Simulating the Usage Acquisition of Two-Word Sentences with a First- or Second-Person Subject and Verb

Naoya Arakawa
Dwango AI Laboratory
naoya.arakawa@nifty.com

Abstract
This article examines a minimalist mechanism in a simple language game to acquire the use of two-word sentences consisting of a first- or second-person pronoun as the subject and a verb as the predicate. In an experiment, a learner agent with minimalist architecture learned to select the subject and verb while interacting with a caretaker agent. The assumptions and implications of the experiment are discussed.

Keywords: Language acquisition, Simulation, Personal pronouns

1 Introduction

This paper shows how a simple simulated agent can learn to use two-word sentences consisting of a first- or second-person pronoun as the subject and a verb as the predicate while interacting with another agent. The environment for the simulation contains two interacting agents: a learner and a caretaker. In such situations, the learner cannot acquire the use of first- and second-person pronouns by only observing the caretaker's utterances, for with a sole speaker, the first-person subject always designates the caretaker and the second-person subject the learner. However, if both the learner and the caretaker speak, then the referents of the first- and second-person pronouns depend on the speaker and thereby the learner can acquire the use of the pronouns by observing the caretaker's reaction. Here, sentences to be used have to contain a non-personal element in order to distinguish the person, hence the use of two-word instead of one-word sentences.

The sentences function either to describe the action of the speaker or the other or to make a preemptive representation (i.e., direction) of the other's action. The caretaker reacts and issues a reward to the learner's utterance or action and takes spontaneous action. While the utterance and action of the learner are chosen randomly (i.e., babbling) at the beginning, the learner acquires the use of sentences and action as their choices are reinforced with external rewards from the caretaker and internal rewards generated when the caretaker takes expected actions.

2 Experiment

This section describes the setting and result of the experiment. The validity of the assumptions for the setting is discussed in Section 3.

2.1 The World

The world constructed in the experiment contains two rambling agents: a caretaker, which uses the language of the experiment, and a learner that learns the language. Each agent knows only its and the other's utterances and actions. Their knowledge of utterances and actions appear in symbolic forms. (Symbol grounding for categorizing exterior things is beyond the scope of this paper.) The agents take three kinds of action: {come, go, turn}.

2.2 The Language

The sentences consist of two words: a subject and a verb. The subject is chosen from four options: {first-person pronoun, second-person pronoun, name of the caretaker, name of the learner}. The verb is chosen from three options: {come, go, turn}. Both the subject and verb are chosen independently.

The caretaker is programmed to act according to the meaning posited for sentences. In particular, the subject specifies an agent and the verb describes the action of the agent specified by the subject. A sentence describes the other's action when uttered immediately after the action, in which the subject represents the other; and the speaker's action if accompanied with it, in which the subject represents the speaker. The subject representing the other is either the name of the other or a second-person pronoun, whereas the subject representing the speaker is the name of the speaker or a first-person pronoun. A sentence is considered to direct the other's action when uttered without any preceding action, in which the subject represents the other.

2.3 The Caretaker

The caretaker executes an action randomly chosen from the repertoire and describes it when the learner has not made any preceding utterance or action; otherwise, the caretaker gives the learner a direction for action randomly chosen from the repertoire. If the learner has not made preceding action but instead made an utterance on the caretaker's action, then the caretaker follows the direction. The caretaker issues a reward to (i.e., acknowledges) the learner when the learner's utterance accompanying its action is correct as the description of the action, when the learner's utterance directly following the caretaker's action is correct as its description, and when the learner's action is correct as the response to the caretaker's previous direction.

2.4 The Learner

The learner has three modes: {reaction, spontaneous action, direction}. It enters into the reaction mode when it has done nothing itself and the caretaker has made an utterance or taken action immediately beforehand. In this mode, it takes action if the caretaker has made an utterance without having taken any action and makes an utterance if the caretaker has taken an action.

