

Research Methods in Child Language

A Practical Guide

Edited by Erika Hoff



Preface

The aim of researchers who study child language is to describe children's language skills and language knowledge at different developmental points and to explain how children progress from their starting state to the achievement of adult-level skills and understandings. Language skills are hard to capture, and both the underlying knowledge and the mechanisms that enable language acquisition are hidden from view inside the mind of the child. Thus, researchers who study child language depend on an array of tools to reveal the object of our study. This book is about those tools. Its aim is to describe the techniques child language researchers use as we go about the business of studying language development.

Some of the methods reviewed in this book are very new, for example the use of functional near infrared spectroscopy (fNIRS) to study the activity of the brain. Others have a longer history – for example the collection, transcription, and coding of speech samples – but have been transformed in recent years by new hardware and software. Each chapter author is a researcher who uses and, in many cases, has contributed to the development of the methods described. The authors were asked to describe the research aims their methods serve, the details of the implementation of those methods, and the type of data the methods yield. Each chapter provides some discussion of the alternative methods available to researchers and their attendant advantages and disadvantages. In many cases, the chapters are part personal travelogue, describing the researcher's journey from research aim to research method.

The book is organized into four parts. The first focuses on laboratory techniques that do not require language production from the participant. Most of these are techniques used to study language in infants. In Chapter 1, Christopher Fennell describes habituation procedures and their use in studying infants' abilities to discriminate the smallest meaningful units of sound. In Chapter 2, Janina Piotroski and Letitia Naigles describe the preferential looking method as it is used to study early language comprehension. In Chapter 3, Daniel Swingley describes the development and use of the looking-while-listening procedure, which provides a window onto the online processing engaged in by prelinguistic infants as they listen to speech. In Chapter 4, Ioulia Kovelman reviews the brain imaging techniques that have been used to peek into the neural activity of infants and older children as they process language. In Chapter 5, the final chapter of this part, Roberta Golinkoff and Kathryn Hirsh-Pasek provide a historical overview of the development of these and other methods that have been used to study language in infants, moving from the early questions of what infants on average can do to the more recently asked question of what individual differences among infants portend for their future language development.

The second part of the book surveys methods that have been used to assess language knowledge in children who do produce speech. In Chapter 6, Cynthia Core provides a short course on phonological development and a survey of methods used to assess phonological development in young children. In Chapter 7, Barbara Pan does likewise for vocabulary development and its assessment. In Chapter 8, Ben Ambridge discusses methods of assessing children's grammatical knowledge, focusing in particular on a new technique - the graded grammaticality judgment paradigm. In Chapter 9, Elaine Reese, Alison Sparks, and Sebastian

Suggate describe the story retelling technique they have used to study children's narratives. Chapters 10, 11, and 12 are introductions to three different techniques that have been used to ask questions about children's underlying linguistic knowledge and online processing in both the lexical and the syntactic domains. In Chapter 10, David McKercher and Vikram Jaswal describe the use of judgment tasks. In Chapter 11, Marina Vasilyeva, Heidi Waterfall, and Ligia Gómez describe priming procedures. In Chapter 12, John Trueswell describes eye movement monitoring techniques.

The focus of the third part of the book is on the use of naturalistic methods to capture the speech children hear and the speech they spontaneously produce. In Chapter 13, Meredith Rowe describes methods that have been and are being used to record, transcribe, and code samples of caregiver-child interaction and illustrates the sort of findings such methods can yield. In Chapter 14, Erica Cartmill, Özlem Demir, and Susan Goldin-Meadow describe the methods used to study the gestures children produce and observe as they communicate and the role of gesture in language development. In Chapter 15, Elena Lieven and Heike Behrens describe the dense sampling procedure they and others have used to capture the nature of children's language input and spontaneous speech. In Chapter 16, Letitia Naigles describes techniques that aim for even more than a dense sample - techniques to capture everything a child hears and/or says. In Chapter 17, David Dickinson describes approaches to capturing teacher-child interactions in preschool classrooms. In Chapter 18, Roberta Corrigan provides an introduction to the data archive and analysis tools that are the Child Language Data Exchange System (CHILDES).

The organizing topic for the fourth part is the question of what we can learn and how we go about learning it when

we study populations other than typically developing, monolingual children acquiring English. In Chapter 19, Aylin Küntay discusses crosslinguistic research. In Chapter 20, Rosario Rumiche and I describe the particular challenges of research with bilingually developing children and their families, and we describe the methods we have used in our research. In Chapter 21, Karla McGregor discusses methodological issues that are unique to the study of children with language impairment, and she reviews the standards of scientific quality that pertain to research that will provide an evidence base for clinical practice. In Chapter 22, Leonard Abbeduto, Sara Kover, and Andrea McDuffie describe their work studying language development in children with intellectual disabilities.

One aim of this collection is to provide the reader with more background and procedural detail about each method than can be included in a journal article. (Another is to be a bit more readable than the necessarily dense prose of an APA-style method section.) The hope is that the information presented in these chapters will be of use to advanced students beginning research in the field of child language, to established researchers embarking in new directions, and to readers of the scientific literature who would like more background on the procedures that yielded the data they are reading about.

