

INTERACTION FOR FACILITATING CONVENTIONALIZATION: NEGOTIATING THE SILENT GESTURE COMMUNICATION OF NOUN-VERB PAIRS

ASHLEY MICKLOS

*Applied Linguistics and Anthropology, University of California, Los Angeles,
Los Angeles, CA
amicklos@g.ucla.edu*

This study demonstrates how interaction – specifically negotiation and repair – facilitates the emergence, evolution, and conventionalization of a silent gesture communication system. In a modified iterated learning paradigm, partners communicated noun-verb meanings using only silent gesture. The need to disambiguate similar noun-verb pairs drove these "new" language users to develop a morphology that allowed for quicker processing, easier transmission, and improved accuracy. The specific morphological system that emerged came about through a process of negotiation within the dyad, namely by means of repair. By applying a discourse analytic approach to the use of repair in an experimental methodology for language evolution, we are able to determine not only if interaction facilitates the emergence and learnability of a new communication system, but also how interaction affects such a system.

1. Introduction

The purpose of this experiment is to investigate the evolution of a communication system, along with other discourse features that accompany such a system, in an experimental setting that models natural interaction. The previous studies (see below) examined these questions, albeit in a limited interactive environment. By expanding the interactive abilities and opportunities of the participants, we hope to gain more insights into specific aspects of language evolution, but in a setting that mirrors more natural communicative practices. That is, we aim to observe the interactive strategies that may be co-opted to facilitate alignment to form-meaning matches in a novel communicative system. With a goal-based paradigm, we hope to approximate a context in which communication is required to benefit individuals, and in which negotiation of new form-meaning matches is required.

1.1. Background

Human language learning and use are inherently interactive, however, the role of interaction in the emergence and conventionalization of new communication systems remains under-investigated in language evolution research. Experimentalists (Kirby, Cornish, & Smith, 2008) have demonstrated in the laboratory how languages evolve compositional structures as they are transmitted culturally through iterated learning from one individual to another,

but only recently have others explored the effect of interaction. Fay *et al.* (2010) extend the iterated learning paradigm to social collaboration in using a graphical sign system to convey abstract concepts. The graphical representations become less iconic through iterated participant use, while task accuracy shows the systems were successful in both the dyadic and community groupings. They conclude interaction was a crucial element to the creation of a new, shared communication system between individuals and communities of individuals.

Similarly, Healey *et al.* (2007) incorporate concurrent communication in a graphical drawing task, allowing participants' "mutual modifiability" of one another's outputs. They find an accumulative history of interaction contributes to lineage specific conventions, supporting the hypothesis of a conventionalized communication system emerging from interaction opportunities. Interaction provided for the use of devices that supported coordination (agreement/disagreement) and served editing functions that identified specific aspects of the pairs' drawings, which facilitated alignment to meaningful elements of the representations. Healey and colleagues suggest a repair-driven co-ordination process in communication could likely be observed in emerging gestural systems, a type of interaction that may involve more use of eye gaze, the face, and the body than graphical systems tasks.

By incorporating the study of repair (as in conversation analysis: Jefferson, 1974; Clark & Krych, 2004) into an interaction-based iterated learning paradigm, we hope to demonstrate how certain aspects of interaction facilitate the conventionalization of a gestural communication system and the effect of repair on such a system.

2. Methodology

In an experimental setting, participants (English-speaking, right-handed, non-signers) engaged in an interactive communication game aimed at producing an evolving communication system through negotiation over simulated generations via a gradual turn-over of participants. Participant dyads were required to interact with one another during a card selection task, alternating between Director and Matcher roles. However, their communication was limited to silent gestures (Goldin-Meadow *et al.*, 2008; Schouwstra, 2012), a communication similar to "Charades" in which the participants can use their hands/arm, as well as the face, to communicate without vocalizations. An Observer watched the dyad interact through one "generation," and then participated in the subsequent.