The learner randomly chooses either the spontaneous action mode or the direction mode when the caretaker has neither taken any action nor made any utterance immediately beforehand. In the spontaneous action mode, it takes an action and makes an utterance. In the direction mode, it makes an utterance based on the 'internal representation' of the caretaker's action. Actions and utterances are randomly chosen at the beginning. The subject and verb are chosen independently. Choices are reinforced based on the information of the context and reward. The context of the action is its own present utterance and the caretaker's previous utterance. The context of the choice of the subject or

verb consists of the mode and either its own present action (in the spontaneous action mode), the caretaker's previous action (in the reaction mode), or the internal representation of the caretaker's action (in the direction mode). Rewards are issued externally by the caretaker or generated internally when the caretaker takes an action following the learner's direction (issued in the direction mode).

Choices are reinforced according to the calculated reward average. Namely, the probability of a choice increases as the expectation of a reward for the choice increases. The reward expectation for choices is calculated with naïve Bayes and random choices are made with the Dirichlet distribution based on estimated occurrences of rewarded choices. The initial expectation values for choices are set uniformly.

The cognitive architecture of the learner is simple: symbolic internal states (consisting of a feature list on the 'blackboard') with the current input (with another symbolic feature list) determine the subsequent internal state. The output (i.e., action or utterance) is taken from part of the internal state.

2.5 Results

Ten sets of 2,500 interactions between the caretaker and learner were performed. The success rate was the average ratio of rewarded learning module choices against 20 choices for every 200 interactions. The average and standard variation of the 10 sets were calculated. Table. 1 and Figs. 1 and 2 show the success rates of utterances, in which S represents subjects and V represents verbs in the reaction, spontaneous action, and direction modes, and action (error bars in the graphs represent ±SD). The success rates of subjects and verbs for each mode were the same, since the choice of subjects and verbs was bound to have the same reward. After 1,200 interactions, the learner learned to utter and act at a 90% rate of correctness.

Table 1: Success rates

Iteration	200	400	600	800	1000	1200	1400	1600	1800	2000	2200	2400
S/V	0.5	0.83	0.9	0.89	0.96	0.96	0.96	0.97	1	0.96	0.97	0.99
react.	±0.05	±0.01	±0.02	±0.02	±0.01	±0.01	±0.01	±0.01	±0.01	±0.01	±0.01	±0.01
S/V	0.29	0.47	0.7	0.85	0.87	0.92	0.94	0.91	0.95	0.94	0.96	0.94
sp.act.	±0.06	±0.05	±0.05	±0.02	±0.03	±0.02	±0.01	±0.02	±0.02	±0.01	±0.01	±0.01
S/V	0.34	0.52	0.71	0.86	0.86	0.93	0.93	0.93	0.95	0.95	0.97	0.95
direction	±0.04	±0.05	±0.05	±0.03	±0.01	±0.02	±0.02	±0.02	±0.01	±0.02	±0.02	±0.01
Action	0.43	0.6	0.78	0.9	0.91	0.93	0.92	0.93	0.95	0.93	0.96	0.94
	±0.05	±0.04	±0.04	±0.02	±0.03	±0.01	±0.02	±0.01	±0.02	±0.01	±0.01	±0.02

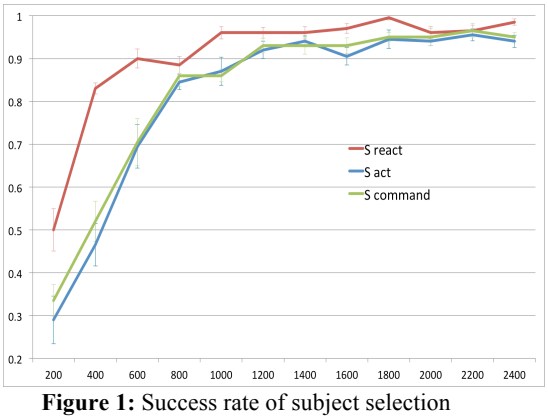

Figure 1: Success rate of subject selection

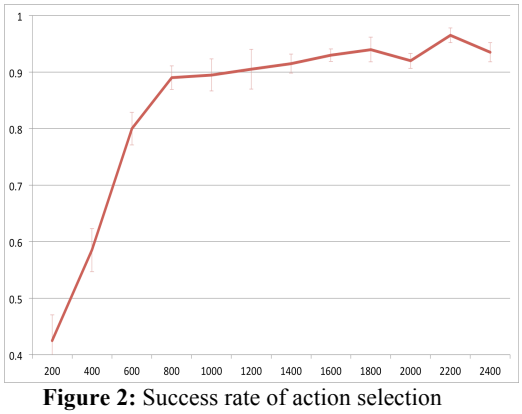

Figure 2: Success rate of action selection

The success rate of the reaction mode was better since it had more choices than the other modes (Table 2). In each mode, the learner's correctness rate reached about 90% after 1,200 trials.

