2016

The Effects of Co-Occurrence on the Collaborative Process of Establishing a Reference

Nicole Maslan
Claremont McKenna College

The effects of co-occurrence on the collaborative process of establishing a reference

Nicole Maslan

Claremont McKenna College

The author presents an analysis of how speakers establish references in conversation. Further, this paper focuses on what words of a reference are conventionalized as speakers coordinate multiple times. The author explores how the co-occurrence of the reference terms with the referent can be a good predictor of what words are conventionalized over time. In order to study this, the author created an online version of the reference game from Clark and Wilkes-Gibbs (1986) experiment, where a matcher and director must describe a set of ambiguous shapes to each other many times. By creating an online version of this reference game the author was able to gather significantly more data and analyze the data with computational tools. Results prove that co-occurrence is a useful predictor of terms which are conventionalized, providing a first step for accounting for statistical inference in the process of conventionalization.

Introduction

An incredible feat of human cognition is how we are able to learn so much from so little. A primary example of this is found in word learning from young children. With only a few examples of what *dog* means, children are able to identify a wide array of dogs with impressive accuracy (Xu and Tenenbaum, 2007). In fact, an infant is still a far better learner than the most complex AI. But this raises the question of how are we able to create such rich inferences from such sparse data. This question is particularly important during the process of conversation, where we are not only comprehending words but also producing them. This paper will focus on how people are able to establish a new reference for something during conversation and how that reference changes and becomes conventionalized after repeatedly using it.

Up to this point, prior research has focused mainly on two forms of learning, deductive inference and associative learning. The deductive approach focuses on deducing what something means by eliminating hypotheses that are given through structured representations of knowledge that the learner already has about the world. The associative learning approach works by keeping track of the frequencies of certain words and contexts. A third model that has helped bridge these two approaches and yielded consider success in replicating human inference is a bayesian probabilities framework. This paper hopes to take a step towards applying bayesian inference to the process of conventionalization by examining how the co-occurrence of a word with a target shape predicts if it will become conventionalized.

Literary Model

For a long time, psychologists and linguists looked at language understanding and comprehension through what Clark & Wilkes-Gibbs (1986) termed as the literary model. In this

literary model a speaker refers as if writing to a distant reader. For example, if the speaker utters the noun phrase *the dog with the black spots* to someone, the speaker assumes that their noun phrase will be clearly understood and uniquely identified by their listener. This act of successful communication is considered to take place only during the time that she issues this one noun phrase. In this case the act of referring begins with *the* and ends with *black spots*. Further, the speaker has entire control of this referential process and there is no input or interaction with the listener. The listener should hear the definite description of *the dog with the black spot* and if he is successful, correctly infers the identity of the referent. His understanding or lack thereof does not change the referential process of the speaker. This view of the referential process often led researchers to study language only in terms of monologue (Pickering & Garrod, 2004; Clark & Wilkes-Gibbs, 1986;Linell, 2004; Tanenhaus & Brown-Schmidt, 2008; Clark, 1996).

When studying monologue, researchers focused on how words were comprehended in isolated settings, such as reading. They have derived most of their predictions about the referential process from generative linguistics and have developed theories from isolated, decontextualized sentences (Pickering & Garrod, 2004). However, some researchers have pointed out the critical importance of studying dialogue (Clark & Wilkes-Gibbs, 1986). If language is in fact for communication then studying it in isolation may be missing important aspects of how it functions to communicate information (e.g. functions that you might find in conversation).

Unique constraints on conversation

There are several important constraints on conversation that should make this referential process look different than the literary model (Clark & Wilkes-Gibbs, 1986). One is that

speakers lack the time to revise and draft out an ideal reference (Christiansen & Chater, 2015). Without having the time to create the clearest reference that they can, we should expect them to use techniques unique to conversation in order to establish a good reference. This also affects the listener, who must both comprehend the speaker's utterance and then reply in a nearly instantaneous manner. In order to have this back and forth type of communication speakers and listeners must synchronize their actions exceptionally well (Clark & Wilkes-Gibbs, 1986). What these constraints and differences lead to is communication that is inherently interactive and contextualized (Pickering & Garrod, 2004). Recent literature in this area even suggests that language understanding can essentially be seen as a special case of social cognition (Goodman & Stuhlmüller, 2012). By only studying language in isolation we would miss important pragmatic uses of language. To Clark, and many others, dialogue captures the collaborative process of how referents are created in real life.

Clark (1996) pointed this out by drawing a distinction between "language-as-product" and "language-as-action" traditions. The language-as-action tradition has focused on how utterances are interpreted with respect to a particular context. This takes into account the goals and intentions of the participants along with how they use and formulate common ground (Clark 1985; 1996, Clark & Marshall 1981). In contrast, the language-as-product tradition has typically tried to merge information-processing psychology with generative grammar to produce a mechanistic account of how people compute different levels of representation (Pickering & Garrod, 2004). As we will see, the literary model is not sufficient for capturing the complexity of creating a referent.

