



## “Proactive” in many ways: Developmental evidence for a dynamic pluralistic approach to prediction

Nivedita Mani, Moritz M Daum & Falk Huettig


To cite this article: Nivedita Mani, Moritz M Daum & Falk Huettig (2016): “Proactive” in many ways: Developmental evidence for a dynamic pluralistic approach to prediction, The Quarterly Journal of Experimental Psychology, DOI: [10.1080/17470218.2015.1111395](https://doi.org/10.1080/17470218.2015.1111395)

To link to this article: <http://dx.doi.org/10.1080/17470218.2015.1111395>



Accepted author version posted online: 23 Nov 2015.  
Published online: 01 Apr 2016.



[Submit your article to this journal](#) 



Article views: 22



[View related articles](#) 



[View Crossmark data](#) 

# “Proactive” in many ways: Developmental evidence for a dynamic pluralistic approach to prediction

Nivedita Mani<sup>1</sup>, Moritz M Daum<sup>3</sup>, and Falk Huettig<sup>2,4</sup>

<sup>1</sup>Psychology of Language Research Group, University of Goettingen, Goettingen, Germany

<sup>2</sup>Donders Institute for Brain, Cognition, and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

<sup>3</sup>Department of Psychology, University of Zurich, Zurich, Switzerland

<sup>4</sup>Department of Psychology of Language, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

(Received 24 March 2015; accepted 18 August 2015; first published online 8 December 2015)

The anticipation of the forthcoming behaviour of social interaction partners is a useful ability supporting interaction and communication between social partners. Associations and prediction based on the production system (in line with views that listeners use the production system covertly to anticipate what the other person might be likely to say) are two potential factors, which have been proposed to be involved in anticipatory language processing. We examined the influence of both factors on the degree to which listeners predict upcoming linguistic input. Are listeners more likely to predict *book* as an appropriate continuation of the sentence “The boy reads a”, based on the strength of the association between the words *read* and *book* (strong association) and *read* and *letter* (weak association)? Do more proficient producers predict more? What is the interplay of these two influences on prediction? The results suggest that associations influence language-mediated anticipatory eye gaze in two-year-olds and adults only when two thematically appropriate target objects compete for overt attention but not when these objects are presented separately. Furthermore, children’s prediction abilities are strongly related to their language production skills when appropriate target objects are presented separately but not when presented together. Both influences on prediction in language processing thus appear to be context dependent. We conclude that multiple factors simultaneously influence listeners’ anticipation of upcoming linguistic input and that only such a dynamic approach to prediction can capture listeners’ prowess at predictive language processing.

**Keywords:** Prediction; Associative strength; Vocabulary development; Production; Children.

Children and adults show an impressive ability to comprehend speech despite the rapidity with which speech input is directed at the listener. This ability derives, in large part, from the fact that listeners anticipate upcoming input based on the spoken input provided thus far, which allows

them to run ahead of the speaker in order to ease the demands on their comprehension of rapidly changing speech stimuli (Borovsky, Elman, & Fernald, 2012; Kamide, Altmann, & Haywood, 2003; Keitel, Prinz, Friederici, von Hofsten, & Daum, 2013; Mani & Huettig, 2012; Nation,

---

Correspondence should be addressed to Nivedita Mani, Psychology of Language Research Group, University of Goettingen, Gosslerstrasse 14, Goettingen 37073, Germany. E-mail: nmani@gwdg.de

This work was funded by the German Initiative of Excellence (University of Göttingen Institutional Strategy).

Marshall, & Altmann, 2003; see Federmeier, 2007; Huettig, 2015). Mani and Huettig (2012), for instance, show that even two-year-olds anticipate upcoming spoken language input. In their study, children were presented with two familiar images—for example, a book and a slice of cheese—and heard the sentence “The boy reads the book”. Analysis of children’s eye movements across the two images suggests that children show increased fixations to the image of the book, shortly after hearing the thematically appropriate word “read”, prior to even hearing the word “book”. The authors interpret these results as support for the notion that young children quickly anticipate upcoming linguistic input—that is, upon hearing “read”, children retrieve information consistent with this word—for example, the word “book” that is likely to be introduced into the discourse context—and use this to fixate related images. In attempting to explain this behaviour, prominent models of language processing and cognition consider the influence of different factors on our ability to predict upcoming linguistic input. Two main factors that have been proposed in the literature are associations from past experience (e.g., Bar, 2007) and prediction based on language production mechanisms (e.g., Chang, Dell, & Bock, 2006; Dell & Chang, 2013; Pickering & Garrod, 2007, 2013). The current study examines the extent to which our prediction behaviour is influenced by these two factors. In particular, we examine the relative influences of (a) the associative strength of the relationship between linguistic stimuli, (b) participants’ production skills in predicting upcoming linguistic input, and (c) whether language context determines which factors drive prediction.

