Published in final edited form as:
Cognition. 2010 July ; 116(1): 143–148. doi:10.1016/j.cognition.2010.04.001.

Sentence Processing in an Artificial Language: Learning and Using Combinatorial Constraints

Michael S. Amato and **Maryellen C. MacDonald**
University of Wisconsin-Madison

Abstract

A study combining artificial grammar and sentence comprehension methods investigated the learning and online use of probabilistic, nonadjacent combinatorial constraints. Participants learned a small artificial language describing cartoon monsters acting on objects. Self-paced reading of sentences in the artificial language revealed comprehenders' sensitivity to nonadjacent combinatorial constraints, without explicit awareness of the probabilities embedded in the language. These results show that even newly-learned constraints have a an identifiable effect on online sentence processing. The rapidity of learning in this paradigm relative to others has implications for theories of implicit learning and its role in language acquisition.

1. Introduction

Language comprehenders rapidly weigh and integrate many partially informative sources of information when understanding language. For example, Kamide, Altmann, & Haywood (2003) found that listeners' expectations about direct object nouns were modulated by information from the conjunction of the sentence subject and verb: for a scene with a man, girl, motorcycle, and carousel, listeners' eye movements showed they rapidly anticipated the carousel given the speech context "The girl will ride the…" but anticipated the motorcycle given "The man will ride the…". While both carousels and motorcycles are plausible direct objects of the verb *ride*, their relative plausibility changes when the subject and verb are considered together. Bicknell, Elman, Hare, McRae and Kutas (2008) obtained similar results in the absence of visual contexts, using sentence reading and EEG measures.

These and many similar results support constraint-based accounts of language comprehension (MacDonald, Pearlmutter & Seidenberg, 1994; Tanenhaus & Trueswell 1995), which hold that comprehenders weigh complex long-distance combinatorial constraints in real time to interpret language. Constraints are assumed to be learned from people's prior experiences with events and language, but many studies of artificial grammar learning have suggested that constraints of this sort are extremely difficult to learn. For example, research using the serial reaction time task has shown that learning, evidenced by reaction time to more vs. less predictable patterns, declines precipitously as the number of elements required to make a prediction increases, and with the introduction of irrelevant elements in the sequence (Cleeremans & McClelland, 1991; Remillard, 2008). Similarly, learning relationships between non-adjacent speech sounds (Newport & Aslin, 2004; Gomez, 2002) or tones (Creel, Newport, & Aslin, 2004) appears to

Address Correspondence to: Maryellen C. MacDonald, Department of Psychology, University of Wisconsin-Madison, Madison, WI 53706, USA, mcmacdonald@wisc.edu.

require the presence of a perceptual or statistical grouping cue that distinguishes critical elements from interveners. These results are striking in light of the fact that long-distance, complex constraints of this sort are said to be critical in sentence comprehension, challenging both the view of statistical learning as a sufficient mechanism for much of language acquisition and also constraint-based comprehension accounts that assume that such learning has occurred.

We addressed this conflict by using a paradigm that combined artificial language learning and sentence processing methods (Wonnacott, Newport & Tanenhaus, 2008) to investigate a combinatorial constraint known to affect sentence comprehension. In our artificial language, similar to Kamide et al.'s (2003) stimuli, a particular direct object was predicted by a combination of a subject noun and verb that were non-adjacent to the object, and neither the noun nor verb alone had predictive value. We gave adult participants experience with the artificial language and a cartoon world that provided meaning to the linguistic elements. We tested participants' knowledge of the combinatorial constraints in a self-paced reading task typical of natural language comprehension studies. If adults can rapidly learn these combinatorial constraints and use them in sentence processing, then reading times on the direct object noun phrase should be shorter when that noun is predicted by the combinatorial constraints than when it is not, as in natural language reading (Bicknell et al., 2008).

2. Method

2.1. Participants

Eighty-one native English-speaking undergraduates received course credit or pay for participation.

2.2. Materials

Our artificial language contained the 19 novel words shown in Table 1. All words were phonotactically and orthotactically probable. Words referring to the (always plural) markings contained the plural suffix *-da*. Sentences had a verb-subject-object structure, with post-nominal adjectives and prepositional phrases. All sentences were six words long, taking the form *Verb monster with markings object color*.