Clark and Wilkes-Gibbs (1986) show that in conversation the referential process is highly interactive between the speaker and listener. Unlike in the literary model, where the entire referential process is contained within one simple noun phrase, speakers tend to self-correct their noun phrase by adding on to it, restarting, and changing course mid phrase (Schegloff et al., 1977; Levelt 1983; Cohen, 1985). Often, the speaker will invite the listener directly into the referential process. An example of this is shown from Cohen (1985):

S. Okay now, the small blue cap we talked about before?

J. Yeah.

S. Put that over the hole on the side of that tube--

J. Yeah.

S. --that is nearest to the top, or nearest to the red handle.

Here, J uses noun phrases such as "the small blue cap we talked about before?" with a rising intonation at the end in order to invite J's input about whether the reference is making sense. With J's input of "yeah," S then continues the process by tacking on another noun phrase. We can see how the reference is not established by simply uttering "the small blue cap" but rather uses J's input of "yeah" as part of the referential process.

When this referential process is seen as a collaborative process rather than a one way transaction, we can begin to see the incredibly complex amounts of coordination that are involved. In fact Lewis (1969) cleverly reframed dialogue as a coordination game. He described it as a game where both participants "win" if both understand the dialogue, and neither "wins" if

they do not understand (Pickering & Garrod, 2004). It can be helpful to think about a complex and messy situation like dialogue as a coordination game since this has been studied quite broadly.

Coordination games

A coordination problem arises when two people have common goals or interests and each person's actions depends on the actions of the other (Clark, 1996). In many cases this requires that participants strategically reason about what the other will do. Thomas Schelling isolated these types of problems with several games he created, now known as Schelling games. In order to "win" at these games, Schelling expressed the importance of mutual expectations. He writes that "What is necessary is to coordinate predictions, to read the same message in the common situation, to identify the one course of action that their expectations of each other can converge on. They must "mutually recognize" some unique signal that coordinates their expectations of each other" (Schelling, p. 54). Humans have created entire languages to provide unique signals to coordinate their expectations of each other. Clark (1996) explains that coordination devices give participants a reason to believe that they and their partner will perform the same joint action. Coordination devices can come in many forms but they must fundamentally be salient to both participants with respect to their shared common ground (Clark, 1996). Some types of coordination devices include explicit agreement, precedent, and convention. But Clark and Lewis argues the the most important of these coordination devices is convention.

Conventions

David Lewis (1969) thought of conventions as a society's best solution for a recurring coordination problem. Lewis defined a convention as a regularity in behavior, partly arbitrary,

common ground in a given community, and a coordination device to solve a recurrent coordination problem. One example of a recurrent coordination problem is how to greet people. Nearly every society has a convention to deal with this. In some countries it is by hugging or kissing on the cheek while in others it is bowing. Not only Lewis, but many others have thought of conventions as the foundation of social and economic life (Young HP, 1993; Lass, 1997; Lewis, 1969; Hechter & Opp, 2001; Centola & Baronchelli, 2015). Similarly, the lexicon of language has grown out of the recurrent coordination problem of having to repeatedly describe certain events and entities (Dell, 2004). If conventions are such an essential part of society and human life, it is worth studying in more detail how exactly they are created. Clark and Wilkes-Gibbs initially began exploring how linguistic conventions were formed in a reference game they experimented with in 1986.

Tangram experiment

In 1986, Clark and Wilkes-Gibbs ran an experiment originally created by Krauss & Weinheimer (1964). In Clark and Wilkes-Gibbs's version of the experiment they placed two students across from each other at a table with an opaque screen between them. Each student had 12 cards placed in front of them in a 2x6 grid where each card contained a different tangram figure. The 2x6 board of the tangram shapes is shown below in Figure 1.

