Anticipatory Looking in Infants and Adults

Johannes Bjerva
Department of Linguistics
Stockholm University
SE - 106 91 Stockholm
bjerva@ling.su.se

Ellen Marklund
Department of Linguistics
Stockholm University
SE - 106 91 Stockholm
ellen@ling.su.se

Johan Engdahl
Department of Linguistics
Stockholm University
SE - 106 91 Stockholm
johan@ling.su.se

Francisco Lacerda
Department of Linguistics
Stockholm University
SE - 106 91 Stockholm
frasse@ling.su.se

ABSTRACT
Infant language acquisition research faces the challenge of dealing with subjects who are unable to provide spoken answers to research questions. To obtain comprehensible data from such subjects eye tracking is a suitable research tool, as the infants' gaze can be interpreted as behavioural responses. The purpose of the current study was to investigate the amount of training necessary for participants to learn an audio-visual contingency and present anticipatory looking behaviour in response to an auditory stimulus. Infants (n=22) and adults (n=16) were presented with training sequences, every fourth of which was followed by a test sequence. Training sequences contained implicit audio-visual contingencies consisting of a syllable (/da/ or /ga/) followed by an image appearing on the left/right side of the screen. Test sequences were identical to training sequences except that no image appeared. The latency in time to first fixation towards the non-target area during test sequences was used as a measurement of whether the participants had grasped the contingency. Infants were found to present anticipatory looking behaviour after 24 training trials. Adults were found to present anticipatory looking behaviour after 28-36 training trials. In future research a more interactive experiment design will be employed in order to individualise the amount of training, which will increase the time span available for testing.

Author Keywords
Anticipatory looking, eye tracking methodology, first language acquisition

INTRODUCTION
In the field of first language acquisition, research often focuses on infants. As such test subjects are unable to verbally provide answers to research questions, carrying out methodological studies is crucial so that reliable research methods that depend on other behavioural responses can be developed. A behavioural response frequently investigated is looking behaviour (e.g. Sobel, Kirkham, 2006; Shukla, Wen, White and Aslin, 2011), as infants tend to look at objects that they are interested in or that cause surprise. Eye tracking provides a convenient means of efficiently investigating this in various situations. It is particularly well suited for research with infants, as their spontaneous looking behaviour may be used as an indication of their linguistic and cognitive capacities, e.g. in speech categorization experiments. This study employs an anticipatory looking paradigm. It explores the infant's expectation that a visual stimulus will appear at a certain location, following the occurrence of some other stimulus, e.g. an auditory stimulus.

BACKGROUND
Anticipatory looking behaviour is shown by infants aged four months when they are trained to expect a visual stimulus in one of two locations depending on which visual stimulus had been previously presented in a third central location (Johnson, Posner and Rothbart, 1991). Furthermore, after being presented with different sounds co-occurring with images in different locations, 6-month-old infants have been shown to be able to predict the location when presented with the sound only (Richardson and Kirkham, 2004; Shukla, Wen, White and Aslin, 2011).

Adults have also been shown to visuo-spatially index auditory information (Richardson and Kirkham, 2004). When presented with different unrelated facts together with faces appearing in one of four locations, and later quizzed on the facts, they consistently looked toward the area in which the face had appeared when the related fact was originally presented. Additionally, when adults hear a word that describes one of several visible objects, they tend to look at the named object (Cooper, 1974; Huettig and Altmann, 2007; Tanenhaus, Spivey-Knowlton, Eberhard and Sedivy, 1995). Furthermore, depending on the linguistic context, they tend to look at potential visual targets before the target word has been presented (Altmann and Kamide, 2007; Kukona, Fang, Aicher, Chen and Magnuson, 2011).

As both infants and adults are able to learn simple contingencies and respond to them with anticipatory looking behaviour, they are suitable participants for eye tracking research employing an anticipatory looking paradigm. However, the amount of training used in previous studies varies. The present study aims to investigate the amount of training

necessary for infants and for adults to learn an audio-visual contingency and present anticipatory looking behaviour in response to auditory stimuli on a group level.

METHOD
Participants were trained to associate an auditory stimulus with a visual event in a specific location on a screen, through being presented with a series of film sequences in which syllables were systematically paired with images appearing in the different locations on the screen. The participants were periodically tested for anticipatory looking behaviour in response to the auditory stimulus only.