Sentences described events depicted in 450×350 pixel bitmap images; a sentence-picture pair is shown in Figure 1 with its English gloss. Two training sets of 81 sentence-picture pairs were created, each with one cartoon monster acting on one nonsense inanimate object. Both training sets manipulated the frequency of each verb+monster+object conjunction, holding the frequencies of all other conjunctions constant, as shown in Figure 2. Specifically, in each training set, each of the three monsters, verbs, and inanimate objects appeared equally often (27 times, and each of the nine first-order conjunctions (verb+monster, monster+object, and verb+object) occurred in nine trials per training set. Second-order conjunctions of verb +monster+object were either rare, occurring once per training set, or frequent, occurring seven times per training set. For example, a *yeen* monster lifted a *sarp* object in seven trials while lifting the other two objects once each. If participants learned these distributions, they could predict the object after reading the verb and monster, even though the predictable element (*object*) was separated from the predictor elements (*verb*, *monster*) by an uninformative prepositional phrase (*with markings*). The inclusion of a color word after the object allowed for the possibility of spillover effects in reading times, in which effects of probabilistic constraints are observed one or more words after some critical word (Just, Carpenter & Woolley, 1982).

The monsters and objects were shown against one of three patterned backgrounds. Assignment of backgrounds to pictures, the color of the inanimate object (one of five colors), the patterned marking on the monster's body (one of three patterns), and whether the monster was left or right of the inanimate object (monsters always faced the object they acted upon) were all controlled to occur equally often across the training set, with no correlation with any other elements. Thus as in the real world, there was nonsystematic variation in the monster world. Some of these elements were noted in the accompanying sentences (monster markings and object color) and some were not (background and monster-object position).

Forty-five pre-training pictures were created depicting a single monster or inanimate object without action, for use in an early vocabulary-training phase. Thirty-six sentences and pictures not used in training were developed for use in a self-paced reading task. Nine sentence fragments without pictures were developed for a text-based fill-in-the-blank task, crossing each monster with each verb but omitting an inanimate object. Nine additional pictures were created for a picture based fill-in-the-blank task, again crossing each monster with each verb but not showing any objects.

2.3. Procedure

Participants completed four training blocks over two days. They were told that they would be learning a new language but received no explicit instructions about its words or structure. Each of the two training sets was used on both training days, with pictures flipped horizontally across training day (i.e. monsters facing left on day1 faced right on day2 and vice versa).

The 45-trial vocabulary pre-training block and the first 81-trial training block used a two-alternative forced choice task. On each trial, a picture was presented on screen above two sentences (or shorter phrases in the pre-training) in the monster language. One sentence correctly described the picture; the other was a foil containing a single incorrect word (e.g., the object word *gorm* when the picture contained a *sarp*) but was otherwise grammatical. Foils focused attention equally on all words of the sentence. Participants were instructed to press a key to indicate whether the sentence on the left or right of the screen described the picture. The correct answer was on the left half of the time. Participants received feedback on accuracy: following correct responses, the border around the picture turned green, the incorrect sentence was removed, and the correct sentence was positioned under the picture, remaining on screen for two seconds. Following incorrect responses, the incorrect choice was removed, the border turned red, and the correct sentence was centered under the picture for six seconds. Avoiding this lengthy feedback on error trials motivated rapid learning. The pretraining items were presented in a fixed order that gradually introduced the language words, paired with simple pictures. They were followed by the items of the training block in random order.

Following a short break, a production task was added for the second training block. On each trial, participants first saw a picture without any text and attempted to produce the artificial language sentence that correctly described it. After speaking, participants pressed the space bar, and the picture was then replaced by two sentences to choose from. Responses were entered and feedback given as in the first block. There was no feedback on participants' productions.

Participants who achieved 90% or greater sentence choice accuracy during their first session were invited to schedule another session for one to four days later. Of the 81 initial participants, 48 were invited to return for a second session. In hindsight, this criterion for continuation appeared to be overly restrictive, as it excluded some participants who made many unlucky guesses at the beginning of the first session but were highly accurate by the end.

The second session began with 20 pre-training items in the simple two alternative forced choice task, in random order, which refamiliarized participants with the vocabulary. Participants then completed the third training block, using the training set 1 materials (with pictures reversed horizontally) and the production-choice task previously used in Block 2.