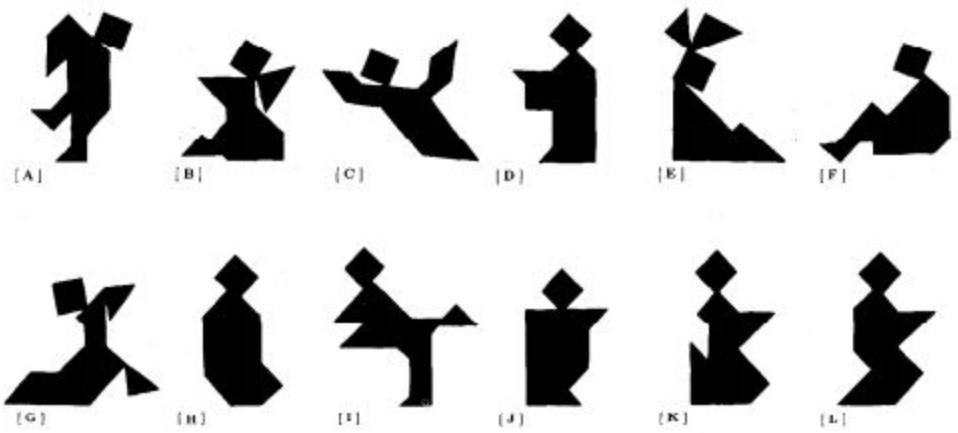


Figure 1. The 12 tangram figures used in the study displayed in a 2x6 grid

In the experiment, one student was assigned the role of the director and the other was assigned the role of the matcher. The director and matcher had the same types of shapes on their board but in different orders. The director's job was to have the matcher rearrange her tangram figures to match the ordering of the director's board. They could talk back and forth as much as was needed but were unable to view each other's boards. After they matched the arrangements of their boards the director's board was rearranged with a new target sequence and the matcher's board was shuffled. They repeated this task six times.

Clark and Wilkes-Gibbs made several predictions of how the task would turn out based on their collaborative view of how references were established. First, they expected that when the matcher and director initially referred to a figure, they would issue many nonstandard noun phrases. Once the participants established a mutual understanding of the referent through this initial negotiation they would be able to refer to it more quickly and more often with a standard noun phrase. This meant that the number of utterances should decrease over the trials and that the noun phrases to refer to the tangrams should become simpler and more standard.

Both of these predictions were supported by the results of their experiment. They found that the number of words the director used decreased dramatically as the trials went on. The directors used an average of 41 words per figure in trial 1 but only 8 words per figure in trial 6, as seen below in Figure 2.

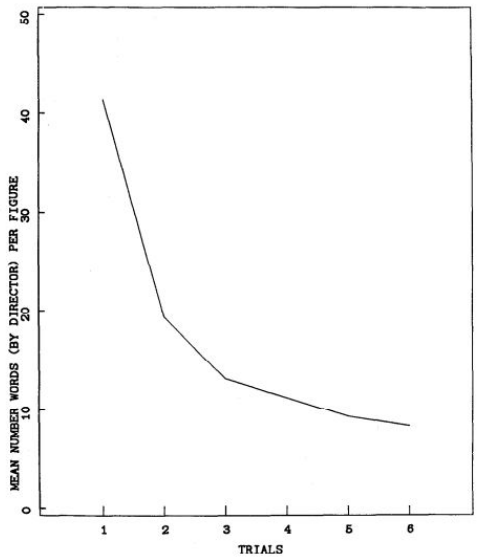


Figure 2. Average number of words used by director per figure on each trial

They also found that the number of exchanges between director and matcher decreased dramatically over the rounds. The director averaged 3.7 turns (as in utterances) per figure on trial 1 but only about one per figure by trial 6. This followed a similar trend as displayed in Figure 2, where the greatest decline occurred between trials 1 and 2 and then leveled off quickly between trials 3 and 6. The also found that the number of elementary noun phrases (considered standard noun phrases) increased from trials 1 to 6, while the number of the nonstandard noun phrases (like episodic and installment) decreased as the trials went on.

From their results, Clark and Wilkes-Gibbs proposed that participants operated under an implicit goal to minimize the collaborative effort it took to establish a reference. Clark and

Wilkes-Gibbs suggested that trying to minimize collaborative effort meant references became as simple as possible and that there were less exchanges between partners when possible. This not only guided what references participants uttered on a turn by turn basis but also acted at large to create general trends of reference making that looked like Figure 3 show below.

For a broad picture of what occurred, consider this very simple series of utterances by one director for figure I on trials 1 through 6:

1. All right, the next one looks like a person who's ice skating, except they're sticking two arms out in front.
2. Um, the next one's the person ice skating that has two arms?
3. The fourth one is the person ice skating, with two arms.
4. The next one's the ice skater.
5. The fourth one's the ice skater.
6. The ice skater.

Figure 3. A snapshot of utterances used to describe a tangram on rounds 1-6

In Figure 3, we can see how the references to this tangram change over each round and became simpler and more concise. However, looking at it another way, we should think that every utterance is the most concise form possible at that time, given their current mutual knowledge. This is where the other central principle that Clark and Wilkes-Gibbs proposed comes in. Conciseness and simplicity of noun phrases was constrained by making sure that each participant understand what was said. Clark and Wilkes-Gibbs described this as the principle of mutual responsibility, claiming: "The participants in a conversation try to establish, roughly by the initiation of each new contribution, the mutual belief that the listeners have understood what the speaker meant in the last utterance to a criterion sufficient for current purposes" (Clark & Wilkes-Gibbs, 1986). This explains how at the start of the game, participants uttered many different noun phrases, corrected what they said and asked questions in order to establish that the other knew what they were referring to. At that point, that was still the minimal collaborative

effort in order to create a reference with each other. However, as common knowledge was built up between these pairs they needed less descriptors in order to establish understanding and thus their phrases contained less words, they used less exchanges, and their noun phrases were simpler. Clark and Wilkes-Gibbs findings were replicated in a coordination game conducted by Garrod & Anderson (1987).

