Spoken language-mediated anticipatory eye-movements are modulated by reading ability - Evidence from Indian low and high literates

Ramesh K. Mishra, Niharika Singh, Aparna Pandey,
Centre of Behavioral and Cognitive Sciences, University of Allahabad, India
Falk Huettig
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, The Netherlands

We investigated whether levels of reading ability attained through formal literacy are related to anticipatory language-mediated eye movements. Indian low and high literates listened to simple spoken sentences containing a target word (e.g., “door”) while at the same time looking at a visual display of four objects (a target, i.e. the door, and three distractors). The spoken sentences were constructed in such a way that participants could use semantic, associative, and syntactic information from adjectives and particles (preceding the critical noun) to anticipate the visual target objects. High literates started to shift their eye gaze to the target objects well before target word onset. In the low literacy group this shift of eye gaze occurred only when the target noun (i.e. "door") was heard, more than a second later. Our findings suggest that formal literacy may be important for the fine-tuning of language-mediated anticipatory mechanisms, abilities which proficient language users can then exploit for other cognitive activities such as spoken language-mediated eye gaze. In the conclusion, we discuss three potential mechanisms of how reading acquisition and practice may contribute to the differences in predictive spoken language processing between low and high literates.

Key words: anticipatory eye movements, attention, individual differences, literacy, prediction, spoken language comprehension

Introduction

Prediction appears to be a fundamental aspect of human cognition (James, 1890; Pezzulo, Hoffmann, & Falcone, 2007). People predict the outcome of others' actions as they unfold (Sebanz & Knoblich, 2009), ensemble musicians generate online predictions by simulating the concurrent productions of their co-musicians (Keller & Koch, 2008; Wolpert, Doya, & Kawato, 2003), and knowing a co-actor's task influences one's own planning and performance even in situations that do not require taking into account the other's task (Sebanz, Knoblich, & Prinz, 2003, 2005). The mere knowledge of another person's upcoming hand movements results in activation of one's own motor system even when no actual movement is seen (Kilner, Vargas, Duval, Blakemore, & Sirigu, 2004). Similarly, motor activation is observed when individuals use visual cues to prepare their own actions as well as when they use the same cues to predict others' actions (Ramnani & Miall, 2004). Even infants' motor development relies strongly on perception and knowledge of up-coming events (von Hofsten, 2004; see also Hunnius & Bekkering, 2010). Anticipatory eye movements have been reported in a great variety of tasks such as tea-making (Land, Mennie, & Rusted, 1999), sandwich-making (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003), driving (Land & Lee, 1994), piano-playing
(Land & Furneaux, 1997); and appear to support subsequent visuo-motor coordination (Mennie, Hayhoe, & Sullivan, 2006).

In the domain of language processing, it has long been known that predictable words are read faster than unpredictable words (Ehrlich & Rayner, 1981; Rayner & Well, 1996; see Frisson, Rayner, & Pickering, 2005, for recent discussion). More recently, eye-tracking studies using spoken language have shown that participants can use semantic (Altmann & Kamide, 1999) and syntactic (Kamide, Scheepers, & Altmann, 2003) information to anticipate an upcoming visual referent. In one such study, for example, participants were presented with semi-realistic visual scenes depicting a boy, a cake, and some toys while concurrently hearing sentences such as "The boy will move the cake" or "The boy will eat the cake". Eye movements to the cake (the only edible object in the scene) started significantly earlier in the "eat" condition than in the "move" condition (and well before the acoustic onset of "cake") which shows that participants used information retrieved from the verb to predict which object was going to be referred to next.

Similarly, research using event-related brain potentials (ERPs), has accumulated strong evidence that language users can use linguistic input to pre-activate representations of upcoming words before they are encountered (DeLong, Urbach, & Kutas, 2005; Federmeier & Kutas, 1999; Van Berkum, Brown, Zwitserlood, Kooljman, & Hagoot, 2005; Wicha, Moreno, & Kutas, 2004). This suggests that individuals predict the most likely continuations for sentences in advance of the actual input. Such linguistic prediction can be based on semantic (Federmeier & Kutas, 1999), syntactic (Van Berkum et al., 2005), and phonological (DeLong et al., 2005) information.