Part I

Studying Infants and Others Using Nonverbal Methods

1

Habituation Procedures

Christopher T. Fennell

Summary

In this chapter, the habituation technique will be described in detail, with a focus on infant visual habituation tasks. This easy-to-implement procedure can be used to measure many domains of early language development; however, it has been primarily used to test questions of language discrimination and word learning. In a typical experiment, an infant sits on a parent's lap, or in an infant seat, listening to a repeated sound or word, which can be paired with a visual display. Once her response (e.g., sucking, heart rate, orienting behavior) decreases to a preset criterion, a test trial is presented where the sound or word changes. An increase in response indicates successful discrimination and potential learning of the sound or word. Key advantages of habituation tasks are their use of autonomic responses, broad age range from fetuses to adults, and applicability to multiple populations, including typically and atypically developing children.

“Your child will get bored, and that’s ok.” Many of us who use habituation to study language development state this to the parents of our participants. In a world of infant videos, playgroups, and exciting toys, the idea that we can glean rich knowledge from purposefully boring an infant can seem strange. Yet, this statement essentially encapsulates the habituation task, a procedure that has answered

fundamental questions about early language acquisition and continues to be of great use to the field of developmental psycholinguistics.

Research Aim

The concept of habituation has a long history in psychology, with forefathers like Wundt and Thorndike exploring human adaptation to a recurring stimulus. Simply defined, habituation is the progressive reduction of an organism's behavior in response to a repeated stimulus. Importantly, the reduction in behavioral response is thought to demonstrate both memory encoding of the stimulus and potential learning. However, as Cohen (2004) and others have stated, the decrease of a response over time may not involve true habituation. The organism may just be fatiguing in general, which may be especially true of infants and children. Thus, habituation procedures require the introduction of a stimulus change after the habituation phase ends. If the organism demonstrates an increase in target behavior in response to the change, the researcher can now state with greater confidence that participants remembered and learned the habituation stimulus on some level. This particularly allows for tests of discrimination by using a similar, but novel, stimulus at test.

Despite its long history, it was not until Fantz's (1964) seminal article that *infant* habituation was broadly introduced into psychological research (see also Golinkoff and Hirsh-Pasek, Chapter 5 this volume). Over the past three and a half decades, there has been an explosion of infant habituation studies, with such pioneers as Leslie Cohen contributing much to our understanding of using this technique with this population. In language development research, habituation tasks have tested such diverse aspects of infants' language abilities as their ability

to tell one language from another, the specificity of their phonological representations (e.g., discriminating a /b/ sound from a /d/), their ability to learn word-object associations (e.g., pairing a novel word with a novel object), and their ability to learn grammatical rules (e.g., learning a word order pattern where nonsense syllables follow an ABA configuration like wo fe wo, and then noticing violations of that arrangement – an ABB pattern like li gi gi).

Why have infant habituation procedures enjoyed such broad use? First, while the response measured could be theoretically any form of behavior, it often involves an autonomic physiological response such as sucking, heart rate, or orienting behavior (e.g., looking). This is critical because overt conscious behaviors are harder, and sometimes impossible, to elicit in the prenatal through infant phases of development, exactly when important language abilities are unfolding. Further, these autonomic responses are valid measures across wide populations of participants, including different species, both typically and atypically developing populations, and – most importantly for developmental studies – a wide range of age groups, from fetuses to adults. Using the same task across ages is optimal, as it is difficult to track development when using a variety of different methodologies, each with its own task demands. Another strength is the breadth of possible stimuli; any repeated stimulus is a potential candidate, including phonemes, stress patterns, words, signs, sentences, etc. Finally, and perhaps most importantly, habituation reflects the cognitive structure of the infant mind. Infants' ability to habituate and react to novel stimuli is so fundamental that it is part of neonatal assessment (Brazleton Scale) and correlates with later cognitive skills (Berg and Sternberg, 1985).

Procedure

In my own research, I have found the habituation task to be invaluable. I examine infants' abilities to detect and use the smallest meaningful units of sound in a language: phonemes. For example, /b/ and /d/ are English phonemes because they denote meaningful differences in words, like "bad" and "dad." I and my colleagues, chief among them Janet Werker, have used visual habituation tasks to investigate whether infants can discriminate two phonemes in speech perception and if they can then use that same phonological information when acquiring words in the lab. It is a testament to the power of the task that we can use the same procedure to answer both phonetic and lexical questions over various points of development.