Maze task

In their maze task, Garrod & Anderson (1987) placed two people in different rooms and displayed a different maze on each screen. An example of what the mazes looked like is shown below in Figure 4.

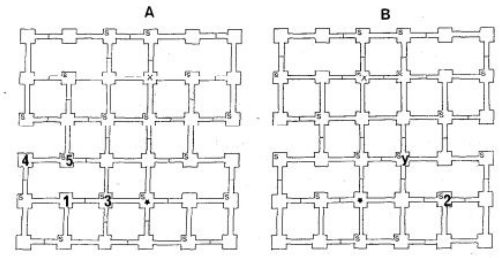


Figure 4. The X marks each player's position and the numbers and the letters refer to different goal positions used in the trials.

The goal for each player was to move their position markers one link at a time until they reached their assigned goal. The challenging part of the game was that each maze had obstacles, located by the bars in A and B of figure 3. Neither player could see these obstacles in their own maze, but could see these obstacles on their partners maze. In order for a player win the game, they needed the guidance of their partner. The general pattern of the gameplay would be making a move and then speaking to the other player about where they should move next. In order to make

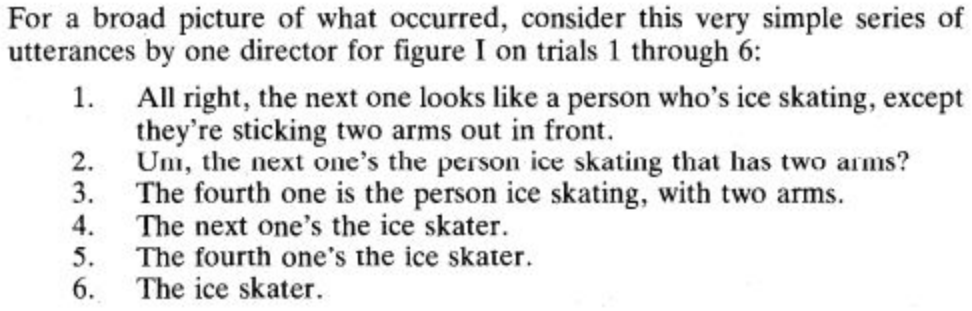
moves, however, participants had to figure out a way to talk about the layout of the maze to each other.

Garrod and Anderson found a high degree of semantic co-ordination emerge between partners very quickly. They found that participants not only exhibited linguistic entrainment but that their various descriptions emerged from a shared conception of the maze and a set of interpretation rules for decoding expressions. They found that participants generally used one of three different mental models participants to describe the maze: path ("see the bottom, go two along, and two up"), figural (depended on referring to a figuration of nodes, using terms like "a square," and "T shape") , and co-ordinate ("I'm on the fourth row and third column"). Using these schemes gave partners useful constraints to decode messages from each other and enabled participants to create local conventions that mitigated the inherent ambiguity of using conventional phrases from the larger linguistic community.

Creating a complete description scene of the maze posed participants with a challenging coordination problem. Yet, Garrod and Pickering were impressed to find that coordination was not reached through explicit coordination but through an interactive co-ordination of input and output. Clark and Wilkes-Gibbs (1986) found something quite similar in their experiment. Participants rarely explicitly talked about what name they would call each shape. The essential question remains about how they are able to converge on common names without talking about this.

Clark and Wilkes-Gibbs (1987) and Garrod and Pickering (1987) provide significant insight to describe what the conventionalization process looks like. We can see that establishing a referent is a fundamentally collaborative process and that there is often lexical alignment and

even greater schematic alignment to solve coordination problems. However, we don't fully understand why certain words get conventionalized and others don't. Looking at figure 3 again, we can ask ourselves why "sticking arms out in front" doesn't get conventionalized but "the ice skater" does.



For a broad picture of what occurred, consider this very simple series of utterances by one director for figure I on trials 1 through 6:

1. All right, the next one looks like a person who's ice skating, except they're sticking two arms out in front.
2. Um, the next one's the person ice skating that has two arms?
3. The fourth one is the person ice skating, with two arms.
4. The next one's the ice skater.
5. The fourth one's the ice skater.
6. The ice skater.

Figure 3. A snapshot of utterances used to describe a tangram on rounds 1-6

One aspect of the conventionalization process that hasn't been successfully modeled is the effect of context. Perhaps, the term "ice skater" was only used for this shape where the words contained in the phrase "sticking arms out in front" were used to describe a variety of other tangrams as well. It is possible that we could predict what words get conventionalized by looking at how likely they uniquely co-occur with a given tangram. Given what we know about the role bayesian inference in human cognition, it is very feasible that people are more likely to select for words with high co-occurrence values and conventionalize those words over time.

