects into inanimate, animal, and human revealed that animal objects patterned entirely with inanimate objects.

References

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, *21*(3), 435–483.

Aronoff, M., Meir, I., Padden, C., & Sandler, W. (2008). The roots of linguistic organization in a new language. *Interaction Studies*, *9*(1), 131–150.

Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, *332*, 346–349.

Bakker, P. (2008). Pidgins versus creoles and pidgincreoles. In S. Kouwenberg & J. V. Singler (Eds.), *The handbook of Pidgin and Creole studies* (pp. 130–157). Walden, MA: Wiley-Blackwell.

Bossong, G. (1991). Differential object marking in romance and beyond. In D. Wanner & D. A. Kibbee (Eds.), *New analyses in romance linguistics* (pp. 143–170). Philadelphia, PA: John Benjamins.

Branigan, H. P., Pickering, M. J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, *118*, 172–189.

Chomsky, N. (2000). Linguistics and brain science. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 13–28). Cambridge, MA: MIT Press.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*, 257–271.

Coppola, M. (2002). *The emergence of grammatical categories in home sign: Evidence from family-based gesture systems in Nicaragua*. Unpublished doctoral dissertation, University of Rochester, Rochester, NY.

Dayley, J. (1985). Why all languages aren't SOV or VOS, or, how competing motivations lead to natural inconsistency. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 52–64.

Derbyshire, D. S., & Pullum, G. K. (1981). Object-initial languages. *International Journal of American Linguistics*, *47*, 192–214.

Dryer, M. S. (1992). The Greenbergian word order correlations. *Language*, *68*, 81–138.

Dryer, M. S. (2008). Order of subject, object, and verb. In M. Haspelmath, M. S. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures*. Munich: Max Planck Digital Library. Available at http://wals.info.

Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, *473*, 79–82.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *Proceedings of the National Academy of Sciences*, *109*, 17897–17902.

Fischer, S. (1975). Influences on word order change in American Sign Language. In C. N. Li (Ed.), *Word order and word order change* (pp. 3–25). Austin, TX: University of Texas Press.

Gell-Mann, M., & Ruhlen, M. (2011). The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, *108*, 17290–17295.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1–76.

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word order variation. *Psychological Science*, *24*, 1079–1088.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.

Goldin-Meadow, S., & Feldman, H. (1977). The development of language-like communication without a language model. *Science*, *197*, 401–403.

Goldin-Meadow, S., & Mylander, C. (1983). Gestural communication in deaf children: Noneffect of parental input on language development. *Science*, *221*, 372–374.

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, *391*, 279–281.

Goldin-Meadow, S., Özyürek, A., Sancar, B., & Mylander, C. (2008). Making language around the globe: A crosslinguistic study of homesign in the United States, China, and Turkey. In J. Guo, E. Lieven, N. Budwig, & S. Ervin-Tripp (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin* (pp. 27–39). New York: Psychology Press.

Goldin-Meadow, S., So, W. C., Özyürek, A., & Mylander, C. (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences*, *105*, 9163–9168.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language: Report of a conference held at Dobb's Ferry, NY, April 13–15, 1961* (pp. 73–113). Cambridge, MA: MIT Press.

Hall, M. L., Mayberry, R. I., & Ferreira, V. S. (2013) Cognitive constraints on constituent order: Evidence from elicited pantomime. *Cognition*, *129*, 1–17.

Haviland, J. B. (2011). Nouns, verbs, and constituents in an emerging 'Tzotzil' sign language. In R. Gutiérrez-Bravo, L. Mikkelsen, & E. Potsdam (Eds.), *Representing language: Essays in honor of Judith Aissen* (pp. 157–171). Santa Cruz, CA: Linguistics Research Center.

Hawkins, J. (1983). *Word order universals*. New York: Academic Press.

Hawkins, J. (2002). Symmetries and asymmetries: Their grammar, typology, and parsing. *Theoretical Linguistics*, *28*, 95–149.

Hawkins, J. (2007). Processing typology and what psychologists need to know about it. *New Ideas in Psychology*, *25*, 87–107.

Kayne, R. S. (1994). *The antisymmetry of syntax*. Cambridge, MA: The MIT Press.

Keenan, E. L., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, *8*, 63–99.

Kimmelman, V. (2012). Word order in Russian Sign Language. *Sign Language Studies*, *12*, 414–445.

Kouwenberg, S. (1992). From OV to VO: Linguistic negotiation in the development of Berbice Dutch Creole. *Lingua*, *88*, 263–299.

Langus, A., & Nespor, M. (2010). Cognitive systems struggling for word order. *Cognitive Psychology*, *60*, 291–318.

Lee, H. (2001). *Optimization in argument expression and interpretation: A unified approach*. Unpublished doctoral dissertation, Stanford University, Palo Alto, CA.

Levinson, S. C., & Gray, R. D. (2012). Tools from evolutionary biology shed new light of the diversification of languages. *Trends in Cognitive Sciences*, *16*, 167–173.

Li, C. N. (1977). *Mechanisms of syntactic change*. Austin, TX, University of Texas Press.

Maurits, L. (2012). *Representation, information theory and basic word order*. Unpublished doctoral dissertation, University of Adelaide, Adelaide, Australia.

McWhorter, J. (2001). The world's simplest grammars are Creole grammars. *Linguistic Typology*, *5*, 125–166.