Participants
The participants were infants (n=22, mean age 12.1 months, range 11-13 months) and Swedish speaking adults (n=16, mean age 26.3 years, range 19-63 years). All adult participants were students at Stockholm University, and were awarded a cinema ticket for their participation in the study. The infants' caregivers received a diploma for participating in the study.

Stimuli
Visual stimuli consisting of images (cats, dogs, etc.) were created using MS Paint, Adobe Photoshop 7.0 and Adobe Photoshop CS4. Out of 19 images, one of the images was used as an attention-getter, and the remaining 18 appeared in boxes on the left/right side of the screen at syllable presentation during training.

For the auditory stimuli, a female native speaker of Swedish was recorded in an anechoic chamber, using a Brüel & Kjær Type 2669 condenser microphone, a Type 2690—0S2 pre-amplifier set and Adobe Audition 1.5. The speaker repeatedly articulated the syllables /da/ and /ga/. Two exemplars were selected as stimuli for the study. These syllables were similar in terms of acoustic properties (measured by a portable sound level meter and in Praat 5.2.35, see Table 1).

	/da/	/ga/
Mean intensity (dB$_{SPL}$)	55.0	54.2
Mean f0 (Hz)	151.7	152.4
Duration (s)	0.573	0.561
Occlusion duration (s)	0.213	0.192

Table 1: Acoustic properties of the stimuli syllables.

The visual and auditory stimuli were combined to make film sequences, using Adobe Premiere Pro CS4. The sequences presented to the participants consisted of the presentation of an attention-getter, followed by an auditory stimulus, which was succeeded by a visual stimulus appearing on either the left or the right side of the eye-tracker screen (Figure 1, bottom). Each side of the screen was consistently paired with a syllable (/da/ or /ga/). Every fifth film sequence was a test trial, in which no visual stimulus appeared on the screen after syllable presentation (Figure 1, top). The first test sequence was shown prior to any training in order to capture the spontaneous looking behaviour in a test situation, thus functioning as a baseline to which later tests could be compared. A total of 10 test trials and 36 training trials were shown. Four different experiment versions were used in order to counterbalance for syllable-location combination and test trial order (Table 2).

Experiment version	1	2	3	4
Syllable/box pairing	/da/ - L /ga/ - R	/da/ - R /ga/ - L	/da/ - L /ga/ - R	/da/ - R /ga/ - L
Test 1	/da/ - L	/da/ - R	/ga/ - R	/ga/ - L
Test 2	/ga/ - R	/ga/ - L	/da/ - L	/da/ - R
Test 3	/da/ - L	/da/ - R	/ga/ - R	/ga/ - L
...	...	...	...	...
Test 10	/ga/ - R	/ga/ - L	/da/ - L	/da/ - R

