Artificial grammar learning of shape-based noun classification

Jennifer Culbertson (jculber4@gmu.edu)
Linguistics Program, 4400 University Drive
Fairfax, VA 22030 USA

Colin Wilson (colin@cogsci.jhu.edu)
Cognitive Science Department, 3400 N. Charles Street
Baltimore, MD 21218 USA

Abstract

Systems of noun classification serve to categorize entities based on a set of semantic and/or phonological features. Previous work, for the most part focused on gender-based classes, has suggested that learners acquiring such systems rely primarily on phonological cues, while semantic cues are used only weakly. We show, using an artificial language learning task with adults, that semantic information alone is sufficient to learn a realistic *shape-based* classification system, challenging the view of phonology bias. Further, our results show that compared to learners exposed to semantically cohesive categories, learners trained on randomly assigned classes are less successful at recalling the category of exposure items. This finding suggests that, contrary to memory-based theories of learning, categories are not necessarily formed by abstraction from memorized exemplars, but can instead be constructed from lower-level properties that category members share.

Keywords: classifiers; noun classes; language acquisition; artificial language learning; semantic features

Introduction

Systems of noun classification—such as gender, noun class, and classifier systems—distinguish or categorize objects according to salient semantic and/or phonological features. Though such systems may differ in their formal realization, the semantic features on which they are based draw from a common pool that includes physical features (e.g. shape, size), function (e.g. food, tool, habitation), as well as animacy and sociocultural status (Denny, 1976; Dixon, 1986; Lakoff, 1987; Comrie, 1989; Aikhenvald, 2000; Senft, 2000).

For example, in Cantonese, the use of a classifier morpheme is required in constructions involving a numerical or definite noun phrase, as in example (1) below.[1] The choice of classifier in Cantonese is largely determined by the head noun; for example the classifier *go[3]* is used for people, while the classifier *zek[3]* is used primarily with animals. Additional classifiers target shape properties like length, dimension, and flexibility.

(1) a. sam[1] go[3] jan[4]
three CL people
'three people'

b. sam[1] zek[3] gau[2]
three CL dogs
'three dogs'

[1] Although English does not use them productively, there are nevertheless a number of nouns which can appear with a classifier, e.g. "four strands of hair", "two sheets of paper", "a school of fish".

Similarly, in the classifier system of Navajo (Mithun, 1986) nouns are classified according to animacy and shape (among other properties); class marking in this language is found on the verb. Signed languages also commonly have noun classification systems based on shape and other functional properties (Supalla, 1986).

Acquisition of noun classification systems

Previous work on the acquisition of systems of noun classification has largely focused on genders and noun classes. Such studies have documented developmental stages including a period of phonological underspecification, and overgeneralization of frequent or default marking, and have highlighted the apparently weak role of semantic (as opposed to phonological or distributional) information (Karmiloff-Smith, 1981; Perez-Pereira, 1991; Demuth & Ellis, 2008; Mariscal, 2009; Gagliardi, 2012). The acquisition of classifier systems, although perhaps less well-studied, indicates a similar developmental trajectory. For example, Tse, Li, and Leung (2007) report that Cantonese-speaking children (3;0–5;0) tend to show early use of classifiers in required contexts but are not adult-like in their choice of classifier until quite late. In particular, children tend to over-use the classifier *go3*—used for people, but also sometimes referred to as a 'general' classifier (C. Li & Thompson, 1989)—and to over-generalize other more frequent classifiers. Although P. Li, Huang, and Hsiao (2010) show that Mandarin-speaking children generalize classifiers to novel nouns on the basis of shape features, Tsang and Chambers (2011) argue that adult speakers of Cantonese tend to rely on cues *other than* the semantic features of the nouns when processing classifiers.

In this paper we investigate the extent to which adult learners can use semantic information *alone* to acquire category distinctions instantiated in a miniature classifier system. Previous work on artificial language learning suggests that, although the population of most interest may be children within any sensitive period for language acquisition, behavioral patterns exhibited by adults can shed light on both general and language-related learning mechanisms (Wilson, 2006; Culbertson, Smolensky, & Legendre, 2012; Finley & Badecker, 2010). The motivation for using an artificial language learning task rather than natural language learning data in this case comes from our hypothesis of why it has been found that phonological cues—even when these are less statistically reliable than semantic properties—are preferentially used by learners acquiring noun classification systems (Braine, 1987; Frigo & McDonald, 1998; Gagliardi, 2012). It seems likely that children process a great deal of phonological information about dependencies between nouns and nominal modifiers (such as gender-marked determiners or classifiers) before they acquire the meanings of these elements (Polinsky & Jackson, 1999). In some sense, then, it is unsurprising that children privilege phonological information at first during language development. Adults may continue to privilege phonological cues, not because they fail to attend to semantics, but simply because their knowledge of noun classes was initially based in phonological processing.

