

CENTER FOR RESEARCH IN LANGUAGE

March 1996

Vol. 10, No. 5

The Newsletter of the Center for Research in Language, University of California, San Diego, La Jolla CA 92039
Tel: (619) 534-2536 • E-mail: crl@crl.ucsd.edu • WWW: <http://crl.ucsd.edu/newsletter.html>

• • •

FEATURE ARTICLE

Rapid Word Learning by 15-Month-Olds under Tightly Controlled Conditions

Graham Schafer and Kim Plunkett

Department of Experimental Psychology, Oxford University

ANNOUNCEMENTS

Postdoctoral Fellowship: The Center for Research in Language announces postdoctoral fellowships in Language, Communication, and Brain. The fellowships are supported by the National Institutes of Health and provide an annual stipend ranging from \$19,608 to \$32,300. Deadline for submission of applications is May 3, 1996.

WWW: The CRL Newsletter is now on World Wide Web at URL: <http://crl.ucsd.edu/newsletter.html>

EDITOR'S NOTE

This newsletter is produced and distributed by the **CENTER FOR RESEARCH IN LANGUAGE**, a research center at the University of California, San Diego that unites the efforts of fields such as Cognitive Science, Linguistics, Psychology, Computer Science, Sociology, and Philosophy, all who share an interest in language. We feature papers related to language and cognition (1-10 pages, distributed via e-mail and the World Wide Web) and welcome response from friends and colleagues at UCSD as well as other institutions.

SUBSCRIPTION INFORMATION

If you are currently receiving a hardcopy of the newsletter and have access to e-mail, please help us save printing and mailing costs by forwarding your e-mail address to CRL. If you require a hardcopy in addition, please request it and we will be happy to send you one.

If you know of others who would be interested in receiving the newsletter, please forward the e-mail or postal mailing address. Please forward correspondence to:

Jay Moody, Editor
Center for Research in Language, 0526
9500 Gilman Drive, University of California, San Diego 92093-0526
Telephone: (619) 534-2536 • E-mail: crl@crl.ucsd.edu

Back issues of this newsletter are available from CRL in hard copy as well as soft copy form. Papers featured in previous issues include the following:

Hitting the Right Pitch: A Meta Analysis of Sentence Context on Lexical Access

Mark St. John

Department of Cognitive Science, UCSD

Vol. 5, No. 6, August 1991

What is Who Violating? A Reconsideration of Linguistic Violations in Light of Event Related Potentials

Marta Kutas

Cognitive Science and Neurosciences, UCSD

Robert Kluender

Department of Linguistics, UCSD

Vol. 6, No. 1, October 1991

Learning to Recognize and Produce Words: Towards a Connectionist Model

Michael Gasser

Department of Computer Science, Indiana University

Vol. 6, No. 2, November 1991

PDP Learnability and Innate Knowledge of Language

David Kirsh

Department of Cognitive Science, UCSD

Vol. 6, No. 3, December 1991

Why It Might Pay to Assume That Languages Are Infinite

Walter J. Savitch

Department of Computer Science, UCSD

Vol. 6, No. 4, April 1992

Mental Spaces and Constructional Meaning

Claudia Brugman

Center for Research in Language, UCSD

Vol. 7, No. 1, October 1992

Is Incest Best? The Role of Pragmatic Scales and Cultural Models in Abortion Rhetoric

Seana Coulson

Department of Cognitive Science, UCSD

Vol. 7, No. 2, January 1993

Marking Oppositions in Verbal and Nominal Collectives

Suzanne Kemmer

Department of Linguistics, UCSD

Vol. 7, No. 3, September 1993

Connectionist Representations: The State of the Art

Daniel Memmi

LIMSI-CNRS

Orsay (France)

Vol. 8, No. 1, November 1993

Abstract Better Than Concrete: Implications for the Psychological and Neural Representation of Concrete Concepts

Sarah D. Breedin, Eleanor M. Saffran, and

H. Branch Coslett

Center for Cognitive Neuroscience, Department of Neurology, Temple University

Vol. 8, No. 2, April 1994

Analogic and Metaphoric Mapping in Blended Spaces: Menendez Brothers Virus

Seana Coulson

Department of Cognitive Science, UCSD

Vol. 9, No. 1, February 1995

In Search of the Statistical Brain

Javier Movellan

Department of Cognitive Science, UCSD

Vol. 9, No. 2, March 1995

Connectionist Modeling of the Fast Mapping Phenomenon

Jeanne Milostan

Computer Science and Engineering, UCSD

Vol. 9, No. 3, July 1995

Representing the Structure of a Simple Context-Free Language in a Recurrent Neural Network: A Dynamical Systems Approach

Paul Rodriguez

Department of Cognitive Science, UCSD

Vol. 10, No. 1, October 1995

A Brain Potential Whose Latency Indexes the Length and Frequency of Words

Jonathan W. King

Cognitive Science, UCSD

Marta Kutas

Cognitive Science and Neurosciences, UCSD

Vol. 10, No. 2, November 1995

Bilingual Memory: A Re-Revised Version of the Hierarchical Model of Bilingual Memory

Roberto R. Heredia

Center for Research in Language, La Jolla, CA

Vol. 10, No. 3, January 1996

Development in a Connectionist Framework: Rethinking the Nature-Nurture Debate

Kim Plunkett

Oxford University

Vol. 10, No. 4, February 1996

Rapid Word Learning by 15-Month-Olds under Tightly Controlled Conditions

Graham Schafer and Kim Plunkett
Department of Experimental Psychology
Oxford University, UK

Abstract

Infants (12.9 m to 16.8 m) were taught two novel words for two images of novel objects. Learning took place by pairing presentations of the to-be-learned auditory label with presentations of the to-be-associated image. This was followed by a period of testing using the preferential looking task in which the subject was presented with a pair of images, and an auditory stimulus. Infants took longer looks at an image if it matched the auditory stimulus than if the auditory stimulus matched the other image, or matched neither image. The design of the experiment controlled for a variety of possible confounds including pragmatic factors, contrastivity, naming effect, object and word familiarity, visual salience, side preference, auditory and phonological features of the stimulus. Subjects showed some learning after six presentations of the auditory label, and learned to distinguish between the two labels after twelve presentations. These results are discussed in the light of previous attempts to teach young children novel words, and with respect to the utility of the preferential looking task for the study of early word learning.

