Verbal and nonverbal predictors of language-mediated anticipatory eye movements

Joost Rommers · Antje S. Meyer · Falk Huettig

Abstract During language comprehension, listeners often anticipate upcoming information. This can draw listeners' overt attention to visually presented objects before the objects are referred to. We investigated to what extent the anticipatory mechanisms involved in such language-mediated attention rely on specific verbal factors and on processes shared with other domains of cognition. Participants listened to sentences ending in a highly predictable word (e.g., "In 1969 Neil Armstrong was the first man to set foot on the moon") while viewing displays containing three unrelated distractor objects and a critical object, which was either the target object (e.g., a moon), an object with a similar shape (e.g., a tomato), or an unrelated control object (e.g., rice). Language-mediated anticipatory eye movements were observed to targets and to shape competitors. Importantly, looks to the shape competitor were systematically related to individual differences in anticipatory attention, as indexed by a spatial cueing task: Participants whose responses were most strongly facilitated by predictive arrow cues also showed the strongest effects of predictive language input on their eye movements. By contrast, looks to the target were related to individual differences in vocabulary size and verbal fluency. The results suggest that verbal and nonverbal factors contribute to different types of language-mediated eye movements. The findings are consistent with multiple-mechanism accounts of predictive language processing.

Keywords Anticipatory eye movements · Predictive language processing · Visual world paradigm · Spatial cueing

J. Rommers · A. S. Meyer · F. Huettig
Max Planck Institute for Psycholinguistics,
Nijmegen, The Netherlands

J. Rommers
International Max Planck Research School for Language Sciences,
Nijmegen, The Netherlands

A. S. Meyer
Radboud University Nijmegen, Nijmegen, The Netherlands

F. Huettig
Donders Institute for Brain, Cognition, and Behaviour,
Nijmegen, The Netherlands

J. Rommers (✉)
Department of Psychology and The Beckman Institute for Advanced
Science and Technology, University of Illinois, Urbana, IL 61801,
USA
e-mail: jrommers@illinois.edu

Language comes to us at a very high speed; yet, we can usually keep up without much effort. A possible explanation for the rapidity of language comprehension is that listeners anticipate upcoming information. This has, for instance, been established in eyetracking studies in which participants listen to utterances while viewing objects. In such experiments, listeners use aspects of the sentence context to launch eye movements to objects before the objects are referred to. One way to characterize predictive language processing is in terms of basic components such as cues, contents, and mechanisms. The cues are the parts and kinds of context that are used to generate predictions. The contents of predictions refer to the representations that become pre-activated. Finally, the mechanisms refer to the underlying processes that enable readers and listeners to use cues for the generation of one or more predictions. Studies of anticipatory eye movements have shown that the contextual cues that listeners use for prediction range from prosody (Weber, Grice, & Crocker, 2006) and case-marking (Kamide, Scheepers, & Altmann, 2003) to visually presented events (Knoeferle, Crocker, Scheepers, & Pickering, 2005). Furthermore, the contents of predictions can include syntactic structure (Arai & Keller, 2013) and semantic information

(Altmann & Kamide, 1999). Electrophysiological work has demonstrated more contents, including the prediction of specific lexical items (DeLong, Urbach, & Kutas, 2005; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). The mechanisms underlying predictive language processing, however, remain unclear.

We investigated verbal and nonverbal aspects of language-mediated eye movements using an individual differences approach. Underwood (1975) proposed that psychological theories should be formulated in such a way as to allow for individual-differences tests, and moreover that such tests should be the first step in the assessment of a theory (see also Cohen, 1994; Kosslyn et al., 2002; Vogel & Awh, 2008). If the performance in two tasks is substantially correlated, this can be regarded as a "go-ahead signal" for a theory proposing that the cognitive components or abilities assessed in the tasks are related.

We made use of the fact that predictive processing is not only characteristic of language processing but also of many tasks that do not involve language. For instance, in the Posner spatial cueing paradigm (Posner, Nissen, & Ogden, 1978), participants are asked to press a button (left or right) to indicate on which side of the center a symbol (an "X") occurs. Before stimulus onset, a central endogenous cue is presented, which can be valid (an arrow pointing to the side on which the stimulus will occur), neutral ("+"), or invalid (an arrow pointing to the other side). The response time difference between the neutral and valid cue conditions is a measure of the benefit from knowing where the stimulus will occur (Posner, 1980). Electrophysiological studies have provided evidence for perceptual facilitation and response preparation during this task (for reviews, see Coles, 1989; Mangun, 1995), consistent with the idea that participants use the cues to anticipate the location of the upcoming stimulus (Klein, 1994; Nobre, Rohenkohl, & Stokes, 2012; Wang, Fan, & Johnson, 2004).

The question arises to what extent the mechanisms underlying predictive language processing are specific to language tasks or are shared with other domains of cognition. Indeed, several authors have introduced the topic of prediction in language processing by drawing analogies to nonverbal domains (DeLong et al., 2005; Kamide, 2008; van Berkum et al., 2005; van Petten & Luka, 2012). Furthermore, the proposed prediction mechanisms in language include forward models and association-based and event-based processes, inspired by action and perception research (Altmann & Mirković, 2009; Bar, 2007; Pickering & Garrod, 2007, 2013). Finally, prediction has been put forward as a general neural mechanism for perception and action (e.g., Clark, 2013; Friston, 2010).