Table 2: Numbers of choices

S react.	S sp.act.	S direct.	V react.	V sp.act.	V direct.	Action
654	401	393	654	401	393	345

Table 3 presents the correlations (mutual information) between context features and choices, among which irrelevant parts according to the experimental setting are marked as negative. Data were averaged from the latter half of the sets. The choice of subjects correlated highly with the mode, and apparently, the choice of verbs describing its action correlated with the action, and it did for the caretaker's action as well. The feature most correlated with choice of action was the verb in the caretaker's utterance.

Table 3: Correlation between context features and choices

Choice	Mode	Mode	Self Action	CT Action	CT Subj.	CT Verb
	react.		-	0.01	-	-
Subject	sp.act	1.34	0.25	-	-	-
	direct		-	-	-	-
	react.		-	1.48	-	-
Verb	sp.act.	0.00	1.40	-	-	-
	direct.		-	-	-	-
Action	-	-	-	-	0.01	1.27

CT: Caretaker

The subject in correct utterances was broken down to the first person 13%, the second person 39%, the learners name 13%, and the caretaker's name 35%. This reflects the breakdown of the modes in correct utterances: direction 26%, spontaneous action 26%, and reaction 48% in which only correct designation is that of the caretaker.

Sample interactions with dialogs appear in Fig.3 in the Appendix. The source code appears on GitHub at https://github.com/rondelion/Lingadrome/tree/master/Io&Tu/

3 Discussion

3.1 The World

When multiple caretakers talk to each other, learners can learn how to use first- and second-person pronouns by observing the caretakers' language use. In fact, Oshima-Takane et al. [1] and Gold and Scassellati [2] assume plural caretakers for pronominal learning. Although such might occur in human language acquisition, the current paper addresses only the case in which the learner learns a language from a single caretaker.

Agents in the experiment have no concepts of three-dimensional objects. The self-other distinction is manifest only in the attribution of utterances and actions to oneself or the other. They also do not know whether objects to which an action or utterance is attributed are the same, even if they are presumed to be so. The learner does not learn the 'concept' of the first or second person, but only learns to play a minimalist language game to generate adequate utterances containing first- and second-person pronouns and choose actions based on the caretaker's utterances.

3.2 The Language

The meaning of sentences in the experiment was programmed and expressed in the form of the caretaker's use of sentences, i.e., in the production of sentences and in responses to the learner's utterances. In the acquisition of a first language by humans, caretakers also only present the use of linguistic expressions via their utterances and actions.

The experiment uses two-word sentences, which corresponds with the period in which infants use one- or two-word sentences (see [3][4]). Here, note that actual two-word sentences do not always consist of a subject and verb. For example, in pro-drop languages, subjects do not necessarily have forms. Thus, the two-word construction in the experiment is artificial.

3.3 Learning

The experiment used approval as a reward. In the case of human language acquisition, such reward may appear as a smile signaling recognition or approval. The experiment also used internal reward when the caretaker acted in response to the learner's direction, based on the assumption that internal reward is generated when a goal is achieved.

The caretaker and learner acted randomly in the experiment. The random choice of the caretaker was necessary for statistics, whereas the random choice of action (i.e., babbling) by the learner was necessary for the reinforcement of learning.

The experiment assumed three modes for the learner: reaction, spontaneous action, and direction. Although those modes could be acquired by learning in a more radical experiment, they were assumed to be hard-wired since their acquisition was beyond the scope of the experiment.

4 Conclusion

This paper has shown how a simulated agent can learn a simple language game by interacting with another agent. At the turn of the century, researchers conducted a series of studies on language emergence with artificial agents (e.g., [5][6]), many of which explored the origin of language by making new language emerge via the interaction of artificial agents. The focus of the presented experiment, however, is the process of learning an established language from its user.