It appears thus that the importance of prediction for language processing has been well established and consequently (and unsurprisingly) theoretical accounts of predictive language processing have become very influential (e.g., Altmann & Mirkovic, 2009; Chang, Dell, & Bock, 2006; Federmeier, 2007; Kukona, Fang, Aicher, Chen, & Magnuson, 2011; Pickering & Garrod, 2007). One noteworthy aspect of this data however is that almost all studies on predictive language processing have been conducted with undergraduate students (but see Borovsky, Elman & Fernald, in press; Nation, Marshall, & Altmann, 2003). It is, at least, an open empirical question whether the sophisticated language-mediated prediction abilities of Western undergraduate students generalize beyond these narrow samples (see Arnett, 2008; and Henrich, Heine, & Norenzayan, 2010, who argue that the Western student participants used in most experimental studies in psychology are the ‘WEIRDest’ - Western Educated Industrialized Rich Democratic - people in the world and the least representative populations one can find to draw general conclusions about human behavior).

Interestingly, there is increasing evidence from other domains that people's ability to predict and anticipate upcoming events is modulated by their level of expertise on the task at hand. Much of the evidence for this comes from sports psychology. Whether this kind of evidence is considered to be relevant for an investigation of predictive language processing depends perhaps on one's own general theory of cognition but it is (at least) noteworthy that elite basketball players, for instance, predict the success of free shots at baskets earlier and more accurately than people with comparable visual experience (i.e., coaches and sports journalists, Aglioti, Cesari, Romani, & Urgesi, 2008). Similarly, expert volleyball players are superior to novice players in predicting the landing location of volleyball serves (Starkes, Edwards, Dissanayake, & Dunn, 1995). Skilled tennis players are faster than novices in anticipating the direction of opponent's tennis strokes (Williams, Knowles, & Smeeton, 2002), and karate athletes are better than spectators in predicting the target area of an opponent's attack (Mori, Ohtani, & Imanaka, 2002). Such high levels of ability appear to be due to the fine-tuning of specific anticipatory mechanisms that enable athletes to predict other's actions prior to their realization (Aglioti et al., 2008).

Here we sought to establish whether language-mediated prediction is modulated by formal literacy. We compared language-mediated anticipatory eye gaze in high literates (Indian university students with an average of 15 years of formal education) and low literates (Indian manual workers with an average of 2 years of formal education). Does the (in)ability to read and write impact on the tendency to predict which concurrent visual object a speaker is likely to refer to next? In other words, does literacy have effects which go beyond the prediction of words in written texts and increase the likelihood of predictive processing even during spoken language processing?
If prediction is central to language processing, as the empirical results with student participants and the theoretical accounts suggest, then it should be present in all proficient speakers/listeners regardless of their level of formal schooling. Even low literates experience speech every day, and this experience should, by adult age, bring their predictive ability to a ceiling level.

We studied Indian low literates who are particularly suited for such an investigation. More than 35% of the Indian population is considered to be low literate or ‘illiterate’ (UNICEF, 2008). It is important to note here that Indian low literates are fully integrated within Indian society. Low literacy levels are mainly due to poverty and other socioeconomic factors rather than any cognitive impairments or difficulty with reading acquisition (see Huettig, Singh, & Mishra, 2011, for further discussion).

To make the task easy for both participant groups we chose a simple 'look and listen' task reminiscent of everyday contexts. Participants listened to simple spoken sentences while concurrently looking at a visual display of four objects on a computer screen. They were told that they should listen to the sentences carefully, that they could look at whatever they wanted to, but that they should not take their eyes off the screen throughout the experiment (Altmann & Kamide, 1999; Huettig & Altmann, 2005; see Huettig, Rommers, & Meyer, 2011, for further discussion of the method).