Against this background, in the present study we investigated the relation of language-mediated anticipatory eye movements to nonverbal anticipatory attention. We investigated two types of language-mediated eye movements, namely

eye movements to predictable target objects and eye movements to objects that were related to predictable targets. This research followed up on a recent study in which we asked participants to listen to sentences ending in highly predictable words (e.g., "Neil Armstrong was the first man to set foot on the moon"; Rommers, Meyer, Praamstra, & Huettig, 2013, Exp. 1). While listening to the sentences, the participants saw visual displays featuring several objects. We found that prior to the onset of the target word, both target objects (e.g., a moon) and other objects with a similar shape (e.g., a tomato) were fixated more often than unrelated distractors, suggesting that visual representations can be preactivated.

In this study, we examined the relationship between the participants' performance in this task and the Posner spatial cueing paradigm described above, as a measure of anticipatory attention. The spatial cueing task has been used widely in the literature and has good reliability (Fan, McCandliss, Sommer, Raz, & Posner, 2002). If language-mediated anticipatory eye movements are related to nonverbal spatial attention, we should see a correlation between the performance in these tasks. Furthermore, if a single, domain-general mechanism underlies performance, this correlation should hold for anticipatory looks to both the target and the shape competitor.

In addition, we collected data from two verbal tasks. The first was the category fluency task, in which participants produce as many words belonging to a specific category as they can in one minute (Delis, Kaplan, & Kramer, 2001). This task has been shown to correlate with the amplitude of electrophysiological effects associated with predictive language comprehension (Federmeier, Kutas, & Schul, 2010; Federmeier, McLennan, De Ochoa, & Kutas, 2002). Second, the participants' vocabulary size was assessed with the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997). It has previously been shown that individuals with relatively good receptive vocabulary knowledge predict more than those with lower vocabulary scores, possibly because they are better able to use the sentence context for prediction (Borovsky, Elman, & Fernald, 2012; Federmeier et al., 2002). Finally, we used Raven's Advanced Progressive Matrices test (Raven, Raven, & Court, 1998) to assess and control for differences between participants in terms of fluid intelligence.

Method

Participants

Eighty-one adult participants (60 women, 21 men; mean age 21 years, range 18–32 years) were paid for participation in the study. All were native speakers of Dutch and had normal hearing and normal or corrected-to normal vision.

Stimuli, design, and procedure

Participants performed the tasks in the order given below. The fluid intelligence test was administered in a second session within three weeks of the first session. An additional action observation task was initially included, but had to be discontinued due to equipment failure.

Category fluency Participants produced as many members of a category (animals and professions, indicated by a word on the screen) as they could in 1 min. Speech was recorded with a Sennheiser microphone. The total numbers of correct responses (excluding mistakes and repetitions) were counted and averaged across the two categories.

Visual-world experiment Participants listened to predictable words in sentence contexts (e.g., the Dutch translation equivalent of "In 1969 Neil Armstrong was the first man to set foot on the moon") while looking at visual displays of four objects that differed depending on the condition (see Fig. 1). The 96 experimental items and 32 fillers from Rommers et al. (2013) were used. The experimental sentence contexts preceding the predictable word varied in duration, with an average of 6, 735 ms ($SD = 1,193$). The filler sentences were predictive and referred to an object on the screen, so that on 50 % of the trials a target object was displayed. The materials were distributed across three counterbalanced presentation lists such that no picture or sentence was repeated. Trial order and picture positions were randomized differently for each participant.

Participants were asked to listen to the sentences carefully and were told that they were free to look at whatever they wanted to, but that they should not take their eyes off the screen. They were tested individually in a dimly illuminated room, seated in front of a screen with their chin on a chin rest. The movements of the right eye were recorded with an EyeLink 1000 tower mount eyetracker sampling at 1000 Hz.

In the earlier study, the objects had appeared only 500 ms before the onset of the critical word. We now allowed for more variability in anticipatory eye movements by having the objects appear before sentence onset, as has been the case in most published visual-world studies (see Huettig, Rommers, & Meyer, 2011). Each trial started with a central fixation circle that remained on the screen until the participant fixated it, allowing for drift correction. Then four objects appeared on the screen. After 1 s, a spoken sentence began playing. The objects remained on the screen until 2,000 ms after sentence offset. Then a blank screen appeared for 500 ms. Data were coded in terms of fixations, saccades, and blinks using the algorithms provided in the EyeLink software.

Peabody Picture Vocabulary Test We used a computerized version of the task developed by Dunn and Dunn (1997; Dutch translation by Schlichting, 2005). On each trial, participants heard a word and saw four numbered pictures on the screen. They indicated which of the pictures corresponded to the spoken word by typing in the number.

Posner spatial cueing task We used a single-session version of the task developed by Posner, Nissen, and Ogden (1978). Participants pressed one of two buttons (left/right) to indicate the location of an X symbol on the screen. Each trial started with a central fixation dot that participants were asked to keep fixating. Then a cue appeared in the position of the dot for 100 ms. The participants were told that on half of the trials, this cue was neutral (+), and on the other half, an arrow cue pointing left (<) or right (>) indicated the location of the upcoming X with 80 % validity. After cue offset, the fixation dot was presented for another 400 ms. Then the target appeared to the left or right of the center of the screen for 1,700 ms or until a response was made. After every trial, feedback was displayed for 500 ms ("correct," "incorrect," or "too late").