2.1. Stimuli

The stimuli for this experiment were two sets of similar-meaning cards. One set was comprised of verb meanings (with the progressive “-ing”), while the other was comprised of noun meanings (with the indefinite article “a” or null article). However, the verb-noun pairs had minimally contrastive features when gestured with speech (e.g. spontaneous co-speech gesture of “hammering” and “a hammer” are typically produced with a closed fist moving up and down on a real or imagined surface). A total of 32 noun-verb pairs comprise the set (for a total of 64 target tokens). The verb cards were represented in the progressive form (“-ing” suffix), while the noun cards adopted the indefinite article “a” or a null article (as with “snow”). Similar-semantic distractor tokens that contributed to task difficulty and masked the task objective (i.e. some did not conform to the grammatical forms of the target noun-verb pairs) were included as well.

2.2 Procedure

Dyads, using only silent gestures, played an interactive card selection task for four rounds per simulated generation. Each participant in the dyad had a set of cards “in hand” (n=10, 8 target matches, and two similar-semantic distractors) which could only be viewed by the holder. For each round, the target set of noun-verb pairs were grouped for similar semantics (e.g. “shovel” and “rake” cards were in the same round), but were distributed between the participants randomly (for noun-verb pairs, as well as for the number of nouns or verbs in each participants’ set). Participants took alternating turns requesting one card at a time from their partner’s cards-in-hand with silent gesture.

Participants were placed in one of two conditions. In the “Standard” condition, each participant had one opportunity per turn to make a request for the target card. Once the gesture was performed (note that the gesture could be performed or changed any number of times before the Matcher had to provide a guess), the Matcher provided a card. If the card was a correct match, the participants placed it to the side (a “matched matrix”); however, if the card was incorrect, the Matcher returned the card to their in-hand board, and the Director placed the target card to the bottom of their set. The Director would have another opportunity to gesture the target card later in the round. The “Do-Over” condition proceeded in the same manner, except in the instance of an incorrect guess from the Matcher: the Director had an immediate turn for a “do-over” in which they could once again gesture the target meaning. This condition was meant to drive immediate repair and reflect how repair is performed in natural conversation. The hypothesis is that more immediate opportunities to repair would result in more rapid alignment to corrected gesture strategies.

3. Results

Here we present the results collected from three chains (of 4 generations) in each condition.

3.1 Noun Marking Strategies and Handshape

Over generations, participants developed gestures to disambiguate nouns and verbs, focusing especially on nouns. Marking systems were innovated and passed down generations, though not all chains used the same system. The three most common, and consistent, noun markers were “object point” (O-Point, one hand points to the other performing the noun-related gesture), “object emphasis” (O-Emh, in which the emphasis could denote shape, handheld-ness, etc), and use of the “index finger” held up like the numeral “one.” Some generations fluctuated between the use of markers, but eventually one strategy dominated. In addition, each marker conformed to its own word order which was also innovated and recognized by the participants. Nearly all chains made use of a noun marking system to disambiguate noun-verb pairs, and by the final generation, many chains used the noun marker reliably.

Do-Over chains exhibited less variability in marker use, namely in Chains 1 and 2, while Chain 3 showed more alignment to a dominant marker in Generation 3. All Standard chains had non-dominant use of markers initially, wherein Chains 2 and 3 displayed a switch in the marking strategy that eventually overtook the other form in use, occurring in generations 2 and 3, respectively. Overall, the Standard condition marked targets more than the Do-Over condition. The Standard condition participants marked more nouns than were left unmarked, and while the primary overall strategy was pointing to the imagined object, each chain exhibited different dominant strategies.

The semantic category to which a noun belonged influenced how its gestural structure manifested. Instrument nouns (23 of 32 noun targets, such as “A Hammer” or “A Shovel”) were most often gestured as “Handled;” that is, the participant held the imagined object in their hand as they performed the action related to the instrument. This is an action-variant (Ortega et al, 2012), by which an object is represented by the action associated with it. A notable exception is “A Phone,” which is most frequently embodied as a hand in the shape of a phone. Overall, there was a trend to use action-variants (“Handled”) over object/perceptual variants (“Embodied”) for noun targets.

3.2 Accuracy and Timing

Accuracy in the Do-Over condition remained similar over generations (an average of 85%), while accuracy in the Standard condition increased from 75%

to 87%. Accuracy spiked most drastically from Generation 1 to Generation 2, in both conditions. We might attribute this trend to the role of the Observer, who comes into the interacting dyad as a knowing participant having seen the previous dyad perform the task. Using the average accuracy of guesses, each condition demonstrated higher average accuracy when guessing noun-targets with the Index Finger marker. The average accuracy using the Index Finger marker was also higher than not using a marker at all; though, O-Emph and O-Point markers resulted in similar accuracy as not marking at all (with O-Point being slightly more advantageous).