Introduction

In this paper we describe a method for studying the learning of novel words by 15-month-old children. We use a preferential looking task in which the child's orientation towards two visual stimuli is used to gauge comprehension of a simultaneously presented auditory stimulus. We are not the first to use preferential looking to measure children's language comprehension (Behrend 1988, Golinkoff, Hirsh-Pasek, Cauley and Gordon 1987, Naigles & Gelman 1995, Reznick 1990, Thomas, Campos, Shucard, Ramsey and Shucard 1981). However, we will show that it is possible to adapt the preferential looking task to measure children's comprehension of recently-learned novel words under carefully controlled experimental conditions.

Word learning is a complex skill involving the acquisition and integration of information across different modalities. Children must identify the concept underpinning the new word. This is a non-trivial problem: there are in principle an infinite number of targets in a single episode of ostensive naming (Quine 1960). Having somehow identified a referent concept, children must be able to map it onto some invariant aspect of the word's acoustic signal. However, there is a lack of invariance in this signal between different tokens of the same word. For example, the acoustic properties of individual phonemes vary according to adjacent and nearby phonemes (e.g., Mann and Repp 1980), speaking rate (Slowiczek and Nusbaum 1985), and identity of speaker (Lieberman, Cooper, Shankweiler & Studdert-Kennedy 1967).

This problem is compounded for the young child by the continuous nature of the speech signal: there are no reliable silences between words in speech. Once a word has been identified as an item which may stand for a concept, the child must be able to assign it a linguistic role. This occurs in a complex manner as yet not understood (Baker 1979, Pinker 1989). Failure to learn a new word may result from difficulties in mastering any of these component skills. In order to identify the factors that facilitate word learning in young children, it is important therefore to have at one's disposal an experimental methodology that can tease apart those components which may cause the process to break down.

Teaching novel words to young children confers a variety of advantages in studying the processes involved in word learning:

- The history of exposure of the child to the new word can be controlled.
- The pragmatic aspects of the learning experience can be controlled.
- The non-linguistic nature of the to-be-learned concept can be controlled.
- The learning environment can be manipulated independently of the word or concept in question.

Woodward, Markman and Fitzsimmons (1994) report recent experiments in the learning of novel words by young children, and discuss two approaches widely adopted in this field: looking, and preferential looking. Oviatt (1980) performed

experiments of the former type. The subject was trained on a novel word ('rabbit' or 'hamster') by an adult who pointed at a live animal in a cage and named it. Subjects were subsequently judged to have learnt the novel word if they looked more at the target in response to a question such as "where's the rabbit?" than in response to a control question of the form "where's the kawlow?". (The amount of looking in these conditions was compared with a baseline measure intended to control for spontaneous looking at the target). There are some problems with this design. Woodward et al. (1994) point out that the subjects in this study (Experiment 1) may have avoided false positives in the control condition simply by looking at the experimenter in confusion when they heard the nonsense word. We could also add that the use of real words and live pet animals must make any conclusive statements about the subjects' previous history with these names and to-be-learned concepts hard to substantiate.

The second technique discussed by Woodward et al. (1994) is that of preferential looking. In this technique, the child's comprehension of words is assessed in terms of her propensity to look at a visual stimulus which matches the words in question. She is simultaneously presented with an alternative visual stimulus, i.e., a distracter. Measuring the subject's responses to an array of stimuli, only one of which is a target, allows orientation away from a specific target to become part of the experiment. In contradistinction, in the looking task with a single target, looking away from the target is confounded with non-participation in the task itself. The introduction of a choice of stimuli to which the subject may respond in any given trial requires, however, that the experimenter control for the effects of relative visual salience of the stimuli. That is, the experimenter must have confidence that the looking effects of interest are mediated by the auditory stimulus and not by the visual attributes of the stimuli used.

Thomas, Campos, Shucard, Ramsey and Shucard (1981) were among the first to use preferential looking as an index of comprehension, and subsequent studies have confirmed or extended their approach (e.g., Golinkoff et al. 1987, Behrend 1988, Reznick 1990). However, these experimenters have been interested primarily in *whether* the child understands a given word, rather than in the properties of the word which caused it to be learned in the first place, or indeed the history of its acquisition. The technique described in this paper seeks to combine the insights which may be gained from teaching children new words (e.g., Lucariello 1987, Nelson and Bonvillian 1973, Oviatt 1980, Ross, Nelson, Wetstone and Tanouye 1986), with the advantages to be gained from using a closely-controlled environment for the exposure to new words and the testing of their comprehension.

There are many potential areas for confusion in the teaching and testing of novel words. Woodward et al. (1994) have shown that word learning, in the form a novel label for a novel artifact, may be effected as early as 13 months, and for

as few as nine presentations of the novel word. In their experiment, subjects played with a pair of unfamiliar objects, both of which were brought to their attention, but only one of which was named; the new label was 'toma'. Each of the two objects was brought to the child's attention nine times. The experimenters initially (Study 1) employed three types of trial. In *new label* trials subjects were presented with the two objects, and asked for the 'toma'. In *preference* trials subjects were asked to use one of the two objects, neither being named. This provided an index of non-linguistic preference for each object. In *familiar label* trials subjects' ability to understand the task was confirmed by asking them to choose between two objects whose names they knew. Whilst older subjects (18-month-olds) were able to show a systematic preference for the target object under these conditions, 13-month-old children were not. The procedure was simplified (Study 2) by dropping the preference trials, and presenting the new label and familiar label trials in blocks. Under these conditions, the subjects (32 13-month-olds) chose the target 64% of the time, a result significantly above chance.

However, the experiment is open to various criticisms. Firstly, as the authors themselves acknowledge, the visual distracters have never been named. Hence, the procedure is open to the criticism that the target was selected only because it had previously been named (cf. Baldwin & Markman 1989). Secondly, the experimenters set out to provide a control for the effects of visual salience by the use of preference trials, as described above. However, their youngest group (13-month-olds) were unable to perform significantly above chance when such preference trials were included. It is also interesting to note that when trials were presented in a between-subjects design, the 13-month-olds performed significantly above chance in selecting the target object, while the 18-month-olds inexplicably did not.