There were seven blocks of 32 trials each; the first block served as practice. Each block consisted of 16 neutral trials, 13 valid trials, and three invalid trials in random order, with counterbalanced target position (left or right) and initial fixation dot duration (400, 800, 1,200, or 1,600 ms). Because we noticed that on a small portion of the trials a response occurred before target onset, which was not registered and usually elicited a second but late response, we excluded trials with response times 2.5 standard deviations from each participant's
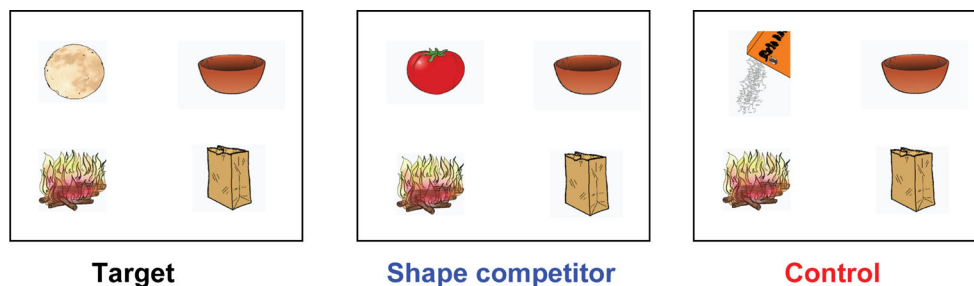
Fig. 1 Example displays for each of the three conditions for the sentence "In 1969 Neil Armstrong was the first man to set foot on the moon." Shown along with the three distractors (a bowl, a fire, and a bag) are the target condition (with a moon), the shape competitor condition (with a tomato), and the control condition (with rice)

mean. The effect of anticipatory attention was quantified as the difference between the mean response latencies of the neutral and valid conditions.

Raven's Advanced Progressive Matrices We used a computerized version of Raven et al. (1998) task. On each trial, participants indicated which of eight possible shapes completed a matrix of geometric patterns by clicking on it with a mouse. Items could be skipped, and were then shown again at the end of the test with the option to click an "I don't know" button. Participants had 40 min to complete 36 items. The time was indicated at the right top corner of the screen. A participant's score was the percentage of correct responses.

Results

Figure 2 shows the time courses of language-mediated eye movements. The main analysis focused on the lead-in sentence up to the critical spoken word. The choice of time window was stimulus-driven and was constrained by the facts that (1) the sentences varied in duration and (2) it takes an average of about 200 ms to program and initiate a saccadic eye movement (Saslow, 1967). Thus, we chose a time window
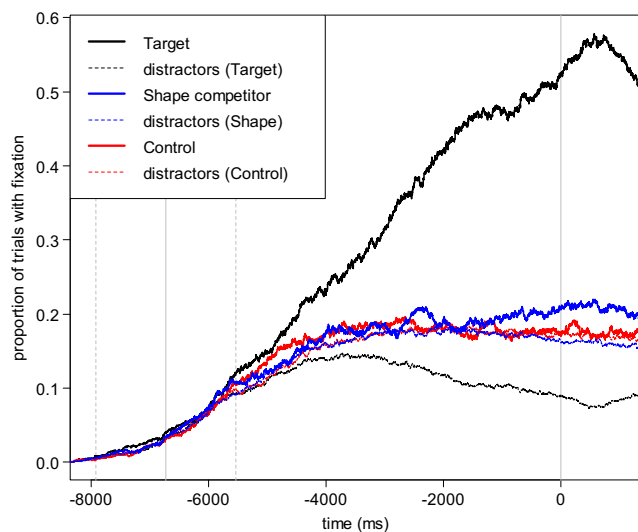
Fig. 2 Language-mediated eye movements: Time-course graph showing fixation proportions to targets, shape competitors, and control objects (solid lines), along with fixation proportions averaged across the three corresponding unrelated distractors (dashed lines). The *y*-axis reflects the proportion of trials in which participants were fixating each type of object. The leftmost three vertical gray lines indicate the average (solid line) and standard deviation (dashed lines) of the sentence onset time. Time zero indicates critical-word onset. Anticipatory effects are visible as lines of the same color diverging before time zero. Participants tended to fixate upon the targets much more than on the competitors or distractors, several seconds before critical-word onset. They were also more likely to fixate upon the shape competitors than upon the unrelated control objects and distractors, especially during the last second before critical word onset

from 200 ms after the item-specific sentence onset until the critical-word onset for analysis. We further examined the time course of the effects by separately analyzing four temporally adjacent 1,000-ms time windows, from 4,000 ms prior to critical-word onset up to critical-word onset. These time windows fell within the duration of the shortest sentence context leading up to the critical spoken word (4,006 ms). The fixation proportions within each time window were transformed to log odds using the empirical logit function (Barr, 2008). The log odds of looks averaged across the three unrelated distractors was subtracted from the log odds of looks to the target/shape/control object to create the dependent variable, which indicated the strength of any bias toward each experimental object over the unrelated distractors. This variable was analyzed using mixed-effects regression models, which allow for simultaneous inclusion of participants and items as random factors (Baayen, Davidson, & Bates, 2008). The fixed effect of display type (control, shape competitor, target) was included along with random intercepts and slopes by participants and by sentences (Barr, Levy, Scheepers, & Tily, 2013), and evaluated by comparing this model to a baseline model with the same random-effects structure, but without the fixed effect of display type. The *p* values associated with the contrasts between the control condition and each of the other two conditions were calculated assuming that the *t* values were drawn from a normal distribution (Barr, 2008).

The results from the full-sentence analysis yielded an effect of display type, $\chi^2(2) = 74.09$, $p < .001$. This was due to a selective bias toward target and shape competitor objects over distractors, since no bias toward experimental objects was detected for the displays with control objects, $\beta = -0.04346$, $SE = 0.04741$, $t = -0.917$, $p = .359$. The bias toward experimental objects was greater for shape competitors than for control objects, $\beta = 0.18240$, $SE = 0.08282$, $t = 2.202$, $p = .028$, and in turn much greater for targets than for shape competitors, $\beta = 2.18502$, $SE = 0.15675$, $t = 13.940$, $p < .001$.