Here, crucially, we use adult English-speakers and construct a miniature language from known objects and their linguistic labels. This removes the problem of acquiring the semantics of nouns and, if our hypothesis is correct, should expose an ability to learn cohesive noun categories on the basis of semantic features alone. While some previous work has suggested that adults can use semantic information to learn classification systems in an artificial language, these studies have exclusively focused on gender-based noun classes (Braine, 1987; Brooks, Braine, Catalano, Brody, & Sudhalter, 1993). Here we target instead shape-based classifiers, which are likely to be less familiar to English-speaking college students (the population typically targeted).

The system is modeled on Cantonese (sortal) classifiers, in particular those which pick out shape properties of objects. As mentioned above, the particular shape properties indicated by Cantonese classifiers—related to the length, flexibility, and dimensions of objects—are representative of those found in classifier systems typologically (Craig, 1986; Dixon, 1986; Comrie, 1989). Table 1 shows the two Cantonese classifiers, along with the semantic features with which they are associated, on which our system was modeled. The examples provided represent nouns which take the relevant classifier in Cantonese, and are also nouns actually used in the task.

Table 1: Shape-based classifiers tested

Classifier	Semantic features	Examples
zi[4]	rigid, narrow, long	knife, twig, candle
jeung[4]	broad, flat, flexible	sheet, card, table

Experiment 1

In Experiment 1, we tested whether adults could learn and generalize categories of nouns, distinguished by their use of the classifiers in Table 1. We compare learning of a system in which classifier use is conditioned on shape-based semantic properties of nouns to learning a random assignment of nouns to classifier categories. We hypothesized that if learners perceive and make use of semantic information in acquiring noun classification systems, they should succeed in learning the semantically-conditioned language. The random-assignment condition was used to establish an experimental baseline against which performance in the shape-based condition can be compared, and in particular to assess the role of memory for individual category members in this task. Exemplar-based models of learning argue that category formation begins with a set of memorized exemplars, abstract categories emerging later due to, e.g., computation of featural similarity among exemplars in a given category (Nosofsky, 1986). This predicts that learners exposed to conditioned and random classifier categories should perform equally well when tested on familiar items—in both cases, the set of exemplars presented during exposure should be stored—but should of course differ on their ability to generalize to novel items.

Participants

Participants were 20 native English-speaking undergraduates from the Johns Hopkins University. They received a small amount of course credit or extra credit for their participation. No subjects reported difficulties hearing or seeing the stimuli.

Materials

The miniature language was comprised of the English numeral words "one" and "two", two nonce classifier morphemes "ka" and "po", and 96 English nouns representing familiar objects. Utterances in the language consisted of a numeral word directly followed by a classifier morpheme, and a noun, as in example (2) below. Utterances were auditorily—using mac text-to-speech, speaker "Alex"—and orthographically presented and were accompanied by a visual image. The image was a single object for numeral "one" or two of the same objects for numeral "two".

(2) a. one-ka hammer
one-CL hammer

b. two-po towel
two-CL towel
'two towels'

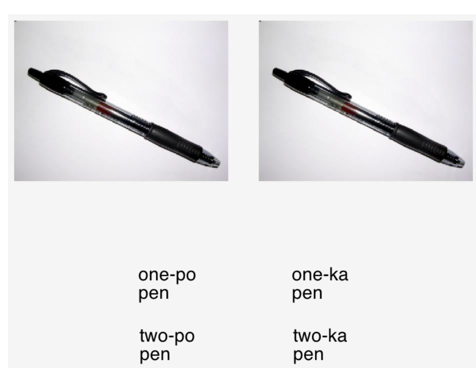

Figure 1: Example trial

Design & Procedure

Participants were seated in front of a computer, and were instructed that the task was about learning a language similar to English but with two ways of saying the words "one" and "two". They then listened to examples of "one-po", "one-ka" and "two-po", "two-ka". This was followed by 48 familiarization trials, half with objects using the classifier "ka" and half using "po". Half of the trials featured a single object and the other half two objects. On each trial, a visual image appeared with four choices below it, one for each possible numeral-classifier combination followed by the object noun pictured. Participants listened to the auditory stimulus and were required to click the choice which matched what they heard. Figure 1 shows an example trial.