The present study

To explore the role of co-occurrence, I will replicate the original experiment done by Clark and Wilkes-Gibbs (1986). This experiment offers an excellent framework to study the development of linguistic conventions. In order to test the predictability of co-occurrence I will

need more data (the original experiment collected 8 pairs of participants) and data that is in a format that can easily be studied with computational tools. In order to collect more data that is in a more usable format, I am digitalizing the experiment so that participants play each other online and communicate through written text rather playing the game in person and communicating verbally. Once I have collected the data I will see if the words with the highest co-occurrence for a given tangram on round 1 match the words used on round 6 for that same tangram. I will measure co-occurrence by finding each words pointwise mutual information value (a measure of association) for a given tangram on round 1. Since the types of words used to describe a tangram typically stabilize after round 4, I will think of the words on round 6 as conventionalized for that given shape. Given what we know about the importance of statistical inference on word learning (Xu & Tenenbaum, 2007) it is possible that the words that are conventionalized at the end of the experiment have a high co-occurrence value at the start of the experiment. Therefore, I predict that the words with the highest co-occurrence values on round 1 will be more likely to get conventionalized by round 6 than random words used on round 1.

Methods

Participants

Twenty-four pairs of participants participated in this study. All participants were recruited from Amazon's Mechanical Turk (mTurk) online crowd-sourcing service and all reported being native English speakers and residents of the United States. Each participant received a payment for participating in the study and bonus payments were awarded based on how well they performed the task.

Materials

For this experiment, I created a digital replication of the tangrams matching game used in the experiment done by Clark and Wilkes-Gibbs (1986). To digitalize their experiment and put it online I principally used Node.js. Participants opened the experimental game in their individual web browsers and were presented with a chat box to type and send messages to their partner along with a 2x6 board of tangrams. The round number they were on (1-6) was displayed on the right side of their screen. The matcher also had a button on their screen in the upper left hand corner to submit the board and begin the next round. Participants each interacted with a screen that looked like Figure 5, shown below:

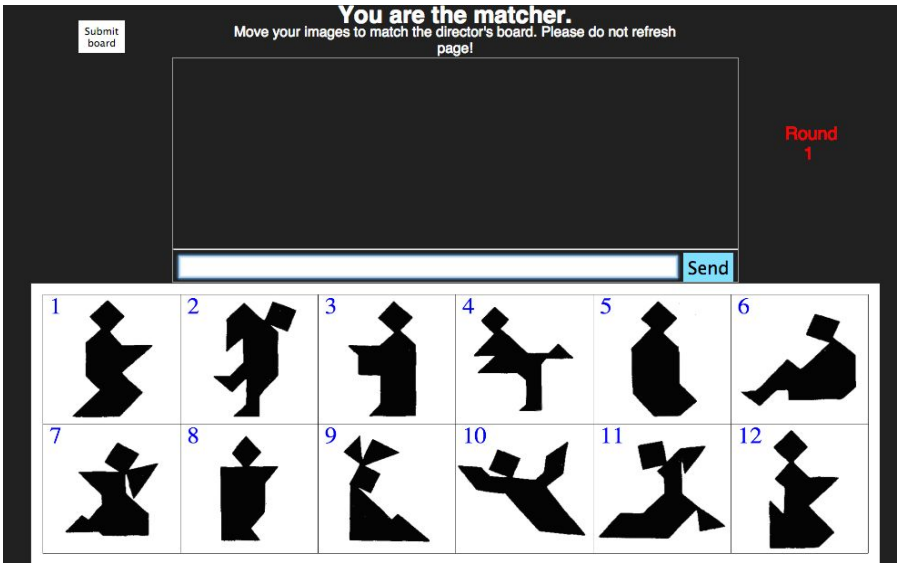


Figure 5. The interface of the game that was displayed on each participant's screen

Instructions and game rules

Before starting the game, participants were presented with a page of instructions on the goals and rules of the game. Participants were told that they would be assigned a matcher or director role at

the start of the game. Each player would be presented with a board of 12 tangrams and a chat box to send messages to each other. The goal of the game would be for the matcher to move her tangrams to match the order of the tangrams on the director's board. Only the matcher would be able to drag her tangrams to different spots on her board. Because of this, the director would have to type and send messages explaining which tangram she should move. The matcher was allowed to ask questions and engage in dialogue with the director. When the matcher and director believed that they had matched up their boards correctly, the matcher would press the submit board button on the screen and the players would enter the next round. The players were informed of their score (how many tangrams they matched correctly with each other) at the start of the next round. While they were notified of their score, they were not told which tangrams they matched correctly. Also, at the start of each round, the matcher and director's boards were shuffled and their chat boxes were cleared. Otherwise, all goals and elements of the game remained the same. After successfully completing a short set of simple multiple choice questions to make sure they understand the game, the game began in their browser and they each were assigned the matcher or director role. Each dyad played a total of 6 rounds of the game.