Time window analyses are shown in Table 1. In accordance with what Fig. 2 suggests, these analyses did not yield a clear bias in looks toward the shape competitor until the time window starting at 1,000 ms before critical word onset. In this window, the shape competitor was looked at more than the unrelated distractors in the same display (indicated by the intercept), and that bias was larger than any bias toward the control object (indicated by the control vs. shape term). In contrast, the bias toward the target object was present in all time windows and was consistently greater than the bias toward the shape competitor.

Group-level results for the Posner cueing task are shown in Fig. 3. From this and the other background tasks, four

Table 1 Time window analysis of the eye movements across the three conditions

Predictor	Time Window	β	SE	t	p
Intercept	–4,000 ms	0.12594	0.11881	1.060	.289
(shape competitor)	–3,000 ms	0.1279	0.1060	1.206	.228
	–2,000 ms	0.07508	0.12895	0.582	.560
	–1,000 ms	0.3292	0.1254	2.625	.009
Control vs. shape	–4,000 ms	–0.02699	0.17392	–0.155	.877
	–3,000 ms	–0.1297	0.1482	–0.875	.381
	–2,000 ms	–0.09630	0.17057	–0.565	.572
	–1,000 ms	–0.4351	0.1620	–2.686	.007
Target vs. shape	–4,000 ms	1.47063	0.24858	5.916	<.001
	–3,000 ms	2.8176	0.2963	9.510	<.001
	–2,000 ms	4.16680	0.29580	14.087	<.001
	–1,000 ms	4.4077	0.3073	14.342	<.001

Time window values indicate the beginning of a window; each window lasted 1,000 ms. All degrees of freedom associated with the χ^2 tests were 1. No bias toward the control object over the unrelated distractors was detected in any time window (not shown in table; all $ps > .264$)

Table 2 Group performance on the background measures

Task	Measure	Mean	SD
Raven's Advanced Progressive Matrices	Answers correct (%)	67	15
Posner spatial cueing	Neutral – valid (ms)	25	19
Peabody Picture Vocabulary Test	Raw score	172	11
Category fluency	Correct responses	19.9	4.2

("pbody"). This was of little concern because, as reported below, these two measures did not appear together as reliable predictors in the same final models. This held true when the predictors had first been orthogonalized by residualizing; the models using the original predictors are reported (see Wurm & Fisicaro, 2014; York, 2012, for arguments in favor of using original predictors).

In the analyses involving the background measures, models were fit separately for the target, shape, and control displays. All background measures were mean-centered, which makes the intercept interpretable as the grand mean. A backward elimination procedure was used, starting from an initial model where all background measures were entered as fixed effects, as well as by-participants and by-sentences random intercepts. Nonsignificant predictors were then removed on the basis of p values calculated assuming that the t values were drawn from a normal distribution (Barr, 2008). For the remaining fixed effects, maximal by-sentence random slopes were fitted and evaluated using model comparisons.

The results from the whole-sentence analysis are shown in Table 3. For anticipatory target bias, removal of the Ravens and cueing predictors ($ps > .213$) did not affect the model fit, $\chi^2(2) = 1.9561$, $p = .376$. Participants looked more at the target object (e.g., the moon) than at the unrelated distractors in the same display, as indicated by the intercept. Furthermore, greater anticipatory bias was associated with high verbal fluency scores and large vocabulary scores.

Regarding the anticipatory shape competitor bias, the verbal fluency and Peabody vocabulary predictors could be removed without affecting the model fit, $\chi^2(2) = 1.241$, $p = .538$. As is summarized in Table 3, participants fixated shape competitors more often than unrelated distractors, as is indicated by the intercept. This effect was stronger for individuals who scored relatively low on the Raven's intelligence test and for individuals whose responses were strongly facilitated by valid cues in the Posner cueing task.

The control condition showed no evidence of a preference for the control objects over the unrelated distractors (all $ps > .132$, including the intercept). The background

measures were derived (see Table 2; the correlations between all measures after averaging by participants are shown in the Appendix). The only reliable correlation between the background measures suggested that participants with higher fluid intelligence ("ravens") also had larger vocabularies
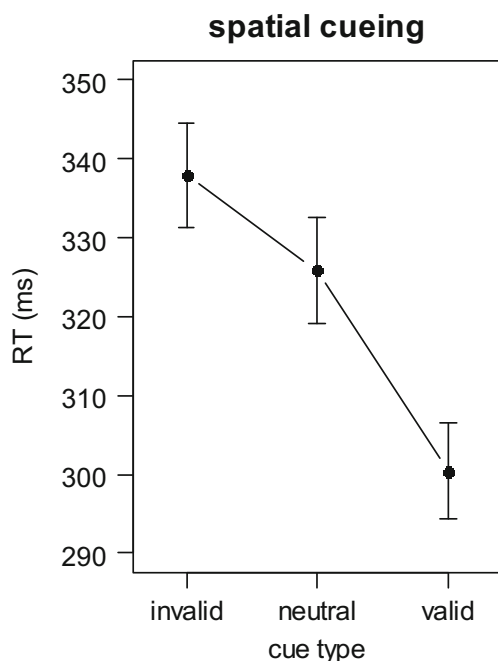
Fig. 3 Average response latencies in the spatial cueing task. Error bars indicate standard errors of the means. Response latencies were longest after invalid cues, shortest after valid cues, and intermediate after neutral cues (all $ts > 4.212$, $ps < .001$). The orienting measure was obtained by subtracting response times after valid cues from those after neutral cues