We chose a frequent Hindi construction which encouraged anticipatory eye gaze to up-coming target objects. These spoken sentences contained adjectives followed by the particle wala/wali and a noun (e.g., ‘Abhi aap ek uncha wala darwaja dekhnge’, literally: Right now you are going to a high door see - You will now see a tall door). The Hindi particle wala/wali is semantically neutral and not obligatory but frequently used for discourse purposes. Adjective (e.g., uncha/unchi, high) and particle (wala/wali) are gender-marked in Hindi and thus participants could use syntactic information to predict the target. In addition, to maximize the likelihood of observing anticipation effects, we chose adjectives which were also associatively related to the target object. We measured at what point in time in the duration of the spoken sentence low and high literates shifted their eye gaze towards the target objects.

Method

Participants

28 high literates (mean age = 24.6 years, SD = 2.3 years; 15 years mean years of formal education) and 30 low literates (mean age = 28.4, SD = 2.6; 2 mean years of formal education) were paid for their participation. All were from the city of Allahabad in the Uttar Pradesh region of India and had Hindi as their mother tongue. All had normal vision, none had known hearing problems. The study was approved by the ethics committee of Allahabad University and informed consent was obtained from all participants.

The assignment to participant groups was based on the mean number of years of formal education. High literates were postgraduate students of Allahabad University. Low literates were recruited on or around the university campus and asked whether they could read or write. All were engaged in public life and supported themselves by working, for instance, in food and cleaning services on or near the university campus. The low literacy group did not include any individuals involved in an adult literacy program. An average of 2 years of formal education in Uttar Pradesh (as in the low literacy group) tends to result in very rudimentary reading skills. To ensure appropriate participant selection a word reading task was administered to participants. 96 words of varying syllabic complexity were presented. High literates on average read aloud 94.2 words correctly (SD = 1.9) whereas low literates only read aloud 6.3 words correctly (SD = 7.77). None of the participants appeared to be socially excluded, none showed any signs of genetic or neurological disease.

Materials and stimulus preparation

There were 60 displays, each paired with a spoken sentence. 30 trials were experimental trials, the other 30 were filler trials. Each sentence contained a lead-in phrase (‘Abhi aap ek’, Right now you are going to), followed by an adjective (e.g., ‘uncha’, high), then the particle (‘wala’/’wali’) and a noun (e.g., ‘darwaja’, door).

A norming study was carried out to select adjectives which are strongly associated with particular (object) names. 15 literate Hindi native speakers participated, none of them took part in the main experiment. Participants saw a list of 30 adjectives and were asked to write down the first 5 nouns that came to mind. The picturable
noun that was produced most frequently for a particular adjective was selected (e.g., for ‘uncha’, high, participants produced most frequently the noun ‘darwaja’, door). The selected adjective (e.g., ‘uncha’, high) was not associated with any of the other objects in the display. Similarly, the grammatical gender of the adjective agreed only with the target but not with the distractor objects in the same display.

Sentences were recorded by a female native speaker of Hindi. Visual displays in the experimental trials (Figure 1) consisted of line drawings of the target object (e.g., door), and three unrelated distractors. All visual stimuli were frequent and common objects known to both participant groups.

Procedure

Participants were seated at a comfortable distance from a 17 inch monitor. A central fixation point appeared on the screen for 750 ms, followed by a blank screen for 500 ms. Then four pictures appeared on the screen. The positions of the pictures were randomized across four fixed positions of a (virtual) grid on every trial. The auditory presentation of a sentence was initiated 1000 ms later. Preview was provided so that participants had time to look at the objects. Participants were asked to perform a ‘look and listen’ task (Altmann & Kamide, 1999; see Huettig & McQueen, 2007, for discussion).

Results

Figure 2 shows a time-course-graph of proportion of trials with a fixation on the target object, or averaged distractors. The curves are synchronized to the acoustic onset of the spoken adjective. The x-axis shows the time in milliseconds from this onset. The calculation excluded all movements prior to the acoustic onset and therefore negative values reflect that (on average) participants moved their eyes away from objects fixated at this onset. Each data point reflects the proportion of trials with a fixation at that point in time minus the proportion of trials with fixations to that region at the acoustic onset of the adjective (see Huettig & Altmann, 2005). The average noun onset occurred 1560 ms after adjective onset.