Meir, I., Lifshitz, A., Ilkbasaran, D., & Padden, C. A. (2010). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht, The Netherlands.

Morgan, H. E. (2012). *Spatial agreement and word order in Kenyan Sign Language*. Talk presented at the World Conference of African Linguistics 7, Douala, Cameroon.

Newmeyer, F. J. (2002). The SO/OS asymmetry and two interpretations of the MaOP: Commentary on John A. Hawkins, 'symmetries and asymmetries: Their grammar, typology, and parsing'. *Theoretical Linguistics*, *28*, 171–176.

Parker, F. (1980). Typology and word order change. *Linguistics*, *18*, 269–288.

Pullum, G. K. (1977). Word order universals and grammatical relations. In P. Cole & J. Sadock (Eds.), *Syntax and semantics 8: Grammatical relations* (pp. 249–277). New York: Academic Press.

Pullum, G. K. (1981). Languages with object before subject: A comment and catalogue. *Linguistics*, *19*, 147–155.

Roland, D., Dick, F., & Elman, J. L. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348–379.

Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences*, *102*, 2661–2665.

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

Senghas, A., Coppola, M., Newport, E. L., & Supalla, T. (1997). Argument structure in Nicaraguan Sign Language: The emergence of grammatical devices. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the Boston University conference on language development* (Vol. 21, pp. 550–561). Boston, MA: Cascadilla Press.

Sevinç, A. M. (2006). *Grammatical relations and word order in Turkish Sign Language (TID)*. Unpublished master's thesis, Middle East Technical University, Ankara, Turkey.

Singleton, J. L., Goldin-Meadow, S., & McNeil, D. (1995). The cataclysmic break between gesticulation and sign: Evidence against a unified continuum of gestural communication. In K. Emmorey & J. Reilly (Eds.), *Language, gesture, and space* (pp. 287–311). Hillsdale, NJ: Erlbaum Associates.

Sinnemäki, K. (2010). Word order in zero-marking languages. *Studies in Language*, *34*, 869–912.

Slobin, D. I. (1977). Language change in childhood and history. In J. Macnamara (Ed.), *Language learning and thought* (pp. 185–214). New York: Academic Press.

Tomlin, R. S. (1986). *Basic word order: Functional principles*. London: Croom Helm.

Vennemann, T. (1973). Explanation in syntax. In J. P. Kimball (Ed.), *Syntax and semantics* (Vol. 2, pp. 1–50). New York: Seminar Press.

Appendix: Stimulus events

Agent	Action	Patient	Reversibility
Boy	Kiss	Ball	Non-reversible
Woman	Kiss	Ball	Non-reversible
Woman	Kiss	Bike	Non-reversible
Woman	Kiss	Box	Non-reversible
Boy	Kiss	Box	Non-reversible
Boy	Kiss	Car	Non-reversible
Man	Kiss	Car	Non-reversible
Woman	Kiss	Cat	Non-reversible
Boy	Kiss	Cat	Non-reversible
Man	Kiss	Dog	Non-reversible
Girl	Kiss	Trike	Non-reversible
Man	Lift	Ball	Non-reversible
Girl	Lift	Ball	Non-reversible
Woman	Lift	Bike	Non-reversible
Boy	Lift	Bike	Non-reversible
Woman	Lift	Box	Non-reversible
Girl	Lift	Box	Non-reversible
Girl	Lift	Car	Non-reversible
Boy	Lift	Cat	Non-reversible
Woman	Lift	Cat	Non-reversible
Woman	Lift	Dog	Non-reversible
Woman	Pet	Ball	Non-reversible
Boy	Pet	Ball	Non-reversible
Woman	Pet	Bike	Non-reversible
Boy	Pet	Bike	Non-reversible
Woman	Pet	Box	Non-reversible
Girl	Pet	Box	Non-reversible
Boy	Pet	Car	Non-reversible
Man	Pet	Car	Non-reversible
Woman	Pet	Cat	Non-reversible
Boy	Pet	Cat	Non-reversible
Boy	Pet	Dog	Non-reversible
Man	Pet	Dog	Non-reversible
Boy	Push	Ball	Non-reversible
Man	Push	Ball	Non-reversible
Woman	Push	Bike	Non-reversible
Man	Push	Box	Non-reversible
Boy	Push	Box	Non-reversible
Man	Push	Car	Non-reversible
Boy	Push	Car	Non-reversible

(continued)

Appendix. (*continued*)

Agent	Action	Patient	Reversibility
Boy	Push	Cat	Non-reversible
Woman	Push	Cat	Non-reversible
Man	Push	Dog	Non-reversible
Boy	Push	Dog	Non-reversible
Girl	Push	Trike	Non-reversible
Woman	Kiss	Boy	Reversible
Boy	Kiss	Girl	Reversible
Woman	Kiss	Man	Reversible
Boy	Kiss	Woman	Reversible
Woman	Lift	Boy	Reversible
Boy	Lift	Girl	Reversible
Man	Lift	Woman	Reversible
Girl	Lift	Woman	Reversible
Woman	Pet	Boy	Reversible
Boy	Pet	Girl	Reversible
Man	Pet	Woman	Reversible
Boy	Pet	Woman	Reversible
Girl	Push	Boy	Reversible
Woman	Push	Girl	Reversible
Woman	Push	Man	Reversible
Boy	Push	Woman	Reversible