Table 3 Whole-sentence analysis of the language-mediated anticipatory eye movements with the background measures as predictors

Condition	Term	β	SE	t	χ^2	p
Target	Intercept	2.32441	0.13959	16.649	136.19	<.001
	Verbal fluency	0.05686	0.02914	1.951	3.8469	.050
	Peabody	0.02439	0.01173	2.078	4.3224	.038
Shape competitor	Intercept	0.13798	0.06772	2.037	4.08	.043
	Ravens	−0.00846	0.00320	−2.647	6.996	.008
	Cueing	0.006024	0.00265	2.278	5.003	.025
Control	Intercept	−0.04349	0.07149	−0.608	0.3734	.541

All degrees of freedom associated with the χ^2 tests were 1

measures did not contribute to explaining the variance in these data, $\chi^2(4) = 3.6105$, $p = .461$.

More detailed time window analyses of the relationships between the background measures and the looks to target objects and shape competitors are reported in Table 4 (for target objects) and Table 5 (for shape competitors). As can be seen from Table 4, the large bias toward the target objects (indicated by the intercept) was associated with high verbal fluency and large vocabulary size in several time windows, confirming the whole-sentence analysis. The β estimates for these verbal measures rose from early on, and the effects were statistically detectable in time windows starting from 3,000 or 2,000 ms prior to the onset of the critical spoken word.

As can be seen from Table 5, the smaller bias toward shape competitors only became evident in a later time window, encompassing the last second prior to the critical spoken-word onset, consistent with the results shown in Table 1. The modulation of this bias as a function of the background measures was visible in a window 2 to 3 s before word onset—that is, when the effect had not yet become detectable at the group level.

Discussion

The participants preferentially fixated the target objects (e.g., a moon) before the target name had been mentioned, consistent with the idea that they predicted the objects to be referred to. We also observed a later and smaller but reliable anticipatory bias toward objects with a shape similar to the upcoming referent (e.g., a tomato). These results replicate our previous findings and are consistent with the notion that predictions can involve the referent's shape.

Our main interest was in relating the participants' tendency to fixate target objects and shape competitors to their performance on verbal and nonverbal tasks. The results dissociated the two types of anticipatory looks from each other. A high proportion of anticipatory looks

Table 4 Time window analyses predicting bias toward the target objects on the basis of the background measures

Predictor	Time Window	β	SE	t	χ^2	p
Intercept	−4,000 ms	1.5966	0.2082	7.669	47.078	<.001
	−3,000 ms	2.94647	0.23548	12.512	103.3	<.001
	−2,000 ms	4.24649	0.25777	16.474	140.98	<.001
	−1,000 ms	4.73612	0.28538	16.596	130.59	<.001
Verbal fluency	−4,000 ms	0.015139	0.026996	0.561	–	(.575)
	−3,000 ms	0.053970	0.037236	1.449	–	(.147)
	−2,000 ms	0.08838	0.05157	1.714	2.7054	.100
	−1,000 ms	0.12925	0.06267	2.062	4.2906	.038
Peabody	−4,000 ms	0.011686	0.011879	0.984	–	(.325)
	−3,000 ms	0.03965	0.01511	2.623	6.7202	.010
	−2,000 ms	0.04673	0.02002	2.335	5.1213	.024
	−1,000 ms	0.04702	0.02510	1.873	3.5554	.059

Time window values indicate the beginning of a window; each window lasted 1,000 ms. The rows containing a χ^2 statistic contain estimates from the final model of the data in that time window. In the other rows, for completeness, estimates from initial models are shown for predictors that were excluded prior to the model comparison stage (based on p values assuming a normal distribution, shown between brackets)

Table 5 Time window analyses predicting bias toward the shape competitor objects on the basis of the background measures

Predictor	Time Window	β	SE	t	χ^2	p
Intercept	–4,000 ms	0.1259	0.1187	1.061	1.1323	.287
	–3,000 ms	0.127487	0.125429	1.016	1.0415	.308
	–2,000 ms	0.076590	0.130192	0.588	0.3528	.553
	–1,000 ms	0.3292	0.1270	2.593	6.5211	.011
Ravens	–4,000 ms	–0.003940	0.006169	–0.639	–	(.523)
	–3,000 ms	–0.012242	0.006257	–1.957	3.7927	.051
	–2,000 ms	–0.010161	0.006281	–1.618	–	(.106)
	–1,000 ms	–0.001011	0.006808	–0.149	–	(.882)
Cueing	–4,000 ms	0.001139	0.005006	0.228	–	(.820)
	–3,000 ms	0.007526	0.005109	1.473	–	(.141)
	–2,000 ms	0.016499	0.005162	3.196	9.3587	.002
	–1,000 ms	0.006568	0.005516	1.191	–	(.234)

Time window values indicate the beginning of a window; each window lasted 1,000 ms. The rows containing a χ^2 statistic contain estimates from the final model of the data in that time window. In the other rows, for completeness, estimates from initial models are shown for predictors that were excluded prior to the model comparison stage (based on p values assuming a normal distribution, shown between brackets)

to the target objects was associated with high vocabulary scores and high category fluency, highlighting the role of lexical knowledge and processes in the task (see James & Watson, 2013; Mani & Huettig, 2014; and Mishra, Singh, Pandey, & Huettig, 2012, for evidence of a link between literacy, experience, and predictive language processing). Time window analyses confirmed the association of the verbal background measures with the eye movements early in time, 2–3 s before critical-word onset. Whereas previous reports that had associated prediction with verbal fluency relied on event-related brain potential modulations occurring after word onset (e.g., Federmeier et al., 2002), the present study linked verbal fluency performance to processes occurring before the acoustic onset of predictable words. To the extent that the verbal fluency task indexes language production processes, the relationship with anticipatory eye movements is consistent with frameworks that suggest that the language production system generates predictions during comprehension (Chang, Dell, & Bock, 2006; Dell & Chang, 2014; Federmeier, 2007; Mani & Huettig, 2012; Pickering & Garrod, 2007, 2013; van Berkum et al., 2005).