It is no surprise that habituation tasks are commonly used to investigate phoneme distinction in infants, as they are designed to be tests of discrimination. But, why choose habituation over other available methodologies? The main advantage of habituation over its main competitor, the conditioned head turn procedure (CHT), is its use of autonomic responses. CHT, on the other hand, involves an initial phase where infants, via operant conditioning, are rewarded for turning their heads when hearing a stimulus change. This places extra demands on infants and the experimental setup, as CHT requires two experimenters and more equipment (i.e., a reward stimulus - usually an animated toy that is triggered by the first experimenter). Further, CHT cannot be reliably used under 6 months of age, yet a large portion of language development, especially that concerning phoneme perception, occurs prior to this age. However, one advantage of CHT over habituation is that one can meaningfully interpret a single individual's data, as there are multiple change trials. Thus, one can see if an infant can reliably discriminate two stimuli, which is of great importance for clinical applications. The visual habituation task, however, typically

involves only one change trial (the novel trial). It is therefore difficult to interpret an individual's data, as she may have had increased, decreased, or equivalent looking times to the novel stimulus compared to a habituation stimulus for any number of reasons (e.g., fatigue, distraction, etc.). For this reason, we can only interpret group data in habituation studies.

Another alternative test of infant phoneme discrimination is the event related potential (ERP) methodology, which measures electrical neural responses to stimuli via electroencephalography (EEG) (see Kovelman, Chapter 4 this volume). There is a standard ERP produced, called the mismatch negativity response, when we detect a change in auditory stimulus. However, more expensive resources and longer experimenter training are required to use ERP than to use the habituation method. Further, there is an added level of difficulty in setup and a corresponding higher level of attrition.

Having chosen a habituation task to test phonological discrimination, for example, it is important to select the appropriate response behavior. Early infant habituation researchers used heart rate and sucking behaviors as dependent variables, demonstrating that these responses decreased as an auditory, olfactory, or visual stimulus repeated. While these measures are less common in modern research, they are both valid, especially with certain populations. For example, during the fetal period, heart rate is one of the few possible behavioral measures. Kisilevsky *et al.* (2009) demonstrated that fetuses will show a novelty response of increased heart rate to maternal speech after being familiarized to a female stranger's speech. In the neonatal and early infancy period, sucking can be preferable to orienting behaviors, such as head turning and looking, given that the very young infant has limited head control and an underdeveloped visual system.

Indeed, the first study to demonstrate that young infants can categorically perceive phonemes used a habituation paradigm with sucking as the dependent measure (Eimas *et al.*, 1971). In another example, Shi, Werker, and Morgan (1999) used this response to establish that newborns can discriminate a spoken list of grammar words (e.g., “in,” “on”) from a list of nongrammatical, or content, words (e.g., “mommy,” “chair”). However the major drawback of using heart rate and sucking measures is that they require monitoring equipment to be in contact with infants’ bodies: an electrocardiogram or a pacifier outfitted with a pressure transducer. These can be expensive and subject to equipment problems. Sucking is also limited in terms of the age range: a 20-month-old may not easily accept a pacifier to be placed continuously in her mouth during testing.

A third measurement option for auditory language skills – a visual orientation response to a pattern on a screen – was validated by Horowitz’s (1975) demonstration of a positive relationship between attention to an auditory stream and visual fixation. This measure is very advantageous, as it requires nothing to be in physical contact with infants. It can even be done without monitoring equipment (i.e., an experimenter observing infants’ looks with her own eyes), although it is recommended that a video record be obtained for coding purposes. It should be noted that another orienting behavior could be used: head turns. Using this measure, Swain, Zelazo, and Clifton (1993) showed that newborns could remember a habituated word form for up to 24 hours. However, this blunt measure is used less often than looking time, which has more informative small variations (e.g., small glances, look aways). Even its limitation in the neonatal period (i.e., poor visual capabilities) can be overcome through the use of closer, more contrastive visual stimuli. The ease of measuring

looking behavior and its broad age range has led to its wide application,¹ including in my own studies.

We can now turn to an illustration of how to test infants' perceptual discrimination of phonemes using habituation. Note that this same method can be used for any language distinction. Habituation studies necessarily involve two phases, habituation and test, which comprise discrete trials wherein an audio, visual, or audiovisual stimulus is presented. In a phoneme discrimination task, a recording of a female producing syllables in infant-directed speech, which infants prefer to adult-directed speech, is delivered around 65 dB. The visual stimulus is a pattern, usually a checkerboard. Each trial is preceded by an attention-getting stimulus in order to get infants to orient to the screen (e.g., a flashing light; a morphing, colourful shape; the face of a baby accompanied by giggling). Once infants look to the screen, the relevant trial commences. Trials can be of set length, or can terminate once an infant looks away for a set time. The latter infant-controlled option can be more sensitive, as it takes into account individual differences in attention on a trial-by-trial basis. However, it also introduces potential observer bias, as the experimenter decides when infants are no longer attending. Therefore, care should be taken when using this variant, ensuring that the experimenter is blind to both the stimulus and the experimental phases.

The physical setup of the procedure is quite straightforward – another major task advantage. The infant and one parent sit facing a visual display, with the parent wearing sound-masking headphones. Given that looking time is our measure, one important requirement is that the room is dimly lit so that the visual image stands out. Researchers achieve this by turning off any overhead lighting and placing a shaded lamp (or lamps) to the left and/or right of the infant at a 45 degree forward angle.