Several authors have proposed that language comprehension piggybacks off language production to predict what a speaker is likely to say next and uses this to support comprehension of what the speaker does indeed say (Chang et al., 2006; Dell & Chang, 2013; Federmeier, 2007; Pickering & Garrod, 2007, 2013). Chang et al. (2006), for instance, argue that prediction is necessary for language learning. Learning occurs when the learner’s production-based predictions are compared to the actually produced language. The prediction

error (i.e., any discrepancies) is used to adjust the system that generated the predictions. Pickering and Garrod (2013) argue that speakers construct efference copies of their predicted productions. Speakers are assumed to compare these efference copies with the output of a production implementer—that is, what they actually say. Importantly, Pickering and Garrod propose that also during comprehension, listeners use these efference copies to predict what a speaker is likely to say next. A final related possibility is that comprehenders use the fully fledged production system for anticipation rather than the impoverished representations of a forward model (Huettig, 2015).

Support for an influence of production in prediction comes from Mani and Huettig (2012) who find that children’s ability to predict upcoming linguistic input is related to their production skills, such that children with larger production vocabularies are better predictors than children with smaller production vocabularies. Similarly, children’s productive vocabulary has an influence on the prediction of forthcoming action goals (Gampe & Daum, 2014).

A different type of cognitive model of prediction relies, on the other hand, on “proactive” memory-based predictions generated on the basis of current information, with an emphasis on the role of past associations driving the brain’s anticipation of the future (Bar, 2007, 2009; Kuperberg, 2007). According to this proposal, we constantly extract associative regularities and statistical information from our past experiences and use this to generate predictions about the sequences that follow a particular sensory experience. Thus, the frequency with which two sensory inputs co-occur in our memory influences prediction, such that the greater the relative frequency with which two items co-occur in our experience, the more likely we are to predict the occurrence of one item on the basis of the other.

Evidence for an important role for thematic associations from past experience in prediction comes from studies showing that participants are as likely to fixate potential agents as well as patients upon hearing a sentence such as “Toby arrests a ...”. Thus, despite *crook* being a contextually more appropriate continuation of this sentence, participants fixate a *crook* and a *policeman* equally in a visual

display, suggesting that anticipatory eye movements are at least partly driven by simple thematic associations rather than only by the contextual constraints imposed by the sentence or event (Kukona, Fang, Aicher, Chen, & Magnuson, 2011).

Against this background, the current study examines the interplay between production-based influences and past associations on listeners' prediction behaviour in an attempt to scrutinize the validity of a dynamic pluralistic approach to prediction, where different factors weigh in on listeners' prediction based on the specifics of the situation (Huettig, 2015; Mani & Huettig, 2013). In particular, we presented children and adults with sentences such as "The boy reads...", as they viewed two images where either (a) one of the images was strongly associated with the verb "read" (e.g., book), while the other image was a thematically inappropriate continuation of the phrase, or (b) one of the images was weakly associated with the verb "read" (e.g., letter), while the other image was a thematically inappropriate continuation of the phrase, or (c) one of the images was strongly associated with the verb "read", while the other image was weakly associated with the verb "read". Thus, there were no semantic differences between the appropriateness of the two associates with regards to the verb: One can read a book in much the same way that one can read a letter. Support for an influence of associative strength on prediction would come from either (a) the finding that participants fixate the thematically appropriate image more in the strong association condition than in the weak association condition or (b) the finding that participants fixate the strongly associated image more than the weakly associated image in conflict trials. Support for an influence of production skills on prediction would be a replication of Mani and Huettig's (2012) finding that the size of the prediction effect displayed by individual children—that is, the extent to which they fixate the thematically appropriate image—correlates uniquely with their production vocabulary size. Finally, by examining the extent to which associative effects can be disassociated from production effects within each condition, we can disentangle the separable influences of the two factors on anticipation of upcoming linguistic input.