In contrast to the results for the target objects, a high proportion of anticipatory looks to the shape competitors was associated with sensitivity to cues in the Posner cueing task. This relationship was observed despite the differences between the tasks in terms of the predictive cues (visually presented arrows vs. various aspects of linguistic context), cue validity (80 % vs. 50 %), and the contents of the predictions (the location of an X symbol vs. an aspect of an upcoming referent). Time window analyses confirmed this relationship

within a time window more than 1 s before critical spoken-word onset. This window coincided with the early part of the shape effect, possibly representing data from the fastest participants, before the effect became statistically detectable at the group level. Further investigations could examine whether individual differences in spatial cueing tasks are related to the timing rather than the strength of biases in language-mediated anticipatory eye movements. Overall, in Underwood's (1975) terms, this result provides a go-ahead signal for the idea that anticipatory attentional mechanisms are shared between verbal and nonverbal tasks.

Regarding the limitations of this study, it should be noted that the cueing response times are related to the effectiveness with which the attentional system prepares for detecting stimuli at attended locations (Posner, 1980). Thus, the link to prediction is somewhat indirect. However, the results converge with recent evidence showing similar anticipatory eye movement behavior in action observation and in listening to sentences about those actions (Poljac, Dahlslätt, & Bekkering, 2014). Because the data are correlational, further research will be necessary to determine exactly what is shared between the tasks. One possibility is that the underlying mechanisms are shared, such that overlapping processes are responsible for predictive processing in both verbal and nonverbal tasks. On the other hand, it is also possible that a mediating factor influenced performance on both tasks. It is clear, however, that what is shared between the tasks is not linguistic. To explore the abovementioned possibilities, predictive mechanisms that gain substantial empirical support in attention, perception, or action research should also be tested in language comprehension studies (see also Kutas, DeLong, & Smith, 2011; van Berkum, 2010).

A somewhat surprising aspect of our data was that a relatively high tendency to look at shape competitors was associated with lower, not higher, fluid intelligence, as measured by the Raven's test. A potential explanation builds on the known relationship between fluid intelligence and working memory capacity (measurements of these abilities estimate the amount of shared variance at 50 %; Kane, Hambrick, & Conway, 2005). It has been proposed that working memory capacity reflects an executive–attention control mechanism that helps to maintain goal-relevant information in high-interference conditions in which potent distractors need to be suppressed (Conway, Kane, & Engle, 2003; Engle & Kane, 2004). Individual differences in cognitive control partly mediate the relationship between working memory capacity and fluid intelligence (Unsworth & Spillers, 2010; Unsworth, Spillers, & Brewer, 2009). Thus, cognitive control could explain why individuals with lower fluid intelligence may have struggled to ignore the distracting shape competitor in our experiment. Although the actual inclusion of measures of cognitive control awaits further research, note that interpretations along these lines maintain the nonlinguistic nature of the looks to shape competitors proposed above.[1]

In general, the difference in magnitude that we observed between the looks to the target objects and the looks to the shape competitors was striking. One might argue that, because the sentences varied and contained words associated with the target object, the looks to target objects could (partly) reflect semantic relationship effects (e.g., looks to a trumpet when hearing "piano"; Huettig & Altmann, 2005). We cannot rule this out, and looks to the target object could therefore be argued to form a less strong index of predictive processing than do looks to shape competitors. Another possibility is that this could be a purely statistical phenomenon. Target objects such as the moon overlapped with many possible features of the concept *moon* that could have become preactivated, whereas the tomato only overlapped in visual shape. Thus, any aspect of the moon that becomes preactivated could bias overt attention toward the moon, but only shape properties would bias overt attention toward the tomato. Even if all aspects of concepts were equally likely to become preactivated, the

shape competitor effect should be smaller. However, the effects appear to differ not only in magnitude but also in timing, with the shape effect emerging much later than the target effect. An interesting possibility is therefore that shape information tended to influence eye movements later because the contents of predictions differ over time, developing from general to more specific as contextual evidence from the spoken sentence accumulates. However, this attributes both types of anticipatory looks to similar mechanisms, which is difficult to reconcile with the dissociations revealed by the patterns of individual differences.

We therefore suggest instead that the dissociation between the results seen for targets and shape competitors in terms of strength, timing, and individual differences supports the idea that different mechanisms underlie different kinds of predictive language processing (see also Mani & Huettig, 2013). This is consistent with research using event-related brain potentials, where it has been observed that different components index the disconfirmation of lexical versus conceptual expectations (Thornhill & van Petten, 2012). Despite the initial support for a relationship between predictive processing across verbal and nonverbal domains, the results do not support the full-fledged domain-generality of a single prediction mechanism, because performance in the nonverbal predictive task was related only to anticipatory eye movements to shape competitors, not to targets. Taken together, the results illustrate the different contributions of verbal and nonverbal processes to language-mediated anticipatory eye movements.