The final training block was a center-presentation self-paced reading task. In each trial, the six words of a monster language sentence were presented one at a time in the center of the computer screen; a participant's keypress removed the preceding word and presented the next one. This format was used instead of the more common “moving window” paradigm (Just et al., 1982), because pre-marking word length with dashes on screen would have allowed participants to guess the next word in some cases, given the small size of the language. The center-presentation technique is generally sensitive to on-line comprehension difficulty, but it does yield more reading time spillover, where processing difficulty at a given word affects reading times on the next word (Just et al., 1982).

The keypress following the sentence-final word removed that word and displayed a picture. Participants pressed a key to indicate whether or not the image matched the sentence. Half of the trials had a matching picture; non-matching pictures differed in one or two features (incorrect verb, monster, marking, object, or object color). Feedback was provided similarly to previous blocks, with the trial's sentence displayed beneath a picture correctly depicting it, on a red or green background for 2 or 6 seconds. Thirty-six self-paced reading test trials followed the end of self-paced reading training trials without a break. The test trials occurred in random order and were indistinguishable from training items to participants.

Following the self-paced reading task, participants performed picture-based and text-based tasks designed to assess explicit knowledge of the probabilities manipulated during training. Task order was counterbalanced across participants, and trials were randomized within each task. For text-based trials, a sentence was presented in the center of the screen with a blank line in the object position, for example *Veek pim mog minada ____ skod*. Words for each of the three possible objects, *gorm*, *sarp*, *clate*, were displayed in random order at the bottom of the screen. Participants pressed a key to indicate which word best completed the sentence. Picture based trials were similar; a picture of a monster intransitively performing an action was presented in the center of the screen, pictures of the three objects were displayed below, and participants pressed a key to indicate which object best completed the scene.

3. Results

All analyses included the 48 participants who completed both sessions. Self-paced reading times for each word were analyzed with a 2 (conjunction frequency: frequent, rare) × 6 (word position) repeated measures ANOVA, and planned comparisons were carried out at words 5-6, the direct object noun phrase. Sentence-picture matching accuracy was compared with a 2 (conjunction frequency) × 2 (picture match/mismatch) ANOVA. Item (F_2) analyses with an artificial language are not interpretable, as there is no larger language to generalize to, and they are not reported here. Fill-in-the-blank responses were submitted to a chi-squared test.

For self-paced reading data (M = 983 ms, SD = 1333), all times greater than 3500ms were removed, affecting 3.70% of the data. Trials with incorrect picture matching responses were also excluded, which removed 9.62% of the data. The remaining trials were converted to length-adjusted residuals (residual RTs saved after regressing raw RTs onto subject and subject*length), a standard analysis for self-paced reading measures (Ferriera & Clifton, 1986). As can be seen in Figure 3, reading times varied through the sentence, yielding a main effect of word position, $F(5, 229) = 15.21$, $p < .001$.

The two positions at which reading times could differ as a function of conjunction frequency were at the direct object (word5) and the color adjective (word6). No difference was found at word5, $F < 1$, but participants read word6 faster when the verb+subject+object formed a frequent conjunction compared to when it formed a rare conjunction, $F(1, 47) = 8.91$, $p < .02$.

Overall accuracy in the picture-matching task was 90.38%, with greater accuracy on picture match (M = 94.91% correct, SD = 22.00%) than on non-match trials (M = 83.18% correct, SD = 37.44%), $F(1, 45) = 28.21$, $p < .001$. Importantly, participants were more accurate when the sentence had contained a frequent conjunction (M = 92.53% correct, SD = 26.31%) than a rare one (M = 88.24% correct, SD = 32.24%), $F(1, 45) = 16.56$, $p < .001$. This effect did not interact with picture match/mismatch, $F < 1$.

In the fill-in-the-blank tasks, participants' choices of the most likely direct object did not differ from chance for either picture-based or text-based versions, $\chi^2 < .1$ for both.

4. Discussion

Participants rapidly learned combinatorial, probabilistic verb+subject+object relationships, even though the subject and object were separated by an uninformative prepositional phrase. Participants brought their new knowledge to bear in reading and comprehending sentences, but explicit judgments showed no awareness of the patterns in the training set. This finding is consistent with other studies in which participants learned statistical patterns without conscious awareness (Cleermans & McClelland, 1991; Remillard, 2008). Although implicit knowledge may affect explicit judgments in some situations, such as cloze tasks, the small set of highly plausible alternative completions presented in our task likely prevented its influence here.