This allows for a clear image of the infants' eyes. Sometimes researchers will surround the visual display with black cloth that stretches the width and height of the room, which will provide a stronger contrast for the presented images. Usually, the task takes place in a small, quiet (or even soundproof) room to aid in the acoustic presentation. Finally, infants' looking times are recorded using a video camera, with this record being used for reliability coding. This camera should be hidden from infants' view so as to not distract them from the visual stimuli.

The software to run the procedure can be created in the lab, or one can use a common freeware program called Habit (Cohen, Atkinson, and Chaput, 2004), which will order stimuli presentation, compute habituation criteria, and accumulate looking time data. Stimuli are usually played from digitized files on the computer and are sent to the display and speaker in the testing room. The experimenter, who should be blind to the audio stimuli being presented and to whether a trial was a habituation or test trial, remotely monitors the infant's looking times. A designated key is pressed on the computer keyboard during infant looks, which this program records.

The habituation phase is, of course, of prime importance to the procedure and where the researcher has many options available. For example, one could end the habituation phase once an infant accrues a certain amount of looking time (e.g., 2 minutes total). These studies are usually termed familiarization rather than habituation and, while they can provide rich data, they can be prone to problems (see Cohen, 2004), as there is no guarantee that all infants would require the same amount of training. Due to individual differences in attention and cognitive skills, one infant may require 2 minutes to learn the stimuli and another may require 3 minutes. To ensure that all infants

are on the same page, a true habituation criterion should be used (e.g., looking time across a block of trials falls to 65% of the highest total looking time summed across the same-sized block).² Importantly, habituation criteria change based on infant age, decreasing as they get older (using 65% at 14 months, but 50% at 20 months). Selecting the appropriate criterion is critical because if the criterion is too strict, the attrition rate will increase; infants become too bored with the stimuli. If it is too lax, infants will not yet be habituated and one may obtain null results that are not indicative of their true ability. A maximum number of trials should still be included in the design (e.g., 24 trials) so that the experiment does not continue indefinitely. One should compare the results of infants who achieved the true criterion (habitutors) and those who reached the maximum number of trials without habituating (nonhabitutors) to investigate if there are any performance differences. For example, Werker *et al.* (1998) found that nonhabitutors did not successfully learn their habituation stimuli, whereas habitutors did.

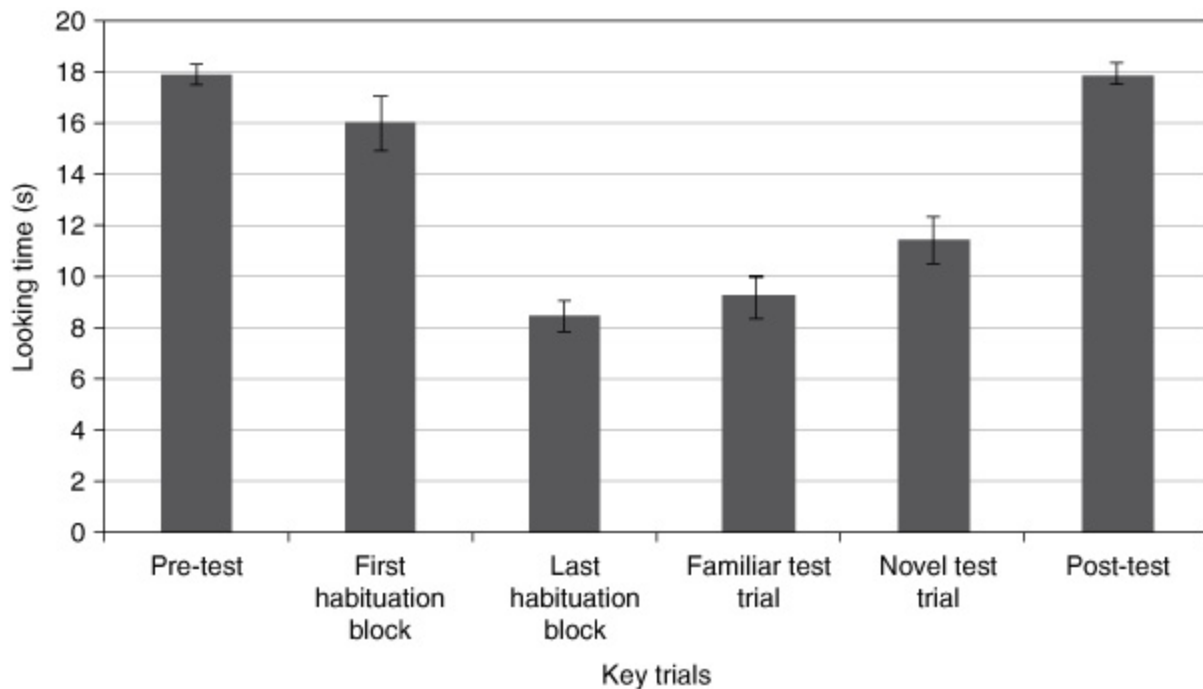


Figure 1.1 Average looking times to key trials in an infant habituation experiment involving a phoneme change at test ($N = 16$). At post-test, infants recovered to pre-test levels and thus were not generally fatigued. Infants habituated, having significantly less looking to the final habituation block than to the first (block = two trials). There is no significant difference between the last habituation block and the familiar test trial, indicating that infants are still bored with the habituation stimulus. Finally, and most importantly, infants noticed the change in stimulus, as the novel test trial is significantly different from both the familiar test trial and the last habituation block.