In the following experiment, we use a preferential looking task which balances frequency of presentation of target and distracter images, and frequency of presentation of target and distracter words, and introduces a variety of additional controls which may enable researchers to tease apart the various factors influencing learnability of wugs¹ by young children. Below we list the major problems confronting preferential looking studies and which we believe our procedure manages to avoid.

Pragmatic factors

Baldwin (1993) has shown that the extent to which 16- and 17-month-old children share attention with the instructor can determine whether new labels are learned or not. Because one

¹ In the remainder of this paper, in recognition of the work of Jean Berko (Berko 1958) we will refer to the to-be-learned novel word as a 'wug'. This reflects its status as something new to the child which may appear to do symbolic work.

cannot easily constrain the child's attention, a paradigm which does not involve the physical presence of an instructor is to be preferred.

The contrastivity trap.

Barrett (1978) and Clark (1987) have emphasised the so-called Principle of Contrast in lexical acquisition. That is, every word form contrasts in meaning with every other, and language learners may exploit this (but see Gathercole 1987 for an alternative view). To ensure that the subject is attending to a given image *because* that image is a target (i.e., matches the auditory stimulus) it is necessary to exclude the possibility that she attends because she recognises that a *distracter* does *not* match the auditory stimulus.

Naming effects.

Problems may arise if one item has a label and one does not. A 31-month-old child presented with two objects, one of which has a label, will assume that any new label refers to the unnamed object (Golinkoff, Hirsh-Pasek, Bailey and Wenger 1992). In addition, it is important to control the frequency of naming of the distracter during training, in order to avoid the problem that the target is selected only because it has previously been named. Baldwin and Markman (1989) have shown that a novel object which is being named is more likely to be looked at subsequently by 10- to 14-month-old infants. The only sure way to avoid the effects of differential repetition is to equate the number of times that the distracter is named during training with the number of times that the target is named during training.

Use of real objects.

Most experimenters (e.g., Nelson & Bonvillian 1973, Oviatt 1980) have used objects from the home, or living animals, as 'concept items' upon which to map new words. This approach is subject to the criticism that subjects may have seen the items, or similar items, or pictures of similar items, before: the object-name mapping may be distorted by past experience.

Use of familiar words.

A similar objection arises with the use of adult English words. For example, Oviatt's (1980) suggestion that her 10, 13, and 16-month-old subjects had not encountered the words 'rabbit' or 'hamster' in their lives before is open to question.

Control of relative salience of the visual stimuli.

If preferential looking is to be used as a measure of comprehension, any within-subject or between-subject bias to a given visual stimulus may contaminate the effect. Hence, the experiment must incorporate controls for relative perceptual salience of an image. Reznick (1990) addressed this issue by adopting a measure in which percentage fixation on the target during the period before onset of the auditory stimulus was subtracted from percentage fixation time on the target when the subject was instructed to look at the target. Unfortunately, this design fails to control for the 'contrastivity

trap' outlined above, i.e., the subject may orient towards the target image only because she knows the distracter doesn't match the auditory stimulus. In contrast, Behrend (1988) and Thomas et al. (1981) compared looking times at the target object when it was named with looking times when a neutral word was presented. Comprehension was indexed by the difference between the two looking times. This approach controls for the relative visual salience of the stimuli but similarly falls into the contrastivity trap. Only a design which employs objects as targets in one trial and as distracters in another will robustly and simultaneously avoid the contrastivity trap and the problem of relative visual salience. This approach has been successfully adopted by Naigles & Gelman (1995), who used words known to the child in an investigation of overextensions.

Side Preference.

Mount, Reznick, Kagan, Hiatt & Szpak (1989) studied infants' responses to pairs of side-by-side identical pictures, and reported that gaze direction is increasingly asymmetric from 13 to 20-months, being biased to the right hemifield. If, as is likely, the effect is a correlate of increasingly sophisticated linguistic performance, and given the need to entertain individual differences, it makes sense to control for a pure, within-subject, bias to side.

Control of auditory stimulus.

In our experiment we wished to investigate the role of phonemic features in lexical uptake. In this case we felt it important to maintain a consistent acoustic stimulus between subjects. In normal speech there is no one-to-one mapping from acoustic signal to phonemic representation (the so-called *invariance* problem). There are also individual differences in the use of motherese. Procedures which use variable tokens of the mother's voice when measuring comprehension (e.g. Behrend 1988, Reznick 1990, Thomas et al. 1981) therefore had to be avoided.

Phonological features of the auditory stimulus.

The phonemic content of the auditory stimulus may also be important in determining performance in the preferential looking task. For example, Vihman, Ferguson and Elbert (1986) have shown that words containing stop consonants are used earlier than words containing fricatives and liquids. Developmental patterns in the child's productive phonological repertoire may reflect the child's developing perceptual sensitivity to phonemic contrasts (Stoel-Gammon and Cooper 1984). It is important therefore to take care that the auditory stimuli that are used to name novel objects are within the scope of the child's contrastive phonological repertoire.

The preferential looking task described below avoids both the contrastivity trap and the naming effect, by balancing the frequency of presentation, and frequency of naming, of the visual stimuli. The procedure minimises the influence of pragmatic factors. It presents genuinely novel stimuli, and

controls for individual subject's preference for specific visual stimuli. Bias to side is controlled for, and variability in the auditory stimuli eliminated.

Method

Preferential Looking.

In this technique, the subject is presented with two visual stimuli and an auditory stimulus. The relation of the auditory stimulus to the visual stimuli is varied between trials, so that the amount of looking at a visual target may be compared in several conditions:

1. A condition in which the auditory stimulus matches the visual target. We refer to this as the MATCH condition.
2. A condition in which the auditory stimulus conveys no information about either of the images. We refer to this as the NEUTRAL condition.
3. A condition in which the auditory stimulus matches the distracter. We refer to this as the ANTI-MATCH condition.

Of course, the ANTI-MATCH condition on one side corresponds to a MATCH condition on the other; we analysed looking towards left and right images separately in order to control for any side bias effects.