After familiarization, participants took a brief break, and were then instructed that they would see a visual image and four choices below it, as in the familiarization phase, but they would hear no audio. Instead they were required to choose the phrase they thought was most likely to be used in the language. This testing phase was made up of 96 trials, including all the objects seen during familarization, and 48 novel objects. The seen objects were the same as those seen in the familiarization phase, but appeared with the other numeral (e.g. if a participant heard "one-ka hammer" and saw a single hammer during exposure, they saw two hammers at test). No feedback was given.

Participants were randomly assigned to one of two conditions. In the shape condition, the use of "ka" and "po" was conditioned on the semantic properties shown in Table 1 above. The object nouns in each class were a subset of those which actually use the corresponding classifier in Cantonese. As such, although they generally exhibited the relevant properties, there was some amount of variation in the extent to which they did so. For example, the noun "table" takes the classifier *jeung[4]* in Cantonese even though it does not perfectly exemplify the semantic features of the class.

In the random condition, the use of "ka" and "po" was unconditioned, and nouns were randomly paired with a particular classifier.

Results

In analyzing the results of this experiment we were interested in two main questions: (i) Do learners in the shape condition—in which classifier choice is determined by semantic features of nouns—succeed in learning and generalizing the correct categories? (ii) Are the categories learned those which were intended, namely the shape-based categories shown in Table 1? To address the first question, we compared first the performance on seen items across the two conditions. Performance on seen items gives an indication of how well the familiarization set was learned by a given participant. The light colored bars in Figure 2 shows proportion choice of the correct classifier on average for participants in each condition. Analysis of this data using mixed-effects logistic regression (with participants and items as random effects) reveals a significant effect of condition ($\beta = 1.47, z = 5.32, p < 0.01$), with participants in the shape condition choosing the correct classifier on seen items much more often than those in the random condition (0.86 vs. 0.45). A significant interaction between condition and number was also found ($\beta = -0.29, z = -2.63, p < 0.01$), indicating the participants in the random condition tended to be less accurate on items with the number "two" compared to "one".

We are also interesting in the extent to which participants in the shape condition could generalize the categorization information they learned during familiarization to novel (unseen) objects at test. As Figure 2 suggests, there was little difference in participants' choice of the correct classifier on seen item, and their choice of the classifier which matched the relevant semantic features on novel nouns. Analysis using mixed-effects logistic regression revealed no significant effect of item familiarity ($\beta = 0.27, z = 1.13, p = 0.26$). A significant interaction between item familiarity and number was found however ($\beta = -0.47, z = -1.98, p < 0.05$), indicating the participants tended to be less accurate on seen items with the number "one" compared to "two". Note that for participants in the random condition, there is no expected correct classifier for novel items, as the noun categories used in familiarization were random, containing no semantic cues.

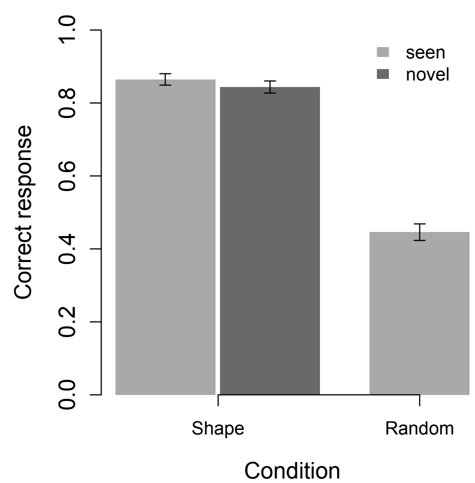

Figure 2: Correct choice of classifier for seen and novel nouns in the shape condition, and seen items in the random condition (NB: there is no correct choice for novel nouns in the random condition, since nouns were categorized randomly).