## EXPERIMENTAL STUDY

### Method

#### *Participants*

Data from 30 children at 2 years of age ( $M = 24.05$  months, range = 23.13 to 25.63 months, 16 girls) were included in the final analysis, which gave us data from five children for each of the six versions of the experiment for counterbalancing. Data from five children were excluded from analysis due to their not providing data for all conditions. Children came from a sample of families who responded to an invitation letter sent to all families with infants living in the area. Parents gave informed consent for participation of their child in the study and received a T-shirt or a book as a token for their participation in the research.

In addition, to ensure the reliability of the results across development, we also tested 26 adults (aged between 18 and 32 years, 16 female) on the same experiment. Of these, data from two adults were excluded due to their not wearing glasses during testing despite having vision problems. This allowed us to analyse the data from 24 adults as is typical in experiments examining similar issues. Adults were psychology students who received course credits for their participation in the experiment.

#### *Stimuli*

We chose 30 verb–noun pairs for the critical conditions, with 10 verb–noun pairs per condition (see Table 1 for a full list of stimuli). The 30 verb–noun pairs were created by combining 10 verbs with three different nouns each. The verbs were inserted into grammatically appropriate sentences as in (1) below

1. Der Junge liest gleich etwas

[The boy] [reads] [now] [something]

Thus, across stimuli presented to participants in the critical condition, the sentences never provided the child with information with regard to the appropriate noun for each verb. In six of 10 sentence-sets, the verb was followed by the words *gleich etwas* [now something], while in the other four

**Table 1.** *Verb–noun combinations that contributed to the different conditions*

<i>Verb</i>	<i>Strong association</i>	<i>Weak association</i>	<i>Unrelated</i>
Read [ <i>lesen</i> ]	Book [ <i>Buch</i> ]	Letter [ <i>Brief</i> ]	Cheese [ <i>Käse</i> ]
Feed [ <i>füttern</i> ]	Duck [ <i>Ente</i> ]	Deer [ <i>Hirsch</i> ]	Bus [ <i>Bus</i> ]
Wash [ <i>waschen</i> ]	Hair [ <i>Haare</i> ]	Trousers [ <i>Hose</i> ]	Radio [ <i>Radio</i> ]
Cut [ <i>schneiden</i> ]	Bread [ <i>Brot</i> ]	Apple [ <i>Apfel</i> ]	Car [ <i>Auto</i> ]
Play [ <i>spielen</i> ]	Ball [ <i>Ball</i> ]	Swing [ <i>Schaukel</i> ]	Sun [ <i>Sonne</i> ]
Eat [ <i>essen</i> ]	Cake [ <i>Kuchen</i> ]	Onion [ <i>Zwiebel</i> ]	Coat [ <i>Jacke</i> ]
Ride [ <i>fahren</i> ]	Train [ <i>Zug</i> ]	Tractor [ <i>Traktor</i> ]	Hat [ <i>Mütze</i> ]
Sit [ <i>sitzen</i> ]	Chair [ <i>Stuhl</i> ]	Cushion [ <i>Kissen</i> ]	Cup [ <i>Tasse</i> ]
Climb [ <i>klettern</i> ]	Slide [ <i>Rutsche</i> ]	Mountain [ <i>Berg</i> ]	Shoe [ <i>Schuh</i> ]
Sleep [ <i>schlafen</i> ]	Bed [ <i>Bett</i> ]	Sofa [ <i>Sofa</i> ]	Phone [ <i>Telefon</i> ]

sentence-sets, the verb was followed by verb-appropriate prepositions—for example, *mit* [with] and *auf* [on] and the words *gleich etwas* [now something].

In addition, we chose 10 verb–noun pairs for the filler conditions, with five verb–noun pairs per filler condition. The filler trials ensured that the child heard a thematically appropriate noun following the verb every three or four trials. Thus, sentences for the filler conditions were structured as in (2) below such that the verb and the adverb *gleich* [now] were followed by a disambiguating noun phrase.