Appendix: Correlations

Figure 4 shows the correlations between all background measures and the eyetracking data in the whole-sentence time window, after averaging within participants. Consistent with the mixed-effects regression models, anticipatory target bias correlated positively with the Peabody vocabulary and the verbal fluency scores, whereas anticipatory shape bias correlated positively with cueing and negatively with the Raven's scores. The correlation between shape bias and the cueing task is unlikely to be due to individual differences in the neutral cueing condition, because residual response times in the valid condition after regressing it onto the neutral condition (cf. DeGutis, Wilmer, Mercado, & Cohan, 2013) yielded the same absolute correlation rho of .25 as in Fig. 4. Other correlations between the eyetracking data and the background measures

---

[1] One might wonder to what extent the spatial cueing task also measures attention control—for instance, within the Attention Network Test (Fan et al., 2002). Indeed, within this framework, executive control can be quantified by measuring response time differences between congruent and incongruent high-conflict flanker stimuli. However, such high-interference situations are different from the valid-versus-neutral cue contrast that we used, which results in a measure called *orienting*. Accordingly, our explanation of the correlation involving the Raven's task differs from the explanation for the correlation involving the cueing task.

did not reach significance, although we did see a trend toward Raven's performance correlating with the control-object bias, $p = .092$ (all other ps $> .178$).
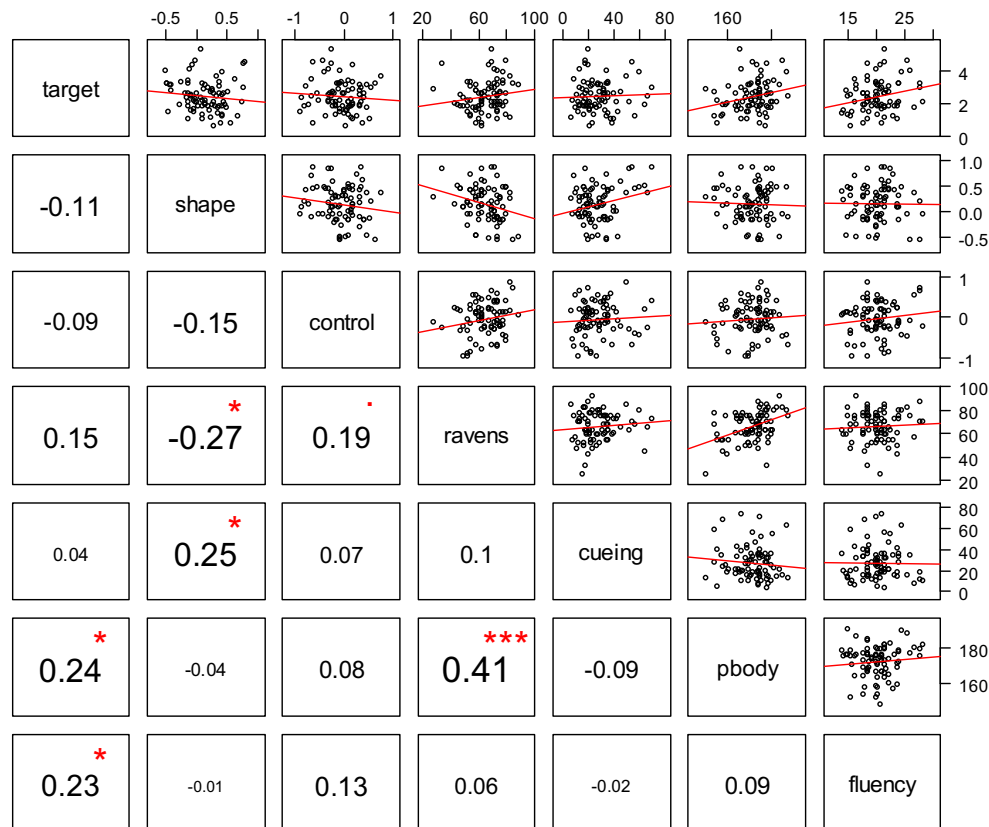
Fig. 4 Correlations between interindividual differences in language-mediated anticipatory eye movements and the background measures. Lower left panels: Pearson correlation coefficients for each pair of measures, with the font size scaled to the absolute rho value. Upper right panels: Scatterplots of each measure plotted against each other measure, along with linear regression lines. "target," target bias (target minus distractors); "shape," shape competitor bias; "control," control-object bias; "ravens," Raven's Advanced Progressive Matrices score; "cueing," facilitation from valid cues in the Posner cueing task; "pbody," Peabody picture vocabulary size; "fluency," category fluency number of correct responses. $^{***}p < .001,\ ^{*}p < .05,\ ^{.}p < .1$

References

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247–264. doi:10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33,* 583–609. doi:10.1111/j.1551-6709.2009.01022.x

Arai, M., & Keller, F. (2013). The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes, 28,* 525–560.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59,* 390–412. doi:10.1016/j.jml.2007.12.005

Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences, 11,* 280–289. doi:10.1016/j.tics.2007.05.005

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59,* 457–474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278. doi:10.1016/j.jml.2012.11.001

Borovsky, A., Elman, J., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with the timecourse of incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112,* 417–436.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113,* 234–272. doi:10.1037/0033-295X.113.2.234

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181–204. doi:10.1017/S0140525X12000477

Cohen, R. L. (1994). Some thoughts on individual differences and theory construction. *Intelligence, 18,* 3–13.