If participants in the shape condition in fact consistently inferred the same set of shape-based categories, we expect to see that their responses on novel test items are highly correlated. On the other hand, participants in the random condition were not expected to infer cohesive categories, and thus we do not expect correlated responses. To assess this, for each pair of participants in the shape condition, we computed the proportion of novel test items that they assigned to the same category. The average agreement proportion for this condition was high (0.74, SE = 0.04). In contrast, a parallel analysis revealed much lower agreement among participants in the random condition (0.50, SE = 0.02); note that 50% agreement would be expected from purely random responding.

Experiment 2

In Experiment 2, we sought to replicate our findings in a more diverse population, namely workers on Amazon Mechanical Turk (a service pairing workers with tasks over the internet). This population includes a range of ages and socio-economic backgrounds that may be more representative of the population at large (Mason & Suri, 2012). In addition, this experiment serves to add to the growing body of linguistic and cognitive research using Mechanical Turk.

Participants

Participants were 24 native English-speaking workers recruited through Amazon Mechanical Turk. They received $1.00 for their participation in the study.

Materials

The materials were the same as those used in Experiment 1, and participants were again randomly assigned to either the shape condition or the random condition.

Results

The results of Experiment 2 replicate the major findings of Experiment 1, as shown in Figure 3. Analysis of this data reveals a significant effect of condition ($\beta = 0.91, z = 3.77, p < 0.01$), with participants in the shape condition choosing the correct classifier on seen items much more often than those in the random condition (0.82 vs. 0.55). A significant interaction between condition and number was also found ($\beta = 0.17, z = 2.08, p < 0.05$), indicating that the participants in the random condition tended to be less accurate on items with the number "one" compared to "two". This interaction is in the opposite direction as what was found in Experiment 1, suggesting that the effect of number may not be reliable.

In terms of generalization to novel items, participants in the shape condition again show a relatively modest but significant increase in accuracy of classifier choice for seen items in comparison to novel items ($\beta = 0.38, z = 2.08, p < 0.05$). No other significant effects were observed, again suggesting that differences in performance driven by number in Experiment 1 may not be reliable.

As in Experiment 1, for each pair of participants in a given condition we computed the proportion of novel test items that were assigned to the same category. Average agreement was above chance for the shape condition (0.65, SE = 0.04), but note that this represents a lower level of agreement than that found in Experiment 1 for the same condition. Just as in Experiment 1, average agreement for the random condition was at the expected chance level (0.50, SE = 0.02).

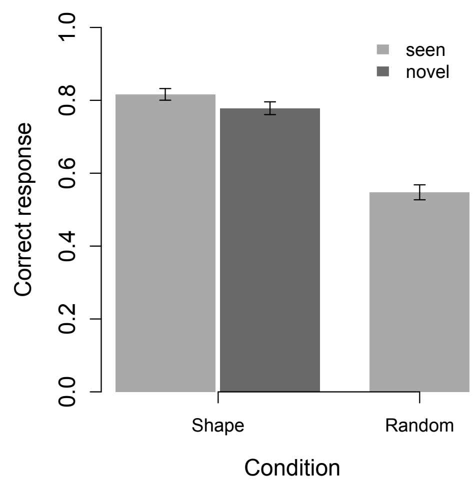

Figure 3: Correct choice of classifier for seen and novel nouns in the shape condition, and seen items in the random condition in Experiment 2 (NB: there is no correct choice for novel nouns in the random condition).

Discussion

In the experiments reported above, we exposed adult English-speakers to a miniature artificial noun classification system. In order to investigate the role of semantic features of nouns in the acquisition of classification systems, we used English words, removing an obstacle present in natural language learning. Child language learners likely go through a stage of development in which phonological but not semantic information is available for the acquisition of noun classification and other grammatical features. The results of our experiments indicate that, when exposed to a realistic classification system (based on two Cantonese sortal classifiers) over known nouns, participants are able to learn the correct categories based on semantic information alone, and can readily generalize this information to new nouns. Learning did not extend to participants exposed to randomly generated noun categories which lacked supporting semantic cues. Our findings were robust in both a population of college students, and among the more diverse population found on Amazon Mechanical Turk—despite a relatively small sample size.

This finding suggests that semantic features of nouns can be quickly used by learners as the basis of a classification system, calling into question the apparently privileged role of phonology cues argued to hold in previous work on this topic (Karmiloff-Smith, 1981; Perez-Pereira, 1991; Tsang & Chambers, 2011; Gagliardi, 2012). While here we have shown that semantically based noun classification can be learned in the absence of phonological cues, in future work we will ask whether phonological information is nevertheless used preferentially over semantic information when both are simultaneously accessible.