At test, there are also options available. One possibility is to compare the novel stimulus to the final block of habituation trials to see if they are significantly different. However, this approach has been criticized, as the final block of habituation trials is necessarily low, and may be artificially so (see Cohen, 2004). For example, perhaps an infant was distracted by their loose shoe for one trial and did not attend to the habituation stimuli. The habituation

phase ends due to the low response; however, the child was not truly habituated. For this reason, the researcher should include two trials at test – the novel stimuli and another trial of the habituation stimulus – with the order counterbalanced across participants (e.g., Stager and Werker, 1997; Werker *et al.*, 1998). A within-subjects comparison of those two test trials, usually termed novel and familiar, will reveal if participants can detect the difference between the habituated stimuli and something new. Alternatively, one can compare two different groups of infants: one group that received the novel stimulus after habituation, and one that received a familiar stimulus after habituation. Although it removes any possible order effects at test, this last method is less frequently used since it runs into problems of matching the two groups and requires twice the number of participants. The statistical analysis to determine if the novel and familiar trials are significantly different can be a *t*-test, either paired-sample or independent depending on the design. However, an ANOVA that includes the test trials as one variable and gender as the other is recommended, as some work has found female advantages in this type of task (e.g., Fennell, Byers-Heinlein, and Werker, 2007; Werker *et al.*, 1998).

Finally, pre- and post-test trials can be included (see Werker *et al.*, 1998). These two trials consist of the identical stimuli, which are maximally different from the habituation and test trials. The pre-test trial occurs prior to the habituation phase and is presented for two reasons: to allow infants to become accustomed to the presentation, and as a comparison trial for the post-test trial. The post-test trial is presented after the test phase. It is expected that if infants are still engaged in the experiment, looking time would recover to near pre-test level during this final trial. Once again, pre- and post-test looking times can be

compared using a *t*-test or an ANOVA that includes gender as a factor.

To determine the reliability of the experimenter's coding, there are three standard procedures. The first requires no extra equipment, but is the least exact. A second trained coder can score the looking times of 25% of the subjects by watching the video records. A Pearson product-moment correlation of original and recoded scores should be greater than 0.95 for the data to be considered reliable. A second method is to have two coders score all the video records using a frame-by-frame analysis to obtain the most exact measures possible. Free software is available to perform this coding (SuperCoder: Hollich, 2005). Finally, but most expensively, a researcher can use an eyetracker (e.g., McMurray and Aslin, 2004). This technology uses the reflection of infrared light to measure the distance between the infant's cornea and pupil, collecting a reading of infant's eye gaze every 20 milliseconds. This provides a very precise recording of an infant's looking behavior (see Piotroski and Naigles, Chapter 2 this volume; Trueswell, Chapter 12 this volume).

Using the recommended setup above, multiple studies have confirmed that after being habituated to audio exemplars from one native-language phoneme category (e.g., /b/), infants look significantly longer to the screen when hearing a new native-language phoneme (e.g., /d/) at test than when hearing the habituated phoneme (e.g., Best *et al.*, 1995; Burns *et al.*, 2007; Polka and Werker, 1994; Stager and Werker, 1997; Sundara, Polka, and Molnar, 2008). The ages tested have ranged from 4 to 20 months and have included both monolingual and bilingual populations.

Visual habituation tasks can therefore be used to test basic language discriminations, including phonemes, quite effectively. When examining infants' use of this phonemic

information in early word learning, it would be efficacious to use the same procedure to test both phoneme discrimination and word learning. In this manner, one can directly compare the same response with the same target stimuli. However, it is important to ensure that habituation procedures can truly measure lexical acquisition, which is more complex and occurs developmentally later than phonological acquisition. After all, other valid measures exist. Many word-learning studies have used face-to-face training sessions to teach older infants and toddlers new words and then tested them via picture selection and pointing tasks. However, face-to-face training opens the door for experimenter bias and lacks strict control (i.e., training differences across participants), whereas habituation tasks present the same pre-recorded stimuli to all participants. Further, these testing methods would be too difficult for younger infants, and perhaps even for older infants. For these reasons, my colleagues and I turned to a visual habituation task that involves word-object associations called the “switch” procedure (Werker *et al.*, 1998). The use of this task allowed us to test infants as close to the beginning of the word-learning period as possible, as it does not place undue demands on participants this age, yet it necessitates that infants associate a word and its referent.