These results demonstrate comprehenders can rapidly learn complex probabilistic constraints that have been assumed to be at the core of constraint-based accounts of sentence processing (e.g., MacDonald et al., 1994; Tanenhaus & Trueswell, 1995), but for which learning mechanisms have rarely been discussed. The results also show that statistical learning can be measured in a task that adults do every day, reading, and which is measured in hundreds of comprehension studies. The reading times were somewhat longer but generally similar to those found in natural language, including the short reading times on the predictable preposition and the spillover effect on word6 reflecting the varying probability of the preceding word5. In contrast to some self-paced reading studies with long, highly complex sentences, there was not a large end-of-sentence wrap-up effect, in which participants have long sentence-final reading times while completing sentence interpretation and preparing for questions about the passage (Just et al., 1982). The low comprehension error rate and absence of large wrap-up effects suggest participants found it fairly natural to read at least these short sentences in a novel language. Reading time differences at word6 may reflect both spillover from integrating word5 into the ongoing sentence interpretation and also the effect of conjunction frequency on more global interpretation processes; future studies with longer materials could distinguish these effects. In either case, the results show the important effect of even newly-learned probabilistic constraints on online sentence interpretation.

These results also inform research on implicit learning, as they demonstrate rapid learning of a complex contingency that was poorly learned in prior studies. If we view our artificial language as only a sequential chain of elements, then participants can be said to have learned a fourth-order contingency, in which the object noun was predictable from an element four words back (the verb) + an element three words back (monster) + two uninformative words (the prepositional phrase). Nonadjacent fourth- and third-order contingencies are extremely difficult to learn in the SRT paradigm, requiring many more trials than in our study (e.g. Cleeremans & McClelland, 1991; Remillard, 2008). Moreover, learning even simpler nonadjacent auditory contingencies appears to require perceptual cues to group informative elements (Newport & Aslin, 2004; Creel, et al., 2004).

Several features of our stimuli and method may have made the learning task much easier than a classic fourth-order contingency. First, the printed stimuli marked word boundaries and

NIH-PA Author Manuscript

NIH-PA Author Manuscript

allowed grouping into six-word (sentence) units in which serial position of each word was informative about category membership. Second, words intervening between critical words had low within-category variability: the *–da* suffix identified monster-marking nouns, and the preposition *mog* never varied. Although Gomez (2002) found that lower variability of intervening syllables prevented learning non-adjacent relationships, in this context with clearly segmented words, low variability may have had the opposite effect by aiding interpretation of the prepositional phrase, thus facilitating learning longer distance relationships. Third, situating the language in a (cartoon) world provided meanings for the words, which likely allowed better retention and learning compared to artificial languages consisting of meaningless elements. Fourth, the visual environment may also have specifically ameliorated the effects of linguistic non-adjacency, in that the prediction-relevant elements (action, monster, object) that emerge over time in the language were all simultaneously co-present in the picture. Of course learners of natural language do not always have a similarly helpful scene, but the powerful learning effects observed here suggest that even occasional conjunctions of sequential and visual information may greatly facilitate learning complex and long-distance relationships in natural language.

Acknowledgments

This research was supported by NCHD Research Grant R01HD47425 and NICHD Training Grant T32HD049899 to the University of Wisconsin-Madison.

Appendix A: Raw self-paced reading times by conjunction frequency, outliers excluded

	word1		word2		word3		word4		word5		word6	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
rare	840	*523*	954	*598*	569	*415*	959	*667*	727	*551*	730	*546*
frequent	865	*567*	921	*615*	551	*400*	974	*686*	730	*543*	678	*457*

References

Bicknell, K.; Elman, JL.; Hare, M.; McRae, K.; Kutas, M. Online expectations for verbal arguments conditional on event knowledge. Proceedings of the 30th annual conference of the Cognitive Science Society; Washington, DC. 2008.

Cleeremans A, McClelland JL. Learning the structure of event sequences. Journal of Experimental Psychology: General 1991;120:235–253. [PubMed: 1836490]

Creel SC, Newport EL, Aslin RN. Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. Journal of Experimental Psychology: Learning, Memory, and Cognition 2004;30:1119–1130.

Ferreira F, Clifton C. The independence of syntactic processing. Journal of Memory and Language 1986;25:248–368.

Gomez RL. Variability and detection of invariant structure. Psychological Science 2002;13:431–436. [PubMed: 12219809]

Just MA, Carpenter PA, Woolley JD. Paradigms and processes in reading comprehension. Journal of Experimental Psychology: General 1982;111:228–238. [PubMed: 6213735]

Kamide Y, Altmann GTM, Haywood SL. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. Journal of Memory and Language 2003;49:133–156.

MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. Psychological Review 1994;101:676–703. [PubMed: 7984711]

Newport EL, Aslin RN. Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology 2004;48:127–162. [PubMed: 14732409]

Remillard G. Implicit learning of second-, third-, and fourth-order adjacent and nonadjacent sequential dependencies. The Quarterly Journal of Experimental Psychology 2008;61:400–424. [PubMed: 17853201]

Tanenhaus, MK.; Trueswell, JC. Sentence Comprehension. In: Eimas; Miller, editors. Handbook in Perception and Cognition, Volume 11: Speech Language and Communication. San Diego, CA: Academic Press; 1995. p. 217-262.

Wonnacott E, Newport EL, Tanenhaus MK. Acquiring and processing verb argument structure: Distributional learning in a miniature language. Cognitive Psychology 2008;56:165–209. [PubMed: 17662707]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

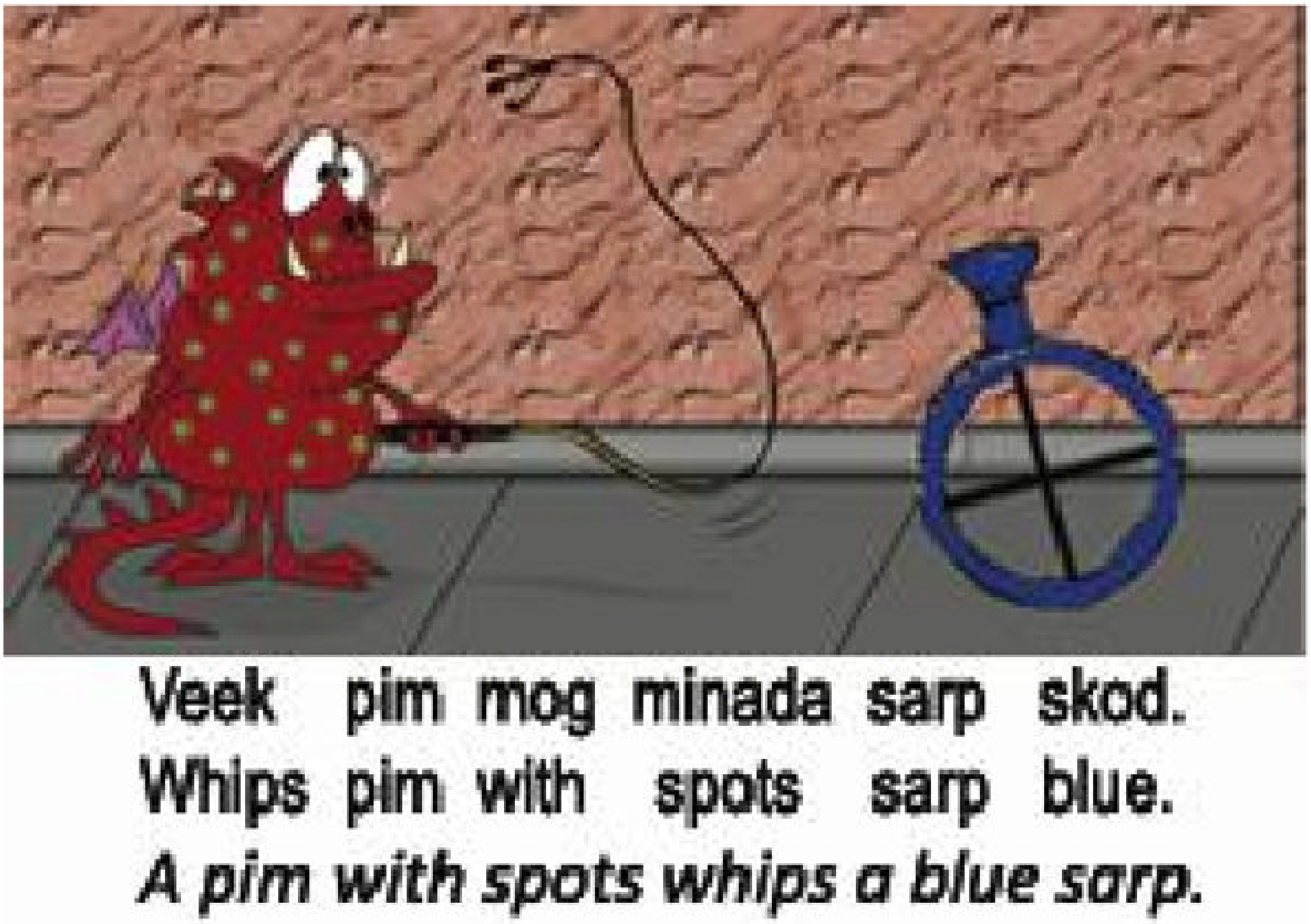

Figure 1.
Sample picture and accompanying sentence, with English gloss and translation.