We believe our results are also relevant to understanding the initial stages of category formation. In particular, the dramatic difference in performance for *seen* items—items which were part of a participant's exposure set—between the two conditions calls into question theories of learning in which categories are formed by abstraction over a set of stored exemplars (Nosofsky, 1986) (see also (Rouder & Ratcliff, 2004) for relevant discussion and detailed model comparison). Under such a view, the prediction would be that learners should store the set of exemplars presented during familiarization regardless of whether the particular classifier-noun pairings are random or semantically conditioned. It would then remain unexplained why participants in the random condition fail to use the stored pairings to perform with high accuracy on seen items at test. Our participants succeeded at remembering (or reconstructing) particular examples only when those conformed to a more abstract generalization across items.

Acknowledgments

The authors would like to acknowledge Kelly Glazer for assistance in constructing the experimental materials.

References

Aikhenvald, A. (2000). *Classifiers: A typology of noun categorization devices*. New York, NY: Oxford University Press.

Braine, M. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–87). Hillsdale, NJ: Lawrence Erlbaum Associates.

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language*, *32*(1), 76 - 95.

Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. Chicago, IL: University of Chicago press.

Craig, C. (1986). *Noun classes and categorization: proceedings of a symposium on categorization and noun classification, Eugene, Oregon, october 1983*. Philadelphia, PA: John Benjamins.

Culbertson, J., Smolensky, P., & Legendre, G. (2012). Learning biases predict a word order universal. *Cognition*, *122*, 306-329.

Demuth, K., & Ellis, D. (2008). *Revisiting the acquisition of Sesotho noun class prefixes*. Mahwah, NJ: Lawrence Erlbaum.

Denny, J. (1976). What are noun classifiers good for? In *Papers from the 12th regional meeting, chicago linguistic society* (pp. 122–132).

Dixon, R. (1986). Noun classes and noun classification in typological perspective. *Noun classes and categorization*, 105–112.

Finley, S., & Badecker, W. (2010). Linguistic and non-linguistic influences on learning biases for vowel harmony. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the Cognitive Science Society* (p. 706-711). Austin, TX: Cognitive Science Society.

Frigo, L., & McDonald, J. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, *39*(2), 218–245.

Gagliardi, A. (2012). *Input and intake in language acquisition*. Unpublished doctoral dissertation, PhD dissertation, University of Maryland.

Karmiloff-Smith, A. (1981). *A functional approach to child language: A study of determiners and reference* (Vol. 24). New York, NY: Cambridge University Press.

Lakoff, G. (1987). *Women, fire and dangerous things–what categories reveal about the mind*. Chicago, IL: University of Chicago Press.

Li, C., & Thompson, S. (1989). *Mandarin Chinese: A functional reference grammar*. Los Angeles, CA: University of California press.

Li, P., Huang, B., & Hsiao, Y. (2010). Learning that classifiers count: Mandarin-speaking children's acquisition of sortal and mensural classifiers. *Journal of East Asian Linguistics*, *19*(3), 207–230.

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: starting small. *Journal of Child Language*, *36*(1), 143–171.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1–23.

Mithun, M. (1986). The convergence of noun classification systems. *Noun classes and categorization*, 379–397.

Nosofsky, R. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General*, *115*(1), 39.

Perez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language*, *18*(03), 571–590.

Polinsky, M., & Jackson, D. (1999). Noun classes: Language change and learning. In B. A. Fox, D. Jurafsky, & L. Michaelis (Eds.), *Cognition and function in language* (pp. 29–50). Stanford, CA: CSLI.

Rouder, J., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General*, *133*(1), 63.

Senft, G. (2000). *Systems of nominal classification*. Cambridge: Cambridge University Press.

Supalla, T. (1986). The classifier system in American sign language. In C. Craig (Ed.), *Noun classes and categorization* (pp. 181–214). Philadelphia, PA: John Benjamins.

Tsang, C., & Chambers, C. (2011). Appearances aren't everything: Shape classifiers and referential processing in Cantonese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1065.

Tse, S., Li, H., & Leung, S. (2007). The acquisition of Cantonese classifiers by preschool children in Hong Kong. *Journal of child language*, *34*(3), 495.

Wilson, C. (2006). An experimental and computational study of velar palatalization. *Cognitive Science*, *30*(5), 945-982.