In the switch procedure, the exact same physical setup as in phoneme discrimination is used, but infants are now habituated to two word-object pairings³ and tested on their ability to detect a switch in a pairing. To assess whether infants not only have learned about the words and objects individually, but have linked object A to word A, or object B to word B, they are given two test trials. On the control trial (the “same” trial), a familiar word and object are presented in a familiar combination, e.g., object A with word A. On the test trial (the “switch” trial) a familiar word

and object are presented, but in a new combination, e.g., object A paired with word B. If the infants have learned about the words and the objects individually but have not learned the associative link, the “same” and “switch” trials will be equally familiar, and should attract equal looking times. However, if the infants have learned the link between the specific words and objects, the “switch” trial, as a violation, should thus attract greater looking time than the “same” trial. The same statistical analyses as in the phoneme discrimination task are used for test trial comparison. Pre- and post-test trials are included in this design (object C paired with word C, both of which are maximally different from the habituation stimuli).

Werker *et al.* (1998) demonstrated that infants as young as 14 months can learn dissimilar sounding words (e.g., “lif” vs “neem”) in the switch task. However, when Stager and Werker (1997) tested the specificity of words by testing phonetically similar labels (e.g., “bih” vs “dih”), they found that 14-month-olds failed to notice the violation at test. This was unexpected as the [b]-[d] contrast is phonemic in English and should therefore be easy for a 14-month-old English-learning infant to discriminate, and therefore use in word learning. Thus, Stager and Werker conducted a series of control studies, all using visual habituation tasks, to further investigate why 14-month-olds failed. Stager and Werker confirmed that infants this age could detect the acoustic difference between “bih” and “dih” by running the phoneme discrimination task discussed earlier. To verify that the problem was specific to word learning, they ran the phoneme discrimination task, but replaced the checkerboard with an object. Infants were thus only habituated to one word-object combination in this task and the “switch” trial entailed a switch from the habituated label (e.g., “bih”) to a minimally different label (e.g., “dih”). Even though infants could succeed in discriminating the

labels if they ignored the object and only attend to the audio, the 14-month-olds once again failed.

Based on all the above findings, it would seem that infants of 14 months only have difficulty accessing phonetic detail when they are placed in a word-learning situation. The fact that all of these controls could involve the same task points to the versatility and power of the design. Nevertheless, one never knows whether 14-month-olds' continued failure to learn phonetically similar words really reveals something about their word-learning abilities, or instead whether, for any other of an infinite number of reasons, the task simply failed to reveal an underlying capability. Indeed, there have been two prominent criticisms of using habituation tasks to measure word learning: the blunt nature of the test and the ecological validity of the training. First, let us examine the issue of measurement.

In order for infants to succeed in the switch task, they need to demonstrate sufficient surprise at the violation of the link between object and label. In the case where both labels are similar, they may have not learned the words with sufficient confidence to trigger a novelty response. To examine the possibility that our testing method masked infants' detailed phoneme use in word learning, we maintained the same habituation phase as before but altered the testing phase of the switch task by applying the methodology of the "looking-while-listening" procedure (Yoshida *et al.*, 2009; see Swingley, Chapter 3 this volume). After being habituated to object A paired with "bin" and object B paired with "din," infants received target test trials where both habituation objects appeared simultaneously as they heard one of the object labels. We counterbalanced side of object presentation and object label. These trials were interspersed with filler test trials that used familiar objects and labels (e.g., car, shoe) to acquaint infants with the task. This testing method allowed

for a subtler analysis of infants' object choice than prior switch studies. We could analyze both proportion looking and latency to the correct object, as well as a timecourse analysis of looking. Test trials were analyzed over 367–2000 ms after the onset of the spoken target word. This window allows for infant reaction time at this age and accounts for the normal duration of their attention. We found that the average proportion of fixation to the target object (53.5%) was significantly greater than chance (50%), an effect generated from the first block of four target test trials (56.8%) rather than the second block (49.4%). Thus, infants of 14 months have encoded enough detail in the word form to distinguish the “bin” object from the “din” object at test. However, as is obvious from the data, it is a very delicate effect, which explains why the traditional switch measure did not reveal it. Future research could use this hybrid procedure, albeit with half the testing trials, to investigate subtle word-learning effects. Nevertheless, this does not necessitate a complete rejection of the traditional testing phase, which has revealed important group differences in the past and may involve a requisite amount of confidence in word knowledge that is more applicable to real-word applications (e.g., correlations between ability in the switch task and phonological processing in later childhood).

To return to the other criticism of the word-learning habituation task, some have argued that the task is too “stripped down” to allow for true word learning; there are no linguistic (i.e., syntax) or pragmatic (e.g., naming routines) cues. Consequently, I and Sandra Waxman took up the challenge to redesign the word-learning version of the visual habituation task to include cues to reference that went beyond the simple temporal contiguity of word and object, while maintaining experimental control (Fennell and Waxman, 2010). We hypothesized that the presence of