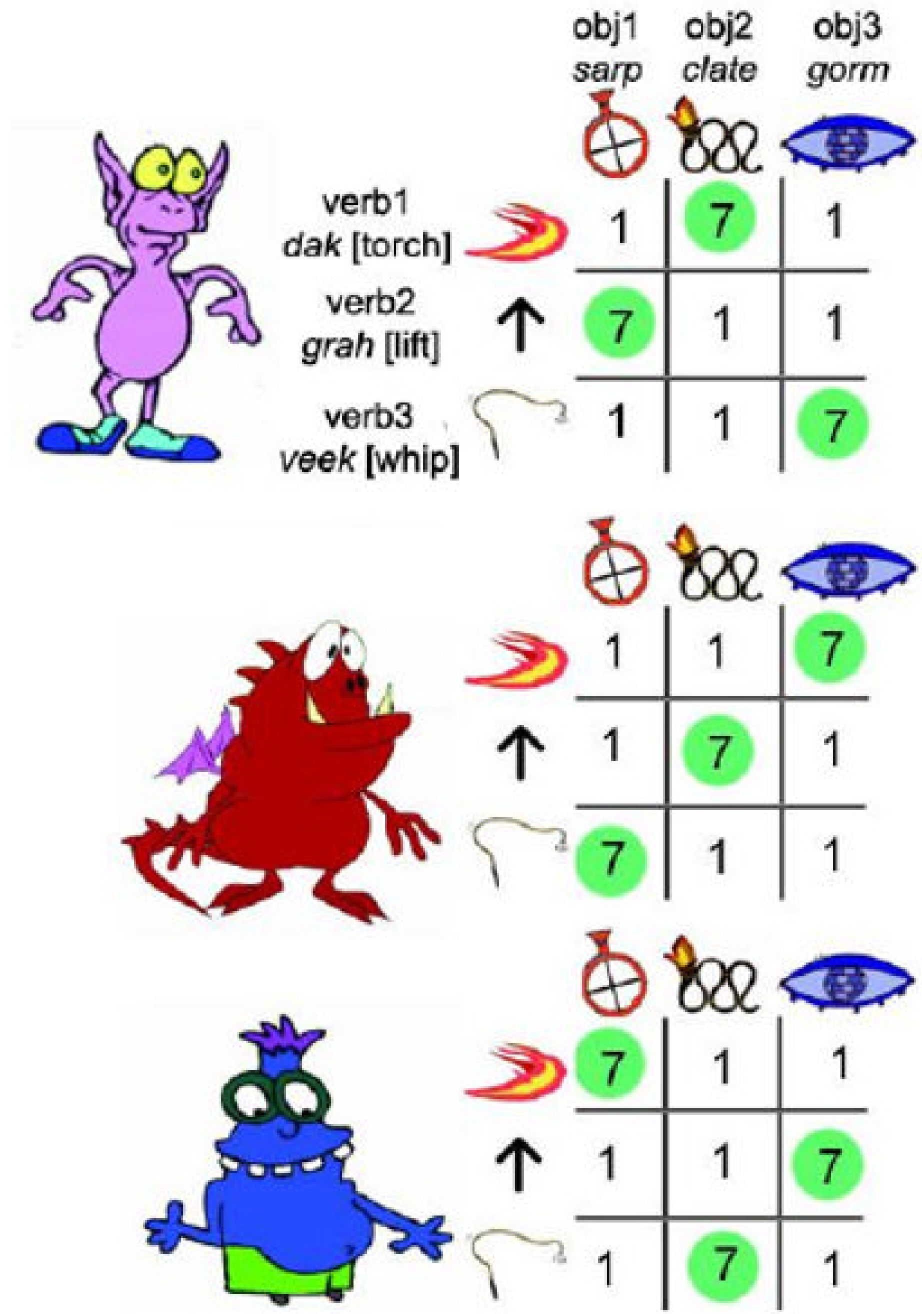

Figure 2.