Table 2: Balancing for syllable-box pairing and test order. In versions 1 and 3, the reinforcement image appeared in the left box (L) when the syllable /da/ was presented and in the right box (R) when the syllable /ga/ was presented. Versions 2 and 4 had the opposite pairing. In experiment versions 1 and 2, the syllable in the first test was /da/, while in versions 3 and 4 it was /ga/. The test syllables then alternated systematically throughout the experiment.
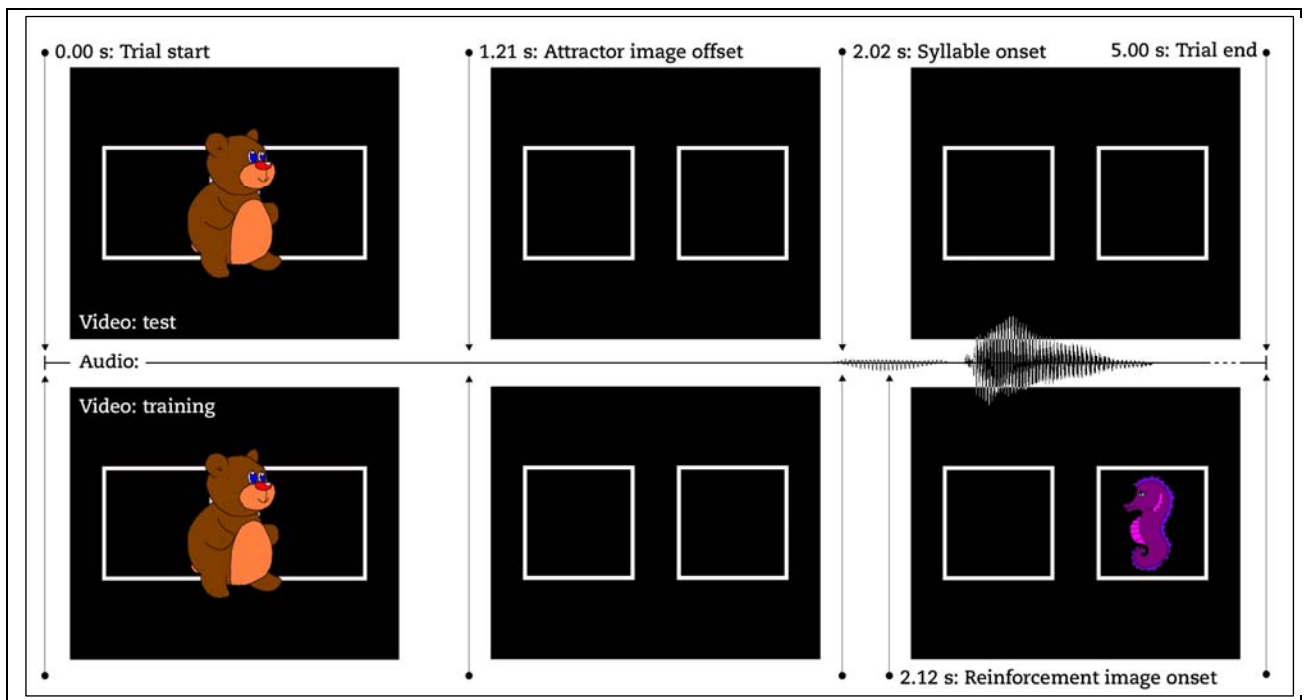
Figure 1: Trial setup. An attractor image was presented at the beginning of each trial, disappearing after 1.21 s. At 2.02 s a syllable was presented. In training trials, a reinforcement image appeared at 2.12 s (bottom), while in test trials no image appeared (top). Total trial duration was 5 s.

Procedure
The infants were seated in their caregiver's lap in front of a Tobii T120 eye-tracking monitor and a set of Creative Inspire T5400 loudspeakers in a sound-attenuated room. In order to prevent the caregivers from influencing the child, they were instructed to remain completely still, and were given sound-insulating PELTOR Workstyle HT7A headphones playing music. The adult test subjects were given no more information than that their eye movements were to be monitored. After calibrating the sound level and the eye-tracking system, the experiment was launched and lasted for approximately four minutes. The eye-tracking system was controlled by the experimenters (using Tobii Studio 2.2.7), who monitored the experiment from a control room, adjacent to the experiment room.

Data preparation and measurements
Data preparation (defining time windows for analysis and Areas of Interest on the screen) was performed in Tobii Studio 2.2.7. The gaze behaviour was investigated by comparing the time to first fixation after syllable onset (TFF) between the baseline test sequence and the other test trials, within the applicable areas. An increased TFF latency within the non-target area (where the image would not have appeared had it been a training trial) was considered to indicate that the participants had shown anticipatory looking behaviour. TFF within the non-target area is preferable over TFF within target (where the image would have appeared had it been a training trial), because it has no lower bound (TFF approaching 0) as participants start to show anticipatory looking behaviour.