Subjects

29 subjects, of mean age 14.8 months (max. 16.8, min. 12.9) participated in the study. There were 15 boys and 14 girls. All were full-term and in good health. All subjects had learned English in the UK from parents for whom English was the first language. None of the subjects had been exposed in the home to languages other than English.

Seven subjects did not complete all experimental trials and were therefore dropped from the analysis. Video-tapes of all those who finished the experiment went on to the blind scoring stage. The average age of these remaining 22 subjects was 14.7 months (there was no effect of age in whether the experiment was completed or not).

Stimuli

We used two non-words consisting of phonotactically legal consonant-vowel-consonant (CVC) strings. This selection was based on the finding (Charles-Luce & Luce 1990) that CVC strings represent the most commonly-occurring word types in the young (five year-old) English speaker's mental lexicon.

In order to investigate whether words containing stop consonants are more readily learned by the child than words containing fricatives and liquids, the two wugs² were chosen

to contain phonemes drawn from these categories. Both wugs had the same central vowel. We used the wugs /bA:d/ and /sA:l/ ("bard" and "sarl"), these being CVCs outside the child's linguistic experience. The phonemes /b/ and /d/ in general emerge earlier than /s/ and /l/ (Vihman et al. 1986). In addition, we required a non-word which would act as a non-informative, or neutral, stimulus contrasting with the wugs used. We used the CVC /gi:k/ ("geek"). This contrasted both in terms of the central vowel (/A/ is low and posterior, whereas /i/ is high and anterior) and in terms of the initial and final consonants (/b/, /d/, /s/, and /l/ are +[anterior], whereas /g/ and /k/ are -[anterior]). The stimuli were recorded as single wugs by a female voice. Stimuli were digitally recorded at 22.05kHz into signed, 16-bit files. Each sample was edited to remove any head and tail clicks, then matched for length and scaled so that maximum peak-to-peak amplitude was the same for all samples.

The visual stimuli were produced by editing images from a CD-ROM children's picture-dictionary. Picture-editing software was used to generate five 320x200 pixel 256-colour pictures, each showing a single 'nonsense object'. We made some attempt to make the pictures of approximately equal visual interest. Each picture had at least two spectral colours in it, had two textures on the object surface, and portrayed depth in some way (i.e., there was some shadow, or one part of the object occluded another part). The general effect of each image was that of a photograph of a single rather strange artifact presented against a white background.

Since the purpose of the experiment was to investigate the learning of auditory labels, rather than visual images, the image associated with each auditory label during the training phase (see below) was systematically varied. Each time the experiment was run, two images were selected from the five images available. Each of the five images was, as far as possible, paired an equal number of times with each auditory label.

Measured variable

The preferential looking test admits a variety of candidate measured variables. These include measures of the total looking at an image, expressed in absolute terms (e.g., Thomas et al. 1981) or proportional terms (e.g. Reznick 1990); and measures of the duration of first look at an image (e.g., Fernald, McRoberts and Herrera, In press). Naigles & Gelman (1995) used both types of measure. We propose to add a further measure: the duration of longest look. Our reasoning is that measuring total looking at two or more targets suffers from two shortcomings: excessive noise and a decreasing effect of subject participation within a given trial. In the case of the former problem, rapid glances are difficult to code accurately. In the latter case, should the subject tend to behave more randomly as the trial proceeds, then any effect of target may be 'washed out'. As regards measuring the length of the first look at an image: This appears more

² Where a token is specifically a spoken instance, it will appear in italics: *wug*. Where a token is specifically a visual presentation, it will appear in bold type: **wug**.

suitable to experiments where the subject is directed to look at a target and knows where that target is to be found (e.g., Naigles & Gelman 1995). For the presentation of single words to young subjects we preferred a measure which did not require that the subject comprehend an instruction. Perhaps most importantly, we selected the duration of longest look as a measure because it had suggested itself to us during pilot work: We had the intuition that there was a qualitative difference in the looks occurring during certain trials, and that this might be reflected in the lengths of the looks themselves.

Procedure

Subjects were seated on their caregiver's lap, facing two eye-level monitors at a distance of approximately 80cm. The screens were placed 44cm apart, centre to centre. Each screen measured 30cm across the diagonal. A loudspeaker, located centrally and above the monitors, delivered the auditory stimuli. A small red LED and a buzzer mounted between the monitors allowed the experimenter to attract the subject's attention and to re-fixate her gaze centrally between trials.

The subject's responses were recorded by hidden video cameras positioned just above each of the two monitors. A third camera, mounted centrally, allowed assessment of the subject's position and angle of gaze relative to the centre-line. Trials were launched individually, when the experimenter judged the subject to be fixating centrally. Order of trials was determined by the computer at run-time and the experimenter was blind to the trial type being launched. The experimenter was invisible to the child throughout the procedure.

There were two phases in the experiment. In a training phase, the subject experienced a sequence of training trials. Each training trial consisted of presentation of one of two auditory stimuli, together with the corresponding image on one of the monitors. The other monitor remained blank. In a testing phase, the subject was presented with a spoken sound stimulus together with a pair of images. The sound stimulus usually, but not always, corresponded to one of the images. The extent to which the subject oriented to an image *given that it corresponded to the speech signal* formed our index of word recognition.

The experiment consisted of a pair of introductory trials followed by two experimental blocks, each block consisting of a training phase followed by a testing phase.

Introductory Trials

At the outset of the experiment the subject was presented (on one of the two monitors) with an image of a **shoe** and then an image of a **cup**, each paired with its auditory label. The order of this initial pair of presentations was varied randomly between subjects. We used **cup** and **shoe** because these are among the earliest words to be acquired by the child, (Fenson, Dale, Reznick, Bates, Thal and Pethick 1994). One of these events occurred on the left monitor, and one on the right, determined randomly by the computer. The purpose of

presenting a pair of real words was to alert the subject to the idea that ostensive naming was occurring (i.e., that the auditory labels were meaningful, matching the image).