Coles, M. G. H. (1989). Modern mind-brain reading: Psychophysiology, physiology, and cognition. *Psychophysiology, 26,* 251–269.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7,* 547–552. doi:10.1016/j.tics.2003.10.005

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition, 126,* 87–100. doi:10.1016/j.cognition.2012.09.004

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Verbal fluency subtest of the Delis–Kaplan Executive Function System.* San Antonio, TX: Psychological Corp.

Dell, G. S., & Chang, F. (2014). The P-Chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B, 369,* 20120394.

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117–1121. doi:10.1038/nn1504

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 44, pp. 145–199). San Diego, CA: Elsevier Academic Press.

Fan, J., McCandliss, B. D., Sommer, T., Raz, M., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience, 14,* 340–347. doi:10.1162/089892902317361886

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology, 44,* 491–505. doi:10.1111/j.1469-8986.2007.00531.x

Federmeier, K. D., Kutas, M., & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain and Language, 115,* 149–161. doi:10.1016/j.bandl.2010.07.006

Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39,* 133–146. doi:10.1017/S0048577202001373

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11,* 127–138.

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

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica, 137,* 151–171. doi:10.1016/j.actpsy.2010.11.003

James, A., & Watson, D. (2013, September). *Language experience and prediction in auditory sentence comprehension.* Poster presented at the International Conference on Architectures and Mechanisms for Language Processing (AMLaP), Marseille, France.

Kamide, Y. (2008). Anticipatory processes in sentence processing. *Language and Linguistics Compass, 2,* 647–670.

Kamide, Y., Scheepers, C., & Altmann, G. T. M. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research, 32,* 37–55.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66–71. doi:10.1037/0033-2909.131.1.66

Klein, R. M. (1994). Perceptual–motor expectancies interact with covert visual orienting under conditions of endogenous but not exogenous control. *Canadian Journal of Experimental Psychology, 48,* 167–181. doi:10.1037/1196-1961.48.2.167

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition, 95,* 95–127. doi:10.1016/j.cognition.2004.03.002

Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., & Rose, R. (2002). Bridging psychology and biology: The analysis of individuals in groups. *American Psychologist, 57,* 341–351. doi:10.1037/0003-066X.57.5.341

Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the brain: Using our past to generate a future* (pp. 190–207). New York, NY: Oxford University Press.

Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology, 32,* 4–18.

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—but only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 843–847.

Mani, N., & Huettig, F. (2013). Towards a complete multiple-mechanism account of predictive language processing [Commentary on Pickering & Garrod]. *Behavioral and Brain Sciences, 36,* 365–366. doi:10.1017/S0140525X12002646

Mani, N., & Huettig, F. (2014). Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. *Journal of Experimental Child Psychology, 126,* 264–279.

Mishra, R. K., Singh, N., Pandey, A., & Huettig, F. (2012). Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates. *Journal of Eye Movement Research, 5*(1), 1–10.

Nobre, A. C., Rohenkohl, G., & Stokes, M. G. (2012). Nervous anticipation: Top-down biasing across space and time. In M. I. Posner (Ed.), *Cognitive neuroscience of attention.* New York, NY: Guilford Press.

Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11,* 105–110. doi:10.1016/j.tics.2006.12.002

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36,* 329–347. doi:10.1017/S0140525X12001495

Poljac, E., Dahlslätt, K., & Bekkering, H. (2014). Shared predictive decision-making mechanisms in action and language. *Language, Cognition and Neuroscience, 29,* 424–434.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25. doi:10.1080/00335558008248231

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role of set for spatial location. In H. L. Pick Jr. & E. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Raven manual section 4: Advanced progressive matrices.* Oxford, UK: Oxford Psychologists Press.

Rommers, J., Meyer, A. S., Praamstra, P., & Huettig, F. (2013). The contents of predictions in sentence comprehension: Activation of the shape of objects before they are referred to. *Neuropsychologia, 51,* 437–447.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57,* 1030–1033.

Schlichting, L. (2005). *Peabody Picture Vocabulary Test-III-NL.* Amsterdam, The Netherlands: Harcourt Assessments BV.

Thornhill, D. E., & van Petten, C. (2012). Lexical versus conceptual anticipation during sentence processing: Frontal positivity and N400 ERP components. *International Journal of Psychophysiology, 83,* 382–392.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist, 30,* 128–134. doi:10.1037/h0076759

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language, 62,* 392–406. doi:10.1016/j.jml.2010.02.001

Unsworth, N., Spillers, G. J., & Brewer, G. (2009). Examining the relations among working memory capacity, attention control, and fluid intelligence from a dual-component framework. *Psychological Science Quarterly, 51,* 388–402.

van Berkum, J. J. A. (2010). The brain is a prediction machine that cares about good and bad—Any implications for neuropragmatics? *Italian Journal of Linguistics, 22,* 181–208.

van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467. doi:10.1037/0278-7393.31.3.443

van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83,* 176–190. doi:10.1016/j.ijpsycho.2011.09.015

Vogel, E. K., & Awh, E. (2008). How to exploit diversity for scientific gain: Using individual differences to constrain cognitive theory. *Current Directions in Psychological Science, 17,* 171–176.

Wang, H., Fan, J., & Johnson, T. R. (2004). A symbolic model of human attentional networks. *Cognitive Systems Research, 5,* 119–134.

Weber, A., Grice, M., & Crocker, M. W. (2006). The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements. *Cognition, 99,* B63–B72.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does *not* do). *Journal of Memory and Language, 72,* 37–48. doi:10.1016/j.jml.2013.12.003

York, R. (2012). Residualization is not the answer: Rethinking how to address multicollinearity. *Social Science Research, 41,* 1379–1386.