RESULTS
For infants, a repeated measures ANOVA of the test by test contrasts revealed significantly longer mean TFF within the non-target area, compared to the baseline test trial, after 24 training trials ($F(1,21)=6.762$, $p<0.017$) (Figure 2, left). The majority of the infants were no longer looking at the screen after 28 training trials, resulting in many instances of missing data. For adults a significant difference was found after 28-36 training trials ($F(1,15)=5.840$, $p<0.029$) (Figure 2, right).
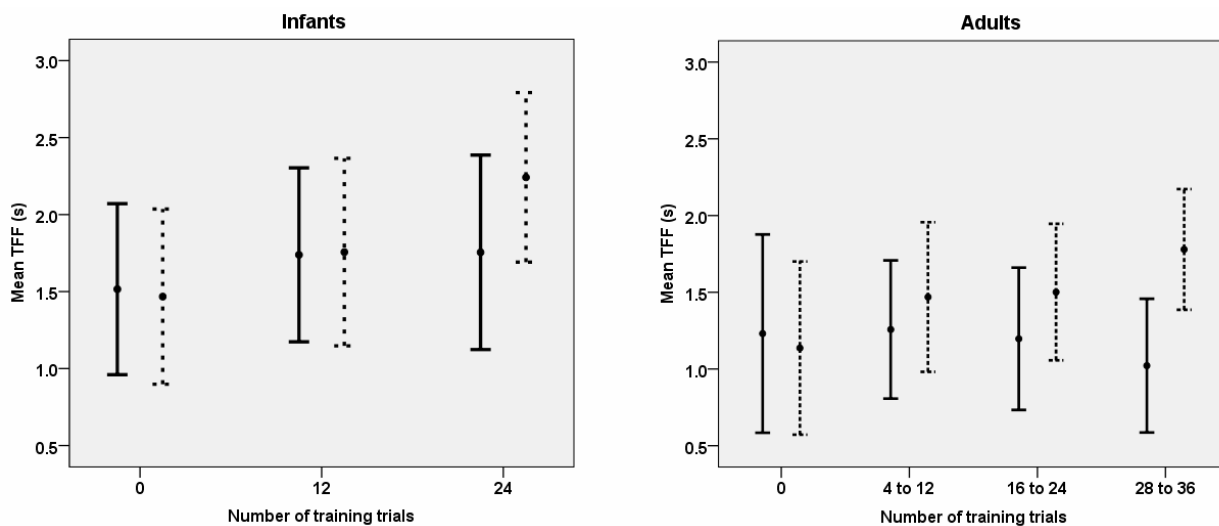
Figure 2: Mean TFF at test trials, within target (solid line) and non-target (dashed line), after 0, 12 and 24 training trials for infants (left) and 0, 4-12, 16-24 and 28-36 training trials for adults (right). The bars depict the 95% confidence interval.

DISCUSSION

The experiment design presented here will be used in experiments with categorical perception in infants, where a continuum of the syllables /da/ and /ga/ will be presented to the test-subject, instead of only two distinctive syllables.

For future research it would be beneficial to employ an adaptive experiment design, in which attractor images can be presented as needed, in order to increase and better utilise the time span during which the infants are attentive to stimuli. An interactive experiment design would also make it possible to individualise the amount of training, so that each participant proceeds to testing only if, and as soon as, training is successful. Such adaptive eye tracking experiments have only recently started to appear in the field of infant research (e.g. Mattsson, 2009; Shukla, Wen, White and Aslin, 2011), but the paradigm has the potential to make important methodological contributions to the field.

REFERENCES

Altmann, G.T.M. and Kamide, Y. The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. Journal of Memory and Language 57 (2007), 502-518.

Cooper, R.M. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. Cognitive Psychology 6 (1974), 84-107.

Huettig, F. and Altmann, G.T.M. Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. Cognition 96 (2005), B23-B32.

Johnson, M.H., Posner, M.I., and Rothbart, M.K. Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. Journal of Cognitive Neuroscience 3 (1991), 335-344.

Kukona, A., Fang, S.-Y., Aicher, K.A., Chen, H., and Magnuson, J.S. The time course of anticipatory constraint integration. Cognition (2011, in press)

Mattsson, L. Prototype of infant hearing test using eye tracking. Master of Science thesis, KTH, Stockholm. (2009)

Richardson, D.C. and Kirkham, N.Z. Multimodal events and moving locations: Eye movements of adults and 6-month-olds reveal dynamic spatial indexing. Journal of Experimental Psychology: General 133, 1 (2004), 46-62.

Shukla, M., Wen, J., White, K.S., and Aslin, R.N. SMART-T: A system for novel fully automated anticipatory eye-tracking paradigms. Behavior Research Methods Online First (2011), 1-15.

Sobel, D.M., and Kirkham, N.Z. Blickets and Babies: The Development of Causal Reasoning in Toddlers and Infants Developmental Psychology 42, 6 (2006), 1103–1115.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., and Sedivy, J.C. Integration of visual and linguistic information in spoken language comprehension. Science 268 (1995) 1632-1634.