Hand coding the tangrams

My analyses required me to be able to look at the words uttered for each tangram. In order to look at this I had to go in and hand code what tangram was being referred to for each utterance. After the games were gathered, I went into the text files which recorded the typed exchanges of the participants and tagged what tangram was referred to during each exchange. I knew what tangram was being talked about by checking which tangram was moved by the matcher after an utterance was said and seeing if that tangram was a correct match for the director's board. As I

continued doing this for many games it became much easier to determine what tangrams were being talked about.

Results

Calculating PMI score

The co-occurrence value of a word was measured by calculating its pointwise mutual information value. The equation to calculate PMI is shown below in figure 6:

$$\operatorname{pmi}(x;y) \equiv \log \frac{p(x,y)}{p(x)p(y)} = \log \frac{p(x|y)}{p(x)} = \log \frac{p(y|x)}{p(y)}.$$ [1]

Figure 6. The PMI equation I used to calculate co-occurrence

Here, for P(x|y), x is a given word on round 1 used to describe a specific tangram, y. To compute this I counted the frequency that the given word x occured when speaking about tangram y. For p(x), I calculated the global frequency of the given word x used to describe all tangrams over round 1. I then took the log in order to normalize the values. The PMI value was a useful way to capture the likelihood for a given word to be referring to a certain tangram. Thus, words with very high PMI values for a given tangram were often words only used to describe that specific tangram in the given round.

Finding matches for the PMI condition

Once the PMI values for the words used to describe a given tangram were gathered I measured whether the word with the highest PMI value was used to describe that same tangram in round 6. Due to the sparseness of the data, multiple words often had the highest PMI value. To deal with this I checked if each highest PMI value word was represented in the round 6 words. I then

[1] https://en.wikipedia.org/wiki/Pointwise_mutual_information

averaged how many words were matched based on the number of highest PMI words there were. If there was a match between a highest PMI word and the words used in round 6, it received a score of 1, if there was no match, it received a 0. A final proportion was calculated by repeating this process for every tangram (12 total) and every game (24) and then dividing the total sum by 288 (the number of tangrams multiplied by the number of games). The proportion statistic was calculated to be 0.2112.

Finding matches for the null condition

To see if the PMI matching proportion was statistically significant I had to compare it to the proportion of randomly selected words on round 1 which matched with words used in round 6. To do this I went through each game and tangram on round 1, selected a random word and checked if it matched with the words used in round 6 for the same game and tangram. If it matched, it received a 1, if not it received a 0. I then calculated the proportion by dividing the sum of 0s and 1s by 288 (the number of tangrams multiplied by the number of games). To create a distribution of these proportions, I ran this process 200 times and gathered 200 proportion statistics. This null distribution is shown below in Figure 7:

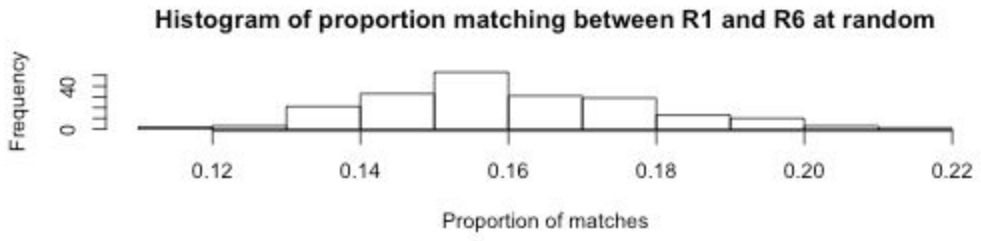


Figure 7. A histogram displaying the null distribution of matching proportions

The 95th percentile of this null distribution is 0.1913. Since the PMI matching proportion is 0.2112, it has a p-value < 0.05 under this null distribution.

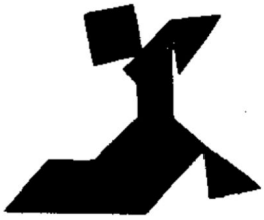
Discussion

The results support the hypothesis that words with high PMI values in round 1 are more likely to remain in round 6 than words chosen at random in round 1. This suggests that a word that initially has a high co-occurrence value with a specific tangram may be a central reason why that word gets conventionalized later on. This is a beginning step to factor the importance of context into modeling the conventionalization process. Given the fact that there are 12 different tangrams that players needed to name, the uniqueness of a given word may have been an important factor in distinguishing a certain tangram among the other 11 tangrams. Further, keeping track of what word correlates with a given shape may be a fast and low-cost way to coordinate actions. It is reasonable that we are implicitly keeping track of these probabilities by paying attention to co-occurrence in order to learn the meaning of new objects. By taking this process into account we can better understand why certain words in this game get conventionalized and others disappear.

The results of this analysis also support the notion that the processes of conventionalization can lead to many different outcomes. The hypothesis that there is a "right" or "true" phrase or name to conventionalize is unsupported by our data. Rather, we found that many different names for each tangram emerged and stabilized by round 6. For example, tangrams [C] and [D] shown below in Figure 8 were characterized with vastly different names that were used on round 6.