Training Phase

After the initial two trials, the first training phase began. Both **wug/wug** pairs (i.e., **bard/bard** and **sarl/sarl**) were trained in the course of a training phase. Each trial consisted of the auditory presentation of one of the *wugs*, accompanied by the appearance of the corresponding colour image on one of the two monitors. Side and order of presentation of stimuli were pseudo-randomly determined by the computer, such that each **wug** image was paired with its corresponding *wug* auditory label three times on the left and three times on the right. Trials in which the same stimulus occurred repeatedly, on the same or different sides, were allowed. Each trial comprised of a single spoken instance of the auditory label; hence a given *wug* was heard 6 times during each training phase. During each training phase the image/label pairs **cup/cup** and **shoe/shoe** were each presented once more to the subject, at a random point during the sequence to underline the procedure of ostensive naming and to re-awaken interest in the stimuli. There were therefore a total of 14 training trials per training phase.

The caregiver was instructed to sit quietly and to listen to instructions played over headphones. The instructions were recorded by the same voice used for the auditory stimuli and were accompanied by white noise. The caregiver could not discern the auditory stimulus being used. She was instructed to look upwards, away from the monitors, thereby minimising the likelihood of influencing the subject's behaviour.

Testing Phase

There were six trials in the testing phase. Each of the three auditory stimuli (*bard*, *sarl* and *geek*) was presented in combination with each of the two possible positions of the two images (i.e., **bard** on the left monitor, **sarl** on the right monitor; or **sarl** on the left monitor, **bard** on the right monitor). On each of the six testing trials, therefore, the same two images were presented: all that varied was the auditory stimulus and the location of the **bard** and **sarl**. Trials were ordered pseudo-randomly by the computer. In order to begin a trial the subject was required to fixate on the light/buzzer display situated centrally between the two monitors. The images appeared on the two monitors, without presentation of the auditory stimulus, for a duration of 2960ms. The auditory stimulus was then presented three times over a period of 7030ms, as a single word with silences between presentations. Throughout this period, the monitors continued to display the **wug** images. The subject's responses were recorded by the video cameras placed above each monitor. Signals from the two cameras were routed via a digital splitter to a VCR which recorded two separate time-locked images of the child onto a single tape.

Table 1: Trial Types

Auditory Stimulus	Left Image	Right Image	Trial Type of Looks at Left	Trial Type of Looks at Right
<i>bard</i>	bard	sarl	MATCH	ANTI-MATCH
<i>bard</i>	sarl	bard	ANTI-MATCH	MATCH
<i>geek</i>	bard	sarl	NEUTRAL	NEUTRAL
<i>geek</i>	sarl	bard	NEUTRAL	NEUTRAL
<i>sarl</i>	bard	sarl	ANTI-MATCH	MATCH
<i>sarl</i>	sarl	bard	MATCH	ANTI-MATCH

After the first testing phase was over, the computer presented a second training phase, with trials presented in a different order from the first training phase. The testing phase was then repeated, again in a new pseudo-random order. With minimal interruption, the entire procedure lasted around five minutes.

Results

As discussed in the Method section, pilot work had led us to surmise that duration of the longest look at a supposed target would prove an effective measured variable, indexing the subject's knowledge that the auditory label matched the image.

Video-tapes of the testing phases were analysed after each experimental session. A button-press apparatus was used to create a file tabulating the time-course of looks to each monitor. Each tape was observed four times, twice to record durations of looks to the monitor on the child's left, and twice to record the durations of looks to the monitor on the child's right. These data were averaged to give the mean longest look per side per trial. Scoring was principally done by the first author; 5% of the recordings were checked for reliability as described below.

There were twelve test trials per subject (two testing phases, of six trials each). Trials were deemed successful if the subject looked at both images in the course of that trial (i.e., during the initial 2960ms and/or the subsequent 7030ms). In this way we could be sure that the subject was aware, at some point during the trial, of the locations of both images.

Recordings scored by the second experimenter were treated similarly. Reliability was assessed by estimating the intra-class correlation coefficient for the total set of longest looks for each experimenter. This yielded an estimate of R of 1.0.³

³ This may seem rather high for a reliability score. The intra-class correlation coefficient is a measure of the ratio of variance in the treatment data, to total variance, when two people do the scoring independently. A value close to unity reflects the relative ease with which

The design was fully counterbalanced between three conditions. With respect to a given side, two trials measured the MATCH effect, where looking at that side was directed at an image which matched the auditory label; two trials measured the NEUTRAL effect, where looking at that side was directed at an image which was neutral with respect to the auditory stimulus; and two trials measured the ANTI-MATCH effect, where looking at that side was directed at one image, whilst the image which matched the auditory label was being displayed on the other monitor. This analysis is summarised in Table 1.

There were up to four data points per condition per subject per side (two wugs x two blocks). These were averaged, to give one figure for the longest look per successful trial per condition per subject per side. With this procedure there were no missing values. Data for each subject in each condition were then averaged between sides. The resulting mean longest look is given in Table 2 as a function of Trial Type, and displayed graphically in Figure 1.

Table 2: Longest Looks

Trial Type	Mean Longest Look
MATCH	2380
NEUTRAL	1975
ANTI-MATCH	1775

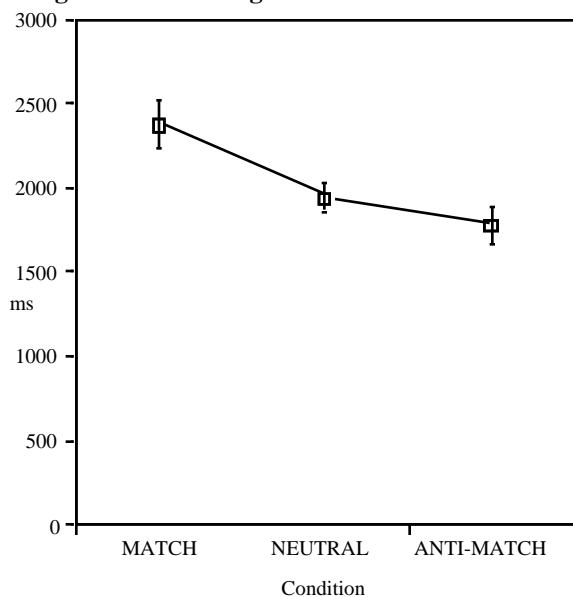
Detailed inspection of the data revealed that it departed from the normal distribution, with skew and kurtosis exceeding their own standard errors (Quenouille 1966). This was probably due to data originating as timed intervals starting at zero (Winer, Brown and Michels 1991). The data were log-transformed, bringing skew and kurtosis within the range for normality.