If artifacts to mimic actual human language learning process are to be constructed, unlike the ones in this rather conceptual experiment, then the knowledge in developmental linguistics would need to be referred to in greater detail; see [3][7] for overviews and [8] for infants' use of pronouns).

In the experiment, even if the types of action for the self and the other were identical, the agents did not conceive them as such. Here, since the type of action was bound to symbolic forms (e.g., "turn") in the learned language, another experiment is possible in which the learner learns to bind the same type of action using symbolic forms as common terms.

A more complex language model (LM) could be used for experiments on language acquisition more generally. The separation of grammar (LM) and semantics, as demonstrated in the current experiment, could accommodate a more complex LM, if, for example, LM is learned with methods such as LSTM while the choice of content words is learned by reinforcement or association. The forms of content words and the choice of internal representation (i.e., semantics) for the words are separated also in the Lemma Model [9] of language production. Moreover, if language is a system of choices [10], then a model in which choices are learned could be developed into a system similar to functional grammar.

References

[1] Oshima-Takane Y, Takane Y, Shultz T. The learning of first and second person pronouns in English: network models and analysis. Journal of Child Language; 1999, 26: 545–575.

[2] Gold K, Scassellati B. Grounded pronoun learning and pronoun reversal. In: The 5th International Conference on Development and Learning; 2006.

[3] O'Grady W. How Children Learn Language. Cambridge University Press; 2005.

[4] Berk S, Lillo-Martin D. The two-word stage: Motivated by linguistic or cognitive constraints? Cognitive Psychology; 2012, 65 (1): 118–140.

[5] Steels L. The Talking Head Experiment: Origins Of Words and Meanings. Language Science Press; 2015.

[6] Vogt P, Sugita Y, Tuci E, Nehaniv C, editors. Symbol Grounding and Beyond. Springer; 2006.

[7] Tomasello M. Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press; 2005.

[8] Moyer M, Harrigan K, Hacquard V, Lidz J. 2-year-olds' comprehension of personal pronouns. Online Proceedings of BUCLD 39; 2015.

[9] Levelt WJ. Producing spoken language: A blueprint of the speaker. In: The Neurocognition of Language. Oxford University Press; 1999: 83-122.

[10] Halliday M. An Introduction to Functional Grammar. Arnold; 1994.

Appendix

Early stage of learning		Immediately before 2500 interactions	
[LL Utt]: Mario vade.	"Mario goes."	[LL Utt]: Tu veni.	"You come."
[LL Act]: turn		[LL Act]:	# direction
[CT Rew]: 0		[CT Rew]: 0	
[CT Utt]: Tu gira.	"You turn."	[CT Utt]: Io veni.	"I come."
[CT Act]: go_away		[CT Act]: come	# CT reacts
[LL Rew]: 0		[LL Rew]: 1	# internal reward
[LL Utt]:		[LL Utt]: Mario veni.	"Mario comes."
[LL Act]:		[LL Act]:	# description
[CT Rew]: 0		[CT Rew]: 1	# reward
[CT Utt]: Tu vade.	"You go."	[CT Utt]:	
[CT Act]:		[CT Act]:	
[LL Rew]: 0		[LL Rew]: 0	
[LL Utt]:		[LL Utt]: Tu vade.	"You go."
[LL Act]: turn		[LL Act]:	
[CT Rew]: 0		[CT Rew]: 0	
[CT Utt]:		[CT Utt]: Mario vade.	"Mario goes"
[CT Act]:		[CT Act]: go_away	# self-description
[LL Rew]: 0		[LL Rew]: 1	# internal reward
[LL Utt]: Luca vade.	"Luca goes."	[LL Utt]: Mario vade.	# description
[LL Act]: come		[LL Act]:	
[CT Rew]: 0		[CT Rew]: 1	# reward
[CT Utt]: Tu veni.	"You come."	[CT Utt]:	

Figure 3 Example interactions
LL: Language Learner (Luca), CT: Caretaker (Mario)
Utt. Utterance, Rew.: Reward
The language for utterances is Interlingua (ia).