referential cues would allow infants to learn new words in full phonetic detail. The first issue to address was that, in all previous implementations of the switch habituation task, novel words were presented in isolation, which has adverse consequences for learning object names. In normal conversation, when words appear in isolation, they tend to be proper names (“Daddy!”), commands (“Stop!”), or exclamations (“Wow!”). Importantly, infants are sensitive to this, failing to map isolated words to objects in categorization tasks. We therefore modified the task by presenting the target words in recorded naming phrases that clearly indicated that the to-be-learned word referred to an object (e.g., “Look! It’s the bin!”). These naming phrases provide both pragmatic cues (ritualized forms used to present words to infants) and syntactic cues (presence of a determiner) to reference. Another method we employed to invoke referential word learning in the task was to introduce a brief training period wherein infants saw three familiar objects (car, shoe, cat) each paired with its familiar basic-level object name presented in isolation (e.g., “Car!”). This highlighted that words and objects belonged together in the task. We then gave infants a habituation phase (e.g., a novel toy object paired with “bin”), following by the standard testing phase of a novel and familiar pairing of the word-object combination. In both manipulations, infants of 14 months mapped the novel word in all its detail to the accompanying object and were surprised when the word changed (e.g., “bin” changing to “din”). This demonstrates that simple changes to the auditory stimuli or design can invoke a more ecologically valid and powerful word-learning task, while maintaining the strict experimental controls that are so advantageous in habituation procedures.

Remaining Data Issues

Let us now turn to two final points on data analysis. The first concerns the habituation phase, while the second relates to findings in the test phase that do not resemble the standard novelty response.

Habituation phase data can potentially reveal important group differences. If one group receives one stimulus (or set of stimuli) and another group receives a different one, or if both groups receive the same stimuli but are from differing populations (e.g., bilingual versus monolingual, or male versus female), researchers can compare the following measures to explore any differences in adaptation or learning. One can examine the length of habituation via the mean looking time across the entire habituation phase, the mean number of habituation trials, or – better yet – the slope of the habituation curve. Other relevant measures are infants' initial interest in the stimuli (mean looking time in the first block of habituation trials) and infants' final habituation level (mean looking time in the last habituation block). One may see group effects in this last measure, wherein one group has a floor effect of habituation and another group just makes the criterion. Differences between groups in any of these measures could explain eventual group differences at test, and thus all of them should be explored.

A strange effect that can possibly occur at test is discovering, contradictory to the habituation model, a significantly greater response to a familiar stimulus over the novel stimulus. Interestingly, this significant difference indicates that infants can discriminate the familiar and novel stimuli, yet they continue to have active interest in the familiar stimulus from habituation and avoid a novel stimulus. Thus, a significant familiarity preference is still interpretable in terms of discrimination ability; however, the counterintuitive nature of the response requires a clear explanation.

In a key paper, Hunter and Ames (1988) posited that familiarity effects relate to the ease or difficulty with which infants process the habituation stimuli. They hypothesized that infants would prefer stimuli that are at their optimal level of stimulation, and will actively avoid or show a lack of interest in stimuli above or below that level. For a low to moderately complex stimulus, infants would initially have a high response as it is at their optimal level, but as they became familiar with the stimulus, their interest would wane – the standard habituation curve. However, a complex stimulus may actually generate an avoidance response at first, and would only generate interest once the infant has become familiar enough with it to start processing its properties. Thus, if the infants in this latter case were run for the same habituation time as infants exposed to a less complex stimulus, they may show a greater response to a familiar stimulus, as they had not reached the end of the habituation curve and are still trying to process the habituation stimuli. Importantly, younger and older infants may have differing responses to the same stimulus, as its complexity is dependent on the maturity of the cognitive system. Younger infants run for the same length of habituation time with the same stimulus as older infants may demonstrate a familiarity effect, as the stimulus would be more complex to their system. Indeed, data from both of these manipulations (age and stimulus complexity) have borne out the predictions of Hunter and Ames. Cohen (2004) therefore advocates for a strict habituation criterion (e.g., 50% reduction in looking time) to avoid possible familiarity effects.

However, it is important to note that familiarity effects can be interpreted. As mentioned earlier, a significant preference for a familiar over a novel stimulus still demonstrates that infants can discriminate the two stimuli and indicates that the habituation stimulus was overly

complex for the developing system of the infants. Hunter and Ames also point out that familiarity effects demonstrate infants' motivation to process information. In an example of the utility of these effects, we recently discovered a strong familiarity preference for a specific novel word in my lab. A subsequent examination of the acoustics of the stimuli revealed that the target phoneme in that word was much more acoustically variable than target phonemes in our other stimuli. That variability increased the complexity of the stimuli and drove the familiarity preference in the study. This familiarity preference therefore led to the finding that infants are acutely sensitive to phonetic variability in stimulus sets.

In conclusion, the habituation task, with its long history and uncomplicated application, is a key tool for developmental psychologists to uncover answers to current and future questions about the very beginnings of language understanding in the mind of the infant, and across the lifespan.

Key Terms

Familiarity preference

An infrequent response in a true habituation task where the participant attends more to the familiar (i.e., habituation) stimulus at test than a novel stimulus. It usually indicates that the wrong habituation criterion was employed and/or that the familiar stimulus was too complex for the participant.

Familiarization study

A study in which all participants experience the target stimuli for the same predetermined amount of time. This is contrasted with a study with a habituation criterion.

Familiar stimulus

The stimulus presented throughout habituation.

Habituation

The progressive reduction of an organism's behavior in response to a repeated stimulus.

Habituation criterion

The set percentage to which the participant's behavioral response must decrease from its highest point before the test phase begins.