The *a priori* hypothesis was that longer looks would be made to an image in the case where the auditory label matched that image rather than in cases where it matched neither image, or in the case where it matched the image on

longest looks are measured combined with the high variability, between trials, of the measured variable.

the other monitor. That is to say, we predicted that the mean longest look in the MATCH condition would be longer than the mean longest look in the NEUTRAL or ANTI-MATCH conditions. Planned comparisons were therefore carried out on the log-transformed data using the Dunn-Sidak procedure for non-orthogonal *a priori* contrasts (Kirk 1982). The mean longest look in the MATCH condition was longer than those in either the NEUTRAL or the ANTI-MATCH conditions ($t_{DS} = 2.57$, $p(\text{one-tailed}) < 0.025$; $t_{DS} = 4.06$, $p(\text{one-tailed}) < 0.005$ respectively). Similar results were obtained with the original, untransformed data.

Figure 1: Mean longest looks in each condition

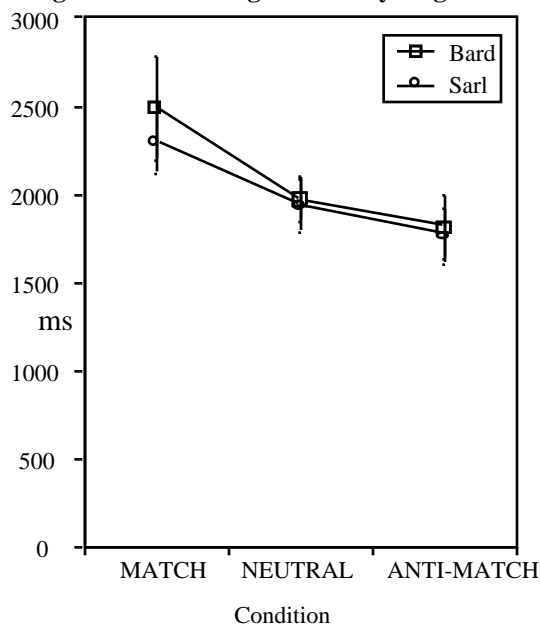


The above result is enough to establish that the subjects' responses to the images are mediated by their previous experience with the auditory labels. However, it is not clear from this analysis which wug has been learned. It might be that knowledge of one wug is enough to drive the effect. Or it might be that knowledge about the *pair* of wugs was somehow being used by the subjects. The data were therefore reanalysed with an additional factor: the previously-associated auditory label for the image in the monitor. We termed this the IMAGE condition. The data are plotted in Figure 2.

To investigate the effect of the IMAGE condition, we carried out four planned contrasts on the log-transformed data using the Dunn-Sidak procedure on this expanded set of conditions. We had made the *a priori* hypotheses that, in both of the IMAGE conditions, the mean longest look in the MATCH condition would be longer than the mean longest look in the NEUTRAL or ANTI-MATCH conditions. In the case of the MATCH-NEUTRAL contrast the null hypothesis could not be rejected, but the MATCH-

ANTI-MATCH contrast was significant in the case of both wugs ($t_{DS} = 2.57$ for **bard**, $t_{DS} = 2.49$ for **sarl**, $p(\text{one-tailed}) < 0.05$ in both cases). Similar results were obtained using the untransformed data. This analysis demonstrated that our subjects distinguished the two wug labels presented during the training phase, mapping the labels to appropriate representations of the images.

Figure 2: Mean longest looks by target label



It remained to investigate whether other factors had mediated the subjects' responses. We had set out to investigate, *inter alia*, whether the phonemic features of a novel label would influence its uptake by the child. Visual inspection of Figure 2 does not support a difference in comprehension for the auditory label *bard* and the auditory label *sarl*. Other factors which might have influenced the subjects' responses were the SIDE to which the subject was orienting and the BLOCK in which the measurement was made. With regard to the SIDE condition, Mount et al. (1989) have reported that gaze direction is increasingly asymmetric between the ages of 13 and 20 months, being biased to the right hemifield. In the case of the BLOCK condition, a propensity only to orient to a target in the second block would demonstrate that six exposures to the image-label pair were insufficient to effect learning. A four-way repeated measures ANOVA was carried out on the longest look data from the original trials, log-transformed. There were some trials (8% of the total) which did not meet the criterion that the subject look at both images at some point during the trial (i.e., during the initial 2960ms and/or the subsequent 7030ms). The measured variable in these trials was replaced with the mean for that condition on that side, as measured across all subjects. The

conditions were TRIALTYPE (3 levels), IMAGE (2 levels), SIDE (2 levels), and BLOCK (2 levels). As expected, there was a strong effect of TRIALTYPE, $F(2,42) = 4.65$, $p = 0.015$, but no other significant effects or interactions. Similar results were obtained using untransformed data. This lack of interactions, taken together with visual inspection of Figure 2 (wherein the main effect changes across trial type but not across label), is evidence against a *bard/sarl* difference but strongly supports an interpretation that the subject takes longer looks at an image when she hears its recently-learned label.

Discussion

We set out to design a task which would provide a flexible yet tightly-controlled framework for the investigation of novel word learning. In the experiment described we have demonstrated the rapid learning of novel words for novel objects, without the intervention of a human instructor. Conditions were controlled to exclude the pitfalls outlined in the introduction. Our subjects took longer looks at a visual target if that target matched the previously-trained auditory label (i.e., in the MATCH condition). Looks were longer in the MATCH condition than in either of the control conditions. We believe this to be the first demonstration of this kind of learning by young children in such tightly-controlled conditions.