Fig. 2 shows that high literates first shifted their eye gaze towards the target from around 800 ms after the onset of the adjective. The average duration of the adjectives was 778 ms (SD = 115). The graph thus suggests that high literates started to predict the up-coming target object well before the acoustic offset of the adjective. Fig. 2 also reveals that the participants in the low literacy group did not show a corresponding early shift in eye gaze towards the target (i.e. fixations to the target only started to diverge from the unrelated distractors around 300 ms after the onset of the noun).

For the statistical analyses we computed mean fixation proportions for each type of object (target object and averaged distractor) per participant and item over a time interval starting from the acoustic onset of the adjective to 100 ms after this onset in order to obtain a baseline of fixation proportions. We can assume that fixations during
this baseline time region were not influenced by information from the critical spoken adjective because of the time considered necessary for programming and initiating an eye movement (Altman, 2011; Saslow, 1967). We calculated fixation proportions during the baseline region to adjust for any bias in overt attention to a type of object before information from the critical adjective became available. Calculating fixation proportions for the baseline time regions (and then comparing these proportions with the mean fixation proportions during subsequent 100 ms time regions) allows us to test for any shifts in overt attention to particular types of objects during times of interest.

Paired t tests showed that for the highly literate participants mean fixation proportions on the target object during the baseline time window (.26) first differed significantly from mean fixation proportions on the target object during the 1000-1099 ms time window (.29), mean difference = 0.032, 95% CI: .055 to .009, d = 0.23; \( t(1, 27) = 2.85, p = .008; t(1, 29) = 2.27, p = .031 \). In contrast, for the low literates, mean fixation proportions on the target object during the baseline time window (.26) first differed significantly from mean fixation proportions on the target only during the 2000-2099 ms time window (.30), mean difference = 0.044, 95% CI: .080 to .009, d = 0.30; \( t(1, 29) = 2.54, p = .017; t(1, 29) = 2.12, p = .043 \).

Discussion

This study was conducted to compare language-mediated anticipatory eye gaze to visual objects in low and high literates. On hearing the biasing adjective and well before the acoustic onset of the spoken target word, high literates started to look more at the target object than unrelated distractors. Low literates’ fixations on the targets only started to differ from looks to the unrelated distractors once the spoken target word acoustically unfolded. Thus high literates shifted their eye gaze towards the target objects about 1000 ms before the low literates.

![Figure 2](image-url)

**Figure 2.** Changes in fixation proportions on the target objects and (averaged) unrelated distractor objects for low literates and high literates. Zero on the timeline is the acoustic onset of the adjective.
We used gender-marked adjectives which are highly associated with the target nouns. High literates' predictions thus relied on either associative (cf. Bar, 2007) or syntactic information. Our data do not conclusively show which type of information high literates used to anticipate the visual object. What our results do show however is that low literates did not consistently use any of the available cues to anticipate the upcoming referent.

One might argue that the anticipation effect in low literates was absent due to noise, or that they understood the sentences in exactly the same way as our highly literate participants, but somehow were less willing or able to shift their eyes to the targets. Our data show that such an account is very unlikely to be correct. The shifts in eye gaze to the target objects of the low literates were closely time-locked to the onset of the noun rather than being randomly distributed across all objects. Figure 2 shows that (taken into account the delay to initiate an eye movement) low literates shifted their eyes towards the target soon after the earliest point in time at which a fixation could reflect a response based on information in the noun. This demonstrates that low literates (as high literates) used information from unfolding spoken words to direct their eye gaze, they just did not use such information for prediction.

A further argument might be that low literates did not process adjective and particle in the same way as the highly literate participants. For instance they may simply have been unable to use the syntactic information of the adjective and the particle for prediction because they did not know that adjective, particle and noun agree in gender. We can also reject this account, our participants did not make any gender errors in their spoken language. It is important to note that our spoken materials were by no means difficult or unusual but simple declarative sentences used in every day situations by high and low literates alike.