[C]

-angel
-ghost
-floating dude
-singer
-arms up
-falling person



[G]

-seal
-rock rolling into ditch
-yoga person
-weird one
-praying mantis

Figure 8. Images of tangram C and G and their corresponding names given by participants

Our evidence that many different types phrases can conventionalize for one tangram is supported well be the co-occurrence word theory that the current analysis tested. This co-occurrence theory supposes that conventionalization is not based on a right or true external response but based on the specific terms one group is using.

Future analyses

It should be noted that the words used in round 1 contain many stop words, including "is," "the," and "an." Given the results of prior research on this topic, it is clear that the number of words dramatically decreases from round 1 to round 6 and that nouns, verbs, and adjectives

have a much higher likelihood of being found in round 6 than do stop words. In order to account for this, a stricter and better null distribution might exclude stop words from round 1. This would help us better hone in on if PMI is good predictor of what words remain in round 6 taking into account that part of speech also tends to be predictive of the words which are conventionalized.

Another way our analysis could be improved is by cleaning up the words participants used. It was not uncommon for participants to misspell words or to use a slightly different variation of words between rounds. An example of this would be counting "iceskater," "ice skater," "iceskaterr," as the same type of word, rather than three distinct ones. In this same vein, we could do a more complicated analysis by merging different words based on their superclass. In this case, we would find the more general concept linking terms such as "flapper," "flaps," "wings," "birdman," "flying person" and tag each word with that same class. Then we could begin to model the co-occurrence of a concept given a certain shape and see if that predicted what was conventionalized.

A more complete model

A central role of AI research today is to create a computational model that learns as quickly as humans do. The first step to doing this is to understand what human learning looks like. In order to do this AI researchers have begun to explore the research done by developmental psychologists.[2] Only by examining learning in action at it's most intense peaks can we gain a clear picture of what it is we want to model. In trying to create an AI that can model how humans coordinate their actions and how they conventionalize language we also need a clear picture of

[2] To learn more view this talk given by Joshua Tenenbaum: https://www.youtube.com/watch?v=hfoeRiZU5YQ

what this process looks like. And in fact, we thankfully already know what some parts of this picture do look like. We have seen through Clark and Wilkes-Gibbs (1986) experiment and from similar results in the present study that the number of words people use and the number of exchanges that they have decreases over repeated coordination. In addition, we know that the conventionalized name becomes simpler and simpler to the point that often only nouns, verbs, and adjectives are represented. The next step is to understand why this picture looks the way it does. Clark and Wilkes-Gibbs (1986) suggested two central mechanisms that account for these patterns.

They first noted that participants strive to use minimal effort to refer to an object. While participants could have described each figure in a lengthy and well thought out paragraph, they used sparse and brief descriptions that contained the maximal amount of information in the least amount of words. We can see the simplification of terms and decrease in exchanges as the end product of repeatedly trying to minimize the effort to coordinate. The other principle force Clark and Wilkes-Gibbs describe is the principle of mutual responsibility. They suggest that conversation is successful only if each participant understands what was said and knows that both of them understand what was said. For this reason, names cannot be simplified immediately but a speaker must make sure their partner understands what they said. The principle of mutual responsibility explains why participants keep supplying additional information until their partner affirms that they understand what is being referred to. While this provides a broad scope of what goes on in the conventionalization process, co-occurrence helps describe the more fine-grained picture of what is going on in the conventionalization process--particularly, which words are

getting conventionalized. In order to build a more accurate model of what establishing references in conversation looks like we need to combine all three of these forces.

Conclusion

Through researching how statistical inference plays into the process of conventionalization we not only better understand how linguistic conventions form but also why certain aspects of our lives take on significant meaning and others do not. A prime example of this is that the most recognizable part of a human body is commonly considered the human hand. One can look at a stomach, knees, or hair on a human bodies, but the average person won't feel the same rush of emotion and knowledge associated with a human as seeing a hand. We can think of the hand as what has been conventionalized for human. Similarly with shape [C] above, you could describe the arms waving, a shirt, floating person, but ghost is what gets conventionalized. Thus given a hand we quickly infer that there is a human and given "ghost" we know that shape C is being referred to.

Given only a little bit of data, humans can make incredible inferences. However, the flip side to this is that when people are given a lot of information they may interpret it based on a few conventionalized parts rather than seeing the whole picture. It interesting to consider that even though conventions are critical for our inference they may come about through a somewhat random process of co-occurrence. From the picture of the ghost shown below, we can see how many different conventions can form given the same recurrent coordination problem.