The finding of Woodward et al. (1994) that 13-month-olds can in favourable circumstances learn novel words from as few as nine instances is of considerable interest. In common with the participants in our study, Woodward et al.'s subjects were 'pre-naming explosion', or 'pre-vocabulary spurt'. That is to say, they were below the age at which children begin to show a marked increase in the rate of addition to their productive vocabularies. This age is generally taken to be 18 months (Dromi 1986, Nelson 1973). Woodward et al. discuss the vocabulary spurt and three families of theories as to the mechanisms underlying it: linguistic development (e.g., Dore 1978, Lock 1980, Plunkett 1993), conceptual development (e.g., Corrigan 1978, Gopnik & Meltzoff 1986), and the advent of constraints on word learning (e.g., Behrend 1990, Markman 1991). Whichever theoretical position is adopted, Woodward et al. characterise the vocabulary spurt within the evolving acquisition of language as follows:

In summary, these explanations imply that before the naming explosion and the insights or cognitive milestones that lead to it, learning a single new word would be a time-consuming process, requiring much exposure to the new word.[...] At the time of the naming explosion, it is argued, children become efficient word learners, capable of learning new words after

only limited exposure to them. (Woodward, Markman and Fitzsimmons (1994) p.554)

According to this interpretation of the field, a pre-vocabulary spurt child should have difficulty in making fast mappings between a new word and its referent. However, Markman and her colleagues have provided evidence that the pre-vocabulary spurt child is indeed capable of learning new mappings from limited exposure to the word and its referent. Our results support their findings.

How many exposures to the new word are necessary for the learning of novel words? The subjects in the Woodward et al. (1994) study showed comprehension after nine exposures to the new label. In our study, subjects experienced each label/image pair twelve times: six times in each training block. (Of course, they heard each auditory label a further four times during the two testing phases). Subjects were tested between blocks. We have shown that preferential looking towards a target picture (the MATCH condition) is already established by the first training phase. There was no effect of BLOCK on the amount of preferential looking towards a target picture. In other words, six presentations of each auditory label were sufficient to bring about some learning.

Did the subjects in our experiment learn *both* new label-object mappings? Across subjects, on average, they did. This can be seen in their responses to specific images in each type of trial, illustrated in Figure 2. Subjects took longer looks in the MATCH condition than in the ANTI-MATCH or NEUTRAL conditions. The difference between the MATCH and ANTI-MATCH conditions was significant. This was true for both **wug** tokens. Subjects did this regardless of the side on which the target was presented, and independently of the actual images used, since these were varied between subjects.

It might be argued that a long look to one image will imply a short look at the other. To investigate this, we examined the temporal patterns of the subjects' responses. Longest looks which exceeded 3500ms, i.e., half the available time for response to the opposite side, occurred in only 13% of trials. Furthermore, there were an average of 2.2 looks to each side during each trial, and this figure did not vary significantly with trial type. We can therefore conclude that the likelihood of ceiling effects appearing disproportionately in ANTI-MATCH condition trials is very low. Nevertheless, the conclusion that two novel words, rather than just one, have been learned would be more secure if the longest look in the MATCH condition were significantly longer than that in the NEUTRAL condition. The trend, however, is in the predicted direction.

The design of the experiment called for two labels to be trained and tested, for reasons of control already discussed. However there is a theoretical interest in training and testing two novel labels rather than just one as is often

the case (e.g. Oviatt 1980, Woodward et al. 1994). In the case where there is only one label under test, no *discrimination* at the auditory level is required to solve the problem. Even if the subject recognised the auditory label she only had to recognise it in the form 'recent label' rather than by recognition of any of its auditory features. In our experiment, the two labels shared a central vowel, and were distinguished only by initial and final consonants. Subjects had to discriminate between the phoneme combinations /bA:d/ and /sA:l/. Phoneme discriminability by infants is addressed in the well-developed phonemic perception literature (e.g., Garnica 1973, Barton 1978). These studies typically employ a method in which a subject discriminates between targets differing by a single phoneme. Often the subjects are taught nonsense words. However, these procedures differ from the one employed here, in that subjects are taught or tested on individual words to a criterion, and then tested for discrimination. The number of presentations of a given word may thus vary. The method we describe is readily adaptable to the study of the acquisition of phonemic perception.

This brings us to the question of whether one of the *wug* tokens was differentially easier to acquire than another. We found no evidence that the differences between the phoneme combinations /bA:d/ or /sA:l/ mediated the uptake of a novel word.

Why did longest look prove to be an effective index of association? At this stage we can only speculate. We found no evidence that in the NEUTRAL condition, when presented unexpectedly with *geek*, subjects glanced rapidly back and forth between the monitors. If this had been the case they would have taken more individual looks in this condition, which they did not. Neither did the result depend upon the learning of a single word. Inspection of Figure 2 reveals that each image, **bard** or **sarl**, attracts approximately equal amounts of looking. This interpretation is confirmed by the lack of interaction in the ANOVA between the TRIALTYPE and IMAGE conditions. More work is clearly needed. We remain confident of our original intuition, arrived at whilst scoring tapes in a pilot experiment: When the child hears the label for an object she recognises, she orients to that object as the referent of the word she hears. This results in longer looks.

What mechanism is at work? Bates (1993) discusses preferential looking as a measure of comprehension. She cites the large literature on preferential looking in children under the age of 6 months. This literature makes the opposite assumption to that adopted in the verbal comprehension literature: young children look longer at surprising stimuli which do not match their expectations (Spelke, Breinlinger, Macomber and Jacobson, 1992). The two effects are not equivalent since one (anomalous displays, e.g. of the sort adopted by Wynn (1992) and Baillargeon (1994)) involves looking at a single display in a

single mode of presentation, whereas the other involves a pair of opposed displays and cross-modal presentation.

The precise nature of the association revealed in the experiment therefore remains to be explored. It may be that learning has been achieved by a "highly effective non-linguistic associative mechanism" (Woodward et al. 1994, p564). Our experiment does not disconfirm the idea that infants deploy a simple associative mechanism for rapid word learning. However, their responses are unlikely to result from simple classical conditioning, insofar as the side to which the subject must turn in the MATCH condition is subject to random variation. Nevertheless, the processing demanded by the auditory labels need not necessarily be linguistic. Savage-Rumbaugh, Murphy, Sevcik, Brakke, Williams and Rumbaugh (1993) discuss what it means for an ape to use a word as referent. They point out that the fact that apes acquire 'namelike' associations does not imply that they understand these names as used by others (Savage-Rumbaugh et al. 1993, p16). The experimental design is readily adaptable to the investigation of such issues as whether non-speech labels would be equally as effective at driving preferential looking, and for the presentation of recently-learned stimuli in a wide variety of linguistic settings.