The length of gestures did not vary greatly, with the exception of repair sequence gestures. Overall, gestures were between 2 and 4 seconds, though, on average, Do-Over condition gestures were shorter than Standard condition ones for all generations. Though the general trend of decreased gesture length, by seconds, appeared trivial, within the context of this study, minute changes have large effects. In general, as gestures were transmitted through generations, their length shortened, indicating the potential constricting of the gesture space and the diminished need to elaborate upon initial gestures.

3.3 Repair

All three repair strategies (Repetition, Clarification, and Reformulation) were present in the two conditions, though to differing degrees. Clarifications, or moderate modifications on the initial (trouble-source) gesture, were the most frequent strategy in either condition (Standard=52, Do-Over=63). A clarification repair required the Director to highlight or emphasize an element of the initial gesture, typically done through the shaking of a hand (in “handled” gestures) or emphasizing a marker. Reformulation repairs are complete modifications of the initial gesture, often involving a change in handshape. Reformulations were more frequent in the Standard condition (n=23) than the Do-Over condition (n=9), which may be related to the inability to perform a repair after an incorrect guess, thereby requiring Directors to re-do gestures without knowing what the trouble source or misinterpretation is. Repetition, either full or partial (but crucially not involving emphasis), was similarly used in both conditions (Standard=21, Do-Over=20), and required no modification to the initial gesture.

The evolution of repair strategies over simulated time can be an index for the degree of conventionalization the communication system has achieved. While the amount of repair does not decrease with each generation (as in Fay *et al.* (2010)), it may still be indicative of the conventionalization process in which repairs conform to prior preferences that are then re-constituted through the repair itself. This study evidences the process of conventionalization in the

repair sequences performed by the Directors themselves (self-repair) or those provided by the Matcher (other-repair). Other-initiated repair is an attempt to get the Director to align with the Matcher's conceptions of the silent gesture system. When a Matcher initiates repair with facial gestures (e.g. furrowed brow) or returned eye gaze to the Director, they communicate a misunderstanding of the prior, and attempt to make the Director repair to a more communicable gesture. If, on the other hand, the Matcher performs a repair (i.e. gives the repair options), they are demonstrating a more effective strategy which may then be taken up in subsequent gestures. Both repair sequence types bring about a saliency in the gesture form (and even order) that allows for participants to fixate on more readily transmittable and communicative form-meaning matches.

The participants' repair strategies also played a role in the conventionalization of noun marking systems. Clarification repairs were the most frequently used repair strategies, particularly with noun targets. In the Do-Over condition, the immediate "repair" turn allowed for more repair use, in which we witness the Director's clarification strategy largely involving the use of a marker. Since clarifications require emphasizing or highlighting a previously gestured feature, greater saliency of those features emerges and becomes a substrate from which participants can build meanings, including compound meanings. Over generations, participants develop conventionalized marking systems from the emphasized forms. Repair, a feature of interaction, does facilitate the systematization of the gestures.

4. Discussion

As generations modify the previous generation's output, descriptive gestures become lexicalized such that an identifiable and systematic marking system develops in all chains. Typically, a marking gesture emerges from the need to disambiguate noun-verb pairs following evidence of inaccurate guesses or confusion. The ubiquity of the marking systems in this experiment demonstrates its efficiency in communicating meanings, and, specifically, in disambiguating nouns from verbs. The immediate opportunity to modify the previous gesture (or gesture space) leads to repair-driven markers that conventionalize within a chain. Clarification repairs could be driving the gestural forms that are conventionalized; as the need to clarify a target meaning arises, and repair is performed, the strategy for repair becomes part of the gestural system. Here, the repair strategy "clarification" often results in an emphasis on the object-ness of the target word, including its handheld-ness, shape, or singularity. Marker clarifications are common, and their saliency could drive the marking system into conventionalization.