Infant controlled

A variation of the habituation task where the length of each test trial is fully determined by infants' attention to the presented stimulus. The trial begins when the infant attends to the stimulus and ends when they stop attending.

Novel stimulus

A stimulus presented at test that is distinct from the familiar stimulus. If participants have adapted to or learned the specific stimulus from habituation, they should have an increased response to this stimulus over the familiar stimulus (see *novelty preference*).

Novelty preference

The classic test response in a habituation task where the participant attends more to a novel stimulus than a familiar one post-habituation.

Switch procedure

An associative word-learning variant of the habituation task where participants receive two word-object associations throughout habituation (object A, word A; object B, word B) and are tested on two test trials, a familiar pairing (object A, word A) and a novel one (object A, word B). If the participants learned the associative link, they will show a novelty response.

Notes

- 1 Colombo has advocated for combining measures of heart rate and looking time behaviors to obtain a more complete picture of infant habituation (e.g., Colombo *et al.*, 2004).
- 2 Different block sizes can be used, but are usually two to four trials long. To give a concrete example of a 65% habituation criterion: if an infant looks 100 seconds across the first three trials, the habituation phase would end once her looking time across a set of three trials fell below 65 seconds.
- 3 When using visual objects paired with audio stimuli, it is important that the object moves across the screen to ensure infant attention. However, one has to ensure that the audio stimulus (i.e., word) does not commence at the same moment that the object changes direction on the screen, as this provides amodal cues to the relationship between word and object (see Werker *et al.*, 1998; 2002).

References

- Berg, C., and Sternberg, R. (1985) Response to novelty: continuity versus discontinuity in the developmental course of intelligence. In H. Reese (ed.), *Advances in child development and behavior* (vol. 19, pp. 1-47). New York: Academic.
- Best, C.T., McRoberts, G.W., LaFluer, R., and Silver-Isenstadt, J. (1995) Divergent developmental patterns in infants' perception of two non-native consonant contrasts. *Infant Behavior & Development*, 19, 339-350.
- Burns, T.C., Yoshida, K.A., Hill, K., and Werker, J.F. (2007) The development of phonetic representation in bilingual

and monolingual infants. *Applied Psycholinguistics*, 28, 455-474.

Cohen, L.B. (2004) Uses and misuses of habituation and related preference paradigms. *Infant and Child Development*, 13, 349-352.

Cohen, L.B., Atkinson, D.J., and Chaput, H.H. (2004) A new program for obtaining and organizing data in infant perception and cognition studies (version 1.0). Austin: University of Texas.

Colombo, J., Shaddy, D.J., Richman, W.A., *et al.* (2004) Developmental course of visual habituation and preschool cognitive and language outcome. *Infancy*, 5, 1-38.

Eimas, P.D., Siqueland, E.R., Jusczyk, P., and Vigorito, J. (1971) Speech perception in infancy. *Developmental Issues*, 171, 303-306.

Fantz, R.L. (1964) Visual experiences in infants: decreased attention to familiar patterns relative to novel ones. *Science*, 146, 668-670.

Fennell, C.T., Byers-Heinlein, K., and Werker, J.F. (2007) Using speech sounds to guide word learning: the case of bilingual infants. *Child Development*, 78, 1510-1525.

Fennell, C.T., and Waxman, S.R. (2010) What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, 81 (5), 1376-1383.

Hollich, G. (2005) Supercoder: a program for coding preferential looking (version 1.5). Computer software. West Lafayette: Purdue University.

Horowitz, F.D. (1975) Visual attention, auditory stimulation, and language discrimination in young infants. *Monographs of the Society for Research in Child Development*, 39, pp. i-x, 1-140.

Hunter, M.A., and Ames, E.W. (1988) A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier and L.P. Lipsitt (eds), *Advances in infancy research* (vol. 5, pp. 69–95). Norwood, NJ: Ablex.

Kisilevsky, B.S., Hains, S.M.J., Brown, C.A., *et al.* (2009) Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32, 59–71.

McMurray, B., and Aslin, R.N. (2004) Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy*, 6 (2), 203–229.

Polka, L., and Werker, J. (1994) Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, 20 (2), 421–435.

Shi, R., Werker, J.F., and Morgan, J.L. (1999) Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, 72 (2), B11–B21.

Stager, C.L., and Werker, J.F. (1997) Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, 388 (6640), 381–382.

Sundara, M., Polka, L., and Molnar, M. (2008) Development of coronal stop perception: bilingual infants keep pace with their monolingual peers. *Cognition*, 108, 232–242.

Swain, I.U., Zelazo, P.R., and Clifton, R.K. (1993) Newborn infants' memory for speech sounds retained over 24 hours. *Developmental Psychology*, 9 (2), 312–323.

Werker, J.F., Cohen, L.B., Lloyd, V.L., *et al.* (1998) Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology*, 34, 1289–1309.

Werker, J.F., Fennell, C.T., Corcoran, K.M., and Stager, C.L. (2002) Developmental changes in infants' ability to learn similar sounding words. *Infancy*, 3 (1), 1–33.