A related question concerns the impact of fluent speech for the learning of novel words. Almost all researchers have used a command to the subject of the form "Where's the ...?" or "See the...?" (Behrend 1988, Golinkoff, Hirsh-Pasek, Cauley and Gordon 1987, Reznick 1990, Thomas et al. 1981). We used single-word stimuli, because we were interested in the possibility of phonemic content mediating uptake of novel words. Single real-word stimuli have been shown to drive preferential looking in 16-month-old to 24-month-old subjects (Plunkett and Schafer, in preparation). However, presenting the stimuli in a continuous phrase or sentence necessarily makes the task a linguistic one. An extension of our procedure in this direction would permit a tightly controlled investigation of how characteristics of the speech signal can facilitate lexical segmentation in continuous speech.

References

- Baillargeon, R. (1994). Physical reasoning in young infants: Seeking explanations for impossible events. *British Journal of Developmental Psychology*, **12**, 9-33.
- Baker, C. L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry*, **10**, 533-581.
- Baldwin, D. A. (1993). Infants' ability to consult the speaker for cues to word reference. *Journal of Child Language*, **20**, 395-418.

- Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: A first step. *Child Development*, **60**, 381-398.
- Barrett, M. D. (1978). Lexical development and overextension in child language. *Journal of Child Language*, **5**, 205-219.
- Barton, D. (1978). The discrimination of minimally-different pairs of real words by children aged 2;3 to 2;11. In N. Waterson & C. Snow (Eds.), *The development of communication* (pp. 255-261). New York: John Wiley.
- Bates, E. (1993). Commentary: Comprehension and production in early language development. In S. Savage-Rumbaugh et al. (Eds.), *Language comprehension in ape and child* Chicago: University of Chicago Press.
- Behrend, D. A. (1988). Overextensions in early language comprehension: evidence from a signal detection approach. *Journal of Child Language*, **15**, 63-75.
- Berko, J. (1958). The child's learning of English morphology. *Word*, **14**, 150-177.
- Charles-Luce, J., & Luce, P. A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language*, **17**, 205-215.
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Eds.), *Mechanisms of Language Acquisition* Hillsdale, NJ: Erlbaum.
- Corrigan, R. (1978). Language development as related to stage 6 object permanence development. *Journal of Child Language*, **5**, 173-189.
- Dore, J. (1978). Conditions for the acquisition of speech acts. In I. Markova (Eds.), *The Social Context of Language* (pp. 87-111). New York: Wiley.
- Dromi, E. (1986). The one-word period as a stage in language development: Quantitative and qualitative accounts. In I. Levin (Eds.), *Stage and structure: Reopening the debate* Norwood, NJ: Ablex.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **59**(5), 1-189.
- Garnica, O. K. (1973). The development of phonemic speech perception. In T. Moore (Eds.), *Cognitive development and the acquisition of meaning* (pp. 214-222). New York: Academic Press.
- Gathercole, V. C. (1987). The contrastive hypothesis for the acquisition of word meaning: A reconsideration of the theory. *Journal of Child Language*, **14**(3), 493-531.
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, **28**(1), 99-108.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, **14**, 23-45.
- Gopnik, A., & Meltzoff, A. N. (1986). Words, plans, things and locations: Interaction between semantic and cognitive development in the one-word stage. In S. Kuczaj & M. Barrett (Eds.), *The development of word meaning* New York: Springer-Verlag.
- Kirk, R. E. (1982). *Experimental Design* (2nd ed.). Pacific Grove, CA: Brooks/Cole.
- Liberman, A. M., Cooper, F., Shankweiler, D., & Studdert-Kennedy (1967). Perception of the speech code. *Psychological Review*, **74**, 431-459.
- Lock, A. (1980). *The guided reinvention of language*. London: Academic Press.
- Lucariello, J. (1987). Concept formation and its relation to word learning and use in the second year. *Journal of Child Language*, **14**, 309-332.
- Mann, V. A., & Repp, B. H. (1980). Influence of vocalic content on the perception of the [sh]-[s] distinction. *Perception and Psychophysics*, **28**, 213-228.
- Mount, R., Reznick, J. S., Kagan, J., Hiatt, S., & Szpak, M. (1989). Direction of gaze and emergence of speech in the second year. *Brain and Language*, **36**, 406-410.
- Naigles, L. G., & Gelman, S. A. (1995). Overextensions in comprehension and production revisited: preferential looking in a study of dog, cat and cow. *Journal of Child Language*, **22**, 19-46.
- Nelson, K. E. (1973). *Structure and strategy in learning to talk*.
- Nelson, K. E., & Bonvillian, J. D. (1973). Concepts and words in the 18 month-old: Acquiring concept names under controlled conditions. *Cognition*, **2**(4), 435-450.
- Oviatt, S. L. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, **51**, 97-106.
- Pinker, S. (1989). *Learnability and cognition*. Cambridge, MA: MIT/Bradford.
- Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, **20**, 43-60.
- Quenouille, M. H. (1966). *Introductory Statistics*. Oxford: Pergamon Press.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Reznick, J. S. (1990). Visual preference as a test of infant word comprehension. *Applied Psycholinguistics*, **11**, 145-166.
- Ross, G., Nelson, K., Wetstone, H., & Tanouye, E. (1986). Acquisition and generalization of novel object concepts by young language learners. *Journal of Child Language*, **13**, 67-83.

- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., & Rumbaugh, D. M. (1993). *Language comprehension in ape and child*. Chicago: University of Chicago Press.
- Slowiaczek, L. M., & Nusbaum, H. C. (1985). Effects of speech rate and pitch contour on the perception of synthetic speech. *Human Factors*, **27**, 701-712.
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review*, **99**, 605-632.
- Stoel-Gammon, C., & Cooper, J. A. (1984). Patterns of early lexical and phonological acquisition. *Journal of Child Language*, **11**, 247-271.
- Thomas, D. G., Campos, J. J., Shucard, D. W., Ramsay, D. S., & Shucard, J. (1981). Semantic comprehension in infancy: A signal detection analysis. *Child Development*, **52**, 798-803.
- Vihman, M. M., Ferguson, C. A., & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, **7**, 3-40.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical Principles in Experimental Design* (3rd ed.). McGraw-Hill, Inc.
- Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, **30**(4), 553-566.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, **358**, 749-750.