Another argument may be that the highly literate participants guessed the purpose of the experiment (i.e. what word will come next in the sentence) and tried to behave accordingly, whereas the low literates did not. If this explanation of the data were correct, then our results may not reflect the ability to anticipate sentence continuation, but instead the ability to guess the purpose of the experiment. We believe this account to be unlikely since our lead-in phrase ('Abhi aap ek', Right now you are going to) was designed to set up an expectation that an object would be referred to. However, further research could usefully explore the extent to which anticipatory processing may be driven by task demands (e.g., in visual world experiments by the limited visual context or the instructions and in ERP experiments using written words by the artificial slow timing of the sentences).

The present group differences are also unlikely to be due to differences in familiarity with 2D representations of real objects. All our objects were line drawings of frequent and common objects familiar to both low and high literates. In a recent study we observed very high naming agreement of similar line drawings in the low literacy group (see Huettig et al., 2011, for further discussion). There is one study (Reis et al., 2001) which has reported a slight difference (approx. 200 ms) in the naming latencies of line drawings between Portuguese illiterates and literates. Our participants however were given a preview of the visual display and thus a small delay in picture naming latencies could not account for the more than 1000 ms delay in shifts in eye gaze to the target objects. Indeed (as mentioned above) our data show that when low literates heard the names of the target objects they quickly shifted their eye gaze to them, which suggests that they had recognized the objects. Thus, we can reject these alternative explanations of our data.

Note that we do not suggest that illiterates and low literates never predict during cognitive processing nor do we claim that they never engage in any form of predictive processing during language processing. When listening to other sentence constructions illiterates/low literates may well be found to engage in some anticipatory processing (though our results do suggest that such context would have to be highly predictive). What we have found is that low literates do not engage in anticipatory eye gaze in Hindi adjective-particle-noun constructions. Our data suggest thus that literacy modulates predictive language processing.

How might formal literacy and language-mediated prediction be related? It has long been known that readers predict up-coming words during reading. As mentioned above, much research has demonstrated that predictable words are read faster than unpredictable words (e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996). We propose that the acquisition and practice of reading increases the likelihood of predictive processing even during spoken language processing. That is we suggest
that literacy has some causal influences that go beyond the prediction of words in written texts. We conjecture that learning to read and write fine-tunes anticipatory mechanisms that involve the retrieval of associated words and the pre-activation of fine-grained (e.g., semantic and syntactic) representations of upcoming words. What could these anticipatory mechanisms be?

One possibility is that the group differences in predictive processing are related to literacy-related differences in adjective-noun associations (cf. Bar, 2007). A related possibility is that illiterates/low literates predict less during language processing because the absence of reading and writing practice in illiterates/low literates greatly decreases their exposure to low level word-to-word contingency statistics. McDonald and Shillcock (2003a,b), in this regard, provided some evidence that readers make use of statistical knowledge in the form of transitional probabilities, i.e. the likelihood of two words occurring together. Moreover, Conway, Bauernschmidt, Huang, and Pisoni (2010) recently demonstrated that performance in implicit learning tasks correlated significantly with the ability to predict the last word of sentences in a written sentence-completion task. Rayner, Warren, Juhasz, and Liversedge (2004) on the other hand have argued that transitional probability effects are unlikely to survive intervening words (cf. Carroll & Slowiaczek, 1986; Morris, 1994). Moreover, Frisson et al. (2005) have questioned whether effects of low level transitional probabilities are independent from ‘regular’ (i.e. higher level) predictability effects (which are typically determined by the use of a Cloze task in which participants are asked to complete sentences or sentence fragments, and predictability is determined by calculating the percentage of times a particular word was given in a particular sentence). Frisson et al. (2005) replicated the findings of McDonald and Shillcock (2003) in a first experiment but, in their second experiment, when items were matched for Cloze values, no effect of transitional probabilities was found.