[C]

-angel
-ghost
-floating dude
-singer
-arms up
-falling person

While this interesting, perhaps there are certain conventions to recurrent coordination problems in our societies that are more helpful and fair to us in the long run. By leaving conventionalization of to chance, we may be taking unnecessary and potentially harmful risks. Through a more mindful understanding of the conventionalization process we can be in a better position to make sure the conventions we develop are healthy and helpful to us.

References

Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary?. *Cognitive Science*, *28*(6), 937-962.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(6), 1482.

Brown‐Schmidt, S., & Tanenhaus, M. K. (2008). Real‐time investigation of referential domains in unscripted conversation: A targeted language game approach. *Cognitive science*, *32*(4), 643-684.

Centola, D., & Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, *112*(7), 1989-1994.

Christiansen, M. H., & Chater, N. (2015). The Now-or-Never Bottleneck: A Fundamental Constraint on Language. *Behavioral and Brain Sciences*, 1-52.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A.K. Joshi, I.A. Sag & B.L. Webber (eds.), Elements of discourse understanding, 10-64. Cambridge: Cambridge University Press.

Clark, H. H. (1985). Language use and language users. *Handbook of social psychology*, *2*, 179-232.

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

Cohen, P.R. (1985). The pragmatics of referring and the modality of communication. Computational Linguistics, 10, 97-146.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. Cognition, 27, 181–218.

Garrod, Simon & Doherty, Gwyneth (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. _Cognition_ 53 (3):181-21

Geisler, WS., Perry, JS., Super, BJ., Gallogly, DP. (2001) Edge co-occurrence in natural images predicts contour group performance. *Vision Res.* 2001 Mar;41(6):711-24.

Glucksberg, S., Krauss, R.M., & Higgins, E.T. (1975). The development of referential communication skills. In F.E. Horowitz (Ed.), Review of child development research. Vol. 4 (pp. 305-345). Chicago: University of Chicago Press.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation.*Journal of Memory and Language*, *49*(1), 43-61.

Hechter, M., & Opp, K. D. (Eds.). (2001). *Social norms*. Russell Sage Foundation.

Horton, W. S., & Gerrig, R. J. (2002). Speakers experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. Journal of Memory and Language, 47, 589–606.

Krauss, R.M., & Glucksberg, S. (1969). The development of communication: Competence as a function of age. Child Development, 40, 255-256.

Krauss, R.M., & Glucksberg, S. (1977). Social and nonsocial speech. Scientific American, 236, 100-105.

Krauss, R.M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. Psychonomic Science, 1, 113-114.

Krauss, R.M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. Journal of Personality and Social Psychology, 4, 343-346.

Krauss, R.M., & Weinheimer, S. (1967). Effect of referent similarity and communication mode on verbal encoding. Journal of Verbal Learning and Verbal Behavior, 6, 359-363.

Lass, R. (1997). *Historical linguistics and language change* (Vol. 81). Cambridge University Press.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41-104.

Lewis D (1969) *Convention: A Philosophical Study* (Blackwell, Oxford UK)

Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives* (Vol. 3). John Benjamins Publishing.

Linell, P. (2004). *The written language bias in linguistics: Its nature, origins and transformations*. Routledge.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(02), 169-190.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(04), 329-347.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, 361-382.

Schelling, T.C. (1960). The strategy of conflict. Cambridge, MA: Harvard University Press.

Schober, M. F. (1998). Different kinds of conversational perspective-taking. In S. R. Fussell, & R. J. Kreuz (Eds.), Social and cognitive psychological approaches to interpersonal communication. Mahwah, NJ: Lawrence Erlbaum.

Searle, J. R. (1969). Speech acts. An essay in the philosophy of language. Cambridge: Cambridge University Press

Stuhlmüller, A., & Goodman, N. D. (2012). A dynamic programming algorithm for inference in recursive probabilistic programs. *arXiv preprint arXiv:1206.3555*.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *language*, 696-735.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*(4), 289-327.

Schelling, T. C. (1960). The strategy of conflict. *Cambridge, Mass*.

Von Neumann, J. (1953). Morgenstern, 0.(1944) Theory of Games and Economic Behavior. *Princeton: Princeton UP*.

Wittgenstein L (1958) *Philosophical Investigations* (Blackwell, Oxford)

Young HP (1998) Social norms and economic welfare. *Eur Econ Rev* 42(3):821–830

Xu, F., & Tenenbaum (2007) Word learning as bayesian inference. *Psychological Review.* Vol. 114, No. 2, 245–272

Special Thanks

I could not have written this thesis without Robert Hawkins, who helped me code this experiment over the summer at the Stanford computation and cognition lab. Thank you Robert for your patience and excitement about reference games and your help with the analyses. I also am indebted to my wonderful advisor Debbie Burke. Without her kindness and openness this thesis would have never existed. In fact, I would have never majored in cognitive science if she hadn't opened her door and let me into her Intro Cognitive Science class even though it was already full and a week and a half into the semester. Thank you Debbie for your help and guidance, I look forward to framing a picture of tangrams that you can hang in your office.