Frequency in each training block of sentences containing conjunctions of the three critical elements. Monsters (yeen, pim, gled) are shown at left, objects (sarp, clate, gorm) across top of grid, and verbs (dak[torch, i.e., breathe fire on], grah[lift], veek[whip]) are nested within each monster in the grid. For example in each training set 7 sentences contained the frequent veek+pim+sarp conjunction shown in Figure 1.

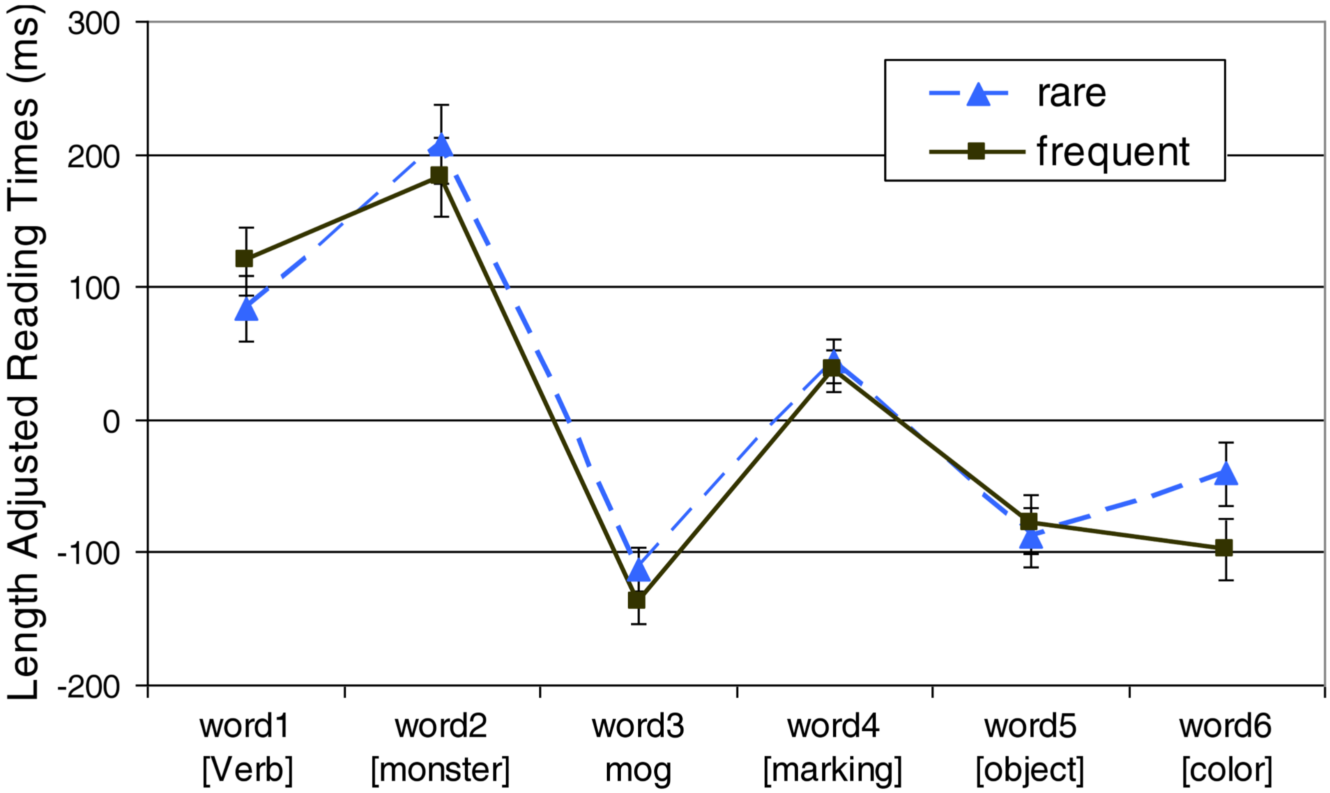

Figure 3.
Participants' mean self-paced reading times by sentence type. Word6 was read faster when the sentence contained a frequent conjunction. Bars show standard errors.

Table 1

Complete list of words in the artificial language

Monsters	Verbs	Colors
pim	*dak*	*vorg*
yeen	*grah*	*skod*
gled	*veek*	*blit*
Objects	Markings	*peka*
sarp	*minada*	*hoon*
clate	*noobda*	Preposition
gorm	*frabda*	*mog*