Thus, a second possibility is that illiterates/low literates predict less during language processing simply because they have acquired less contextual knowledge than high literates. Schwanenflugel and colleagues (e.g., Schwanenflugel & Shoben, 1985) for example have argued that in highly predictive contexts more featural restrictions of up-coming words are generated in advance of the input than in low predictive contexts. These featural (e.g., semantic, syntactic) restrictions may then constrain what words are likely to come up. Individuals with no or low literacy levels, because of the absence of reading, may have had fewer opportunities to increase their general contextual knowledge and may consequently generate fewer featural restrictions of up-coming words, which in turn may result in less online anticipation.

A final possibility we would like to raise here is that reading and predictive language processing may be related to general processing speed. Reading and spoken language comprehension, for instance, differ in the amount of information that is processed per time unit (approx. 250 vs. 150 words/minute). To maintain a high reading speed, prediction, arguably, is helpful if not necessary. Furthermore, it is conceivable that the steady practice of reading enhances readers’ general processing speed. Salthouse (1996) for instance has pointed out that “performance in many cognitive tasks is limited by general processing constraints, in addition to restrictions of knowledge (declarative, procedural, and strategic), and variations in efficiency or effectiveness of specific processes ... it is assumed that general limitations frequently impose constraints on many types of processing and, hence, that they have consequences for the performance of a large variety of cognitive tasks” (pp. 403-404). Stoodley and Stein (2006), for instance, found that dyslexics and poor readers showed a general motor slowing related to a general deficit in processing speed. Of course, this data does not tell us whether the reading problem and the slow processing speed in dyslexics are causally related. It is interesting in this regard however that low literates’ shifts in eye gaze to the target objects, on hearing the acoustic information of the target word, also occurred approximately 200 ms later than for the high literates. In other words, even the ‘non-anticipatory’ shifts in eye gaze when the target objects were named were slightly delayed in the low literacy group.

It is important to note that these potential causal factors underlying the differences in predictive language processing between low and high literates (i.e. low level word-to-word contingency statistics, online generation of featural restrictions, general processing speed) are not mutually exclusive. In fact, they are likely to interact (e.g., a faster general processing speed may result in a greater amount of featural restrictions generated online) and of course there may be other factors, yet to be
explored, which make proficient readers more likely to predict up-coming words.

Finally, we point out that our data cannot tease apart independent effects of formal schooling and learning to read and write. It is notoriously difficult to separate effects of literacy from more general effects of formal schooling since all forms of reading instruction inevitably involve (at least some) aspects associated with formal education. We believe that it is useful to draw a distinction between proximate and distal causes of the observed behaviour. Proximate causes are those which immediately lead to an observed behavior, distal causes are those which are more remote. We suggest that formal schooling is more likely to be a distal cause of the differences in language-mediated prediction between our participant groups whereas literacy is more likely to be a proximate cause. Other distal influences may include parental education, childhood nutrition, and access to medical care. More research could usefully be directed at exploring how these factors influence literacy acquisition.

Conclusions

We observed that high but not low literates showed anticipatory eye movements to concurrent target objects in Hindi adjective-particle-noun constructions. Our data is consistent with the notion that the steady practice of reading and writing enhances individuals' abilities to generate lexical predictions, abilities that help literates to exploit contextually-relevant predictive information when anticipating which object an interlocutor will refer to next in one's visual environment. Our findings highlight the need to investigate a) the degree, and b) the potential mechanisms, of anticipatory language processing in non-student populations.

Acknowledgements

Support was provided by a grant from the Department of Science and Technology, India (RM) and by the Max Planck Society (FH). We thank two anonymous reviewers for their comments on a previous version of this paper.

Address for correspondence: Ramesh Kumar Mishra Centre of Behavioral and Cognitive Sciences (CBCS) University of Allahabad. Allahabad 211002. UPIndia Email:rkmishra@cbcs.ac.in, Ph-91-0532-2460738 (work), Mob-91-9451872007, Fax-91-0532-2460738(work).

References


Altmann, G. T. M. (2011). Language can mediate eye movement control within 100 milliseconds, regardless of whether there is anything to move the eyes to. Acta Psychologica, 137, 190-200.