4.1 The Case for Interaction

The crux of this study is to demonstrate that interaction, namely face-to-face contingent negotiation, has a role in the emergence and evolution of communication systems. Using an iterated learning paradigm, we can simulate language evolution over generations, but incorporating dyadic interaction with an observer highlights the interactive component of language use and transmission. Both conditions followed this model, except that one sought to mimic natural conversation in terms of turn taking in repair sequences. Though some differences existed between the conditions, here we are concerned with the broader role of interaction in facilitating conventionalization.

The modified iterated learning model in this study derived from earlier studies of iterate learning of alien languages, though many did not provide contingent interaction scenarios. Kirby, Cornish, and Smith (2008) found that structure emerged between generations 4 and 9, with variability by chain. These interactions took place in computer-modulated communication, not face-to-face interaction. Nonetheless, a compositional structure emerged that allowed for motion, color, and shape to be encoded by different morphemes that could be strung together to encode meanings such as “black square moving in a circle.” As noted, some chains attained their maximum level of structure by generation, while others did not until generation 9; often early-structure chains collapsed again into less-structured use of the language.

Face-to-face interaction, on the other hand, may promote quicker and more efficient transmission of the newly emerged language structure. Fay *et al.* (2010)’s graphical communication task saw isolated pairs reduce drawing complexity through round 3 (an approximate to “generation 3”); however, community pairs took longer to achieve less complexity in drawings (note: simple drawings require less graphical refinement but still convey meaning). The second condition here, though, is more akin to speakers of different languages – or dialects – coming together to communicate. The current study, in contrast, mimics the language learning environment of novices, including children and foreign language learners.

In a more naturalistic setting, Master, Schumann, and Sokolik's (1989) experimental creation of Persian and German pidgins required participants to extend their pidgin lexicon to new meanings. These systems displayed compositionality, regularity, and stabilized compound noun forms by their fourth use. The rapidity of this systematization may have resulted from the extensive negotiation that took place when innovating new lexical items from a given set. In fact, just as with Fay *et al.*, here noun compounds simplified by 66 % (for example, an initial noun compound may have been comprised of 6 individual

tokens, strung together, but after negotiation, decreased to about 3 tokens). It should be noted that this happened at the community level, as pairs exchanged partners in a gradual-turnover fashion, similar to the generational changes incorporated in our study.

In sum, face-to-face interaction might allow for more efficient and reliable transmission over generations. Interacting dyads, and groups to some extent, reach stabilization earlier than some non-interactive conditions, and may lead to the emergence and fixation of many features of systematicity simultaneously (being promoted by negotiation of varied aspects of meaning through moment-to-moment shifts and alignments). Furthermore, allowing for interaction that more closely approximates language use practices (e.g. turn-taking patterns, use of facial gestures, etc), which may be present even pre-linguistically, is essential to understanding how potential communicators rely on interaction structures to facilitate the conventionalization of linguistic systems. Teaching and learning are naturally interactive activities, therefore studies of the cultural transmission of language should consider how features of interaction impact the trajectory, in terms of evolving systematicity, of those languages.

References

- Clark, H.H., & Krych, M.A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50(1), 62–81.
- Fay, N., Garrod, S., Roberts, L., & Swoboda, N. (2010). The interactive evolution of human communication systems. *Cognitive science*, 34(3), 351–86.
- Goldin-Meadow, S. et al (2008). The natural order of events: How speakers of different languages represent events nonverbally. *Proceedings of the National Academy of Sciences of the United States of America*, 105 (27), 9163-9168.
- Healey, P.G.T., Swoboda, N., Umata, I., & King, J. (2007). Graphical language games: interactional constraints on representational form. *Cognitive science*, 31(2), 285–309.
- Jefferson, G. (1974). Error Correction as an Interactional Resource. *Language in Society*, 2, 181-199.
- Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(31), 10681–6.
- Master, P., Schumann, J., & Sokolik, M.E. (1989). The experimental creation of a pidgin language. *Journal of Pidgin and Creole Languages*, 4(1), 37-63.
- Ortega, G., Sümer, B., & Özyürek, A. (2012). Type of iconicity matters: Bias for action-based signs in sign language acquisition, 1114–1119.
- Schouwstra, M. (2012), *Semantic Structures, Communicative Strategies, and the Emergence of Language*. PhD Thesis, University of Amsterdam.