2. Der Junge reitet gleich das Pferd

[The boy] [rides] [now] [the horse]

The critical sentences were then combined with stationary images of pairs of objects whose labels were strongly associated with the verb (e.g., *liest* [read]–*Buch* [book]), weakly associated with the verb (e.g., *liest* [read]–*Brief* [letter]), or unrelated to the verb (e.g., *liest* [read]–*Käse* [cheese]) to form 60 videos, of which 30 videos were presented to each participant. Filler sentences were, similarly, combined with stationary images of pairs of objects whose labels were either associated with the verb or unrelated to the verb to form 10 videos, of which all 10 videos were presented to each participant. All sentences had either “The boy” or “The girl” as the subject/agent, counterbalanced across participants—for example, half the participants heard “The boy now rides the horse” while the other half heard “The girl now rides the horse”. Images were coloured photographs of the objects, with

one image per word. Movies were formatted such that the target and distractor images (measuring 480 × 380 pixels each) appeared simultaneously to the left and right of the screen (1920 × 1084 pixels) with a separation of 320 pixels between them.

#### *Associative norms*

Associative relationships between verb–noun pairs were determined using the Noun Associations for German database (Melinger & Weber, 2006). As this database gives counts of verbs that were freely associated with the nouns, we collected separate associative rankings of the verb–noun pairs from two sets of participants. First, we asked parents of the children tested in the experiment to rank the nouns presented in order of the strength of their association to the verb as if they were responding for their children. For example, parents were asked how strongly their child would associate the noun *cake* with the word *eat*. Similarly, we asked a separate group of adults ( $n = 22$ ) to rank the nouns presented in order of the strength of their association to the verb. Both groups of participants were asked to provide a rank between 0 to 5 where 0 was “not associated”, while 5 was “highly associated”. Paired-samples  $t$ -tests found that both groups of participants ranked the strongly associated verb–noun pairs as more strongly associated with one another than either the weakly associated verb–noun pairs—parents:  $t(9) = 3.86$ ,  $p = .004$ , 95% confidence interval, CI [0.49, 1.88]; adults:  $t(9) = 3.77$ ,  $p = .004$ , 95% CI [0.45, 1.81]—or

the unrelated verb–noun pairs—parents:  $t(9) = 18.69$ ,  $p < .001$ , 95% CI [3.43, 4.38]; adults:  $t(9) = 16.56$ ,  $p < .001$ , 95% CI [3.52, 4.63]. Furthermore, participants ranked the weakly associated verb–noun pairs as more strongly associated with one another than the unrelated verb–noun pairs—parents:  $t(9) = 10.16$ ,  $p < .001$ , 95% CI [2.11, 3.32]; adults:  $t(9) = 5.82$ ,  $p < .001$ , 95% CI [1.81, 4.08].<sup>1</sup>

### Experimental set-up

During the experiment, gaze data from both eyes were recorded using a Tobii X120 eye tracker. Videos were presented in the middle of a 40" screen located immediately above the eye tracker. The eye tracker was set to record gaze data at 60 Hz with an average accuracy of 0.5° visual angle. The Tobii Studio package was used to present the videos to the children during the experiment. Prior to testing, we calibrated the gaze of each child using a 9-point calibration procedure, in which an attention-getter appeared in every position of a 3 × 3 grid of calibration points. The experiment started if eight or more points were successfully calibrated for at least one of the eyes.

### Procedure

The short movies were combined to form three different lists of trials. Children were randomly assigned to one of these lists. Each list included 30 test trials and 10 filler trials. The 30 critical trials were divided as follows: In 10 trials, the verb was presented with an image whose label was strongly associated with the verb and an image whose label was unrelated to the verb (strong association condition). In 10 trials, the verb was presented with an image whose label was weakly associated with the verb and an image whose label was unrelated to the verb (weak association condition). In 10 trials, the verb was presented with an image whose label was strongly associated with the verb and another image whose label was weakly associated to the verb (conflict

condition). Importantly, the same noun–verb associations were used across strongly associated and weakly associated trials, thereby allowing analyses of the extent to which children's fixations to the same image change across conditions. Thus, for instance *read–book–cheese* was presented in strong association trials, *read–letter–cheese* in weak association trials, and *read–book–letter* in conflict trials. The 10 filler trials were divided as follows: In five trials, the verb was presented with two images, both of whose labels were associated with the verb. In five trials, the verb was presented with an image whose label was associated with the verb and one image whose label was unrelated to the verb. The filler trials ensured that the child heard a thematically appropriate noun following the verb every three trials.

Each trial began with two images presented on screen in silence. The images remained on screen for the entire 7-s duration of the trial. A variable duration after the onset of the trial, the carrier sentence began such that the onset of the verb was at 2.000 ms, and the onset of the ambiguous noun *etwas* [something] in critical trials was at 4.000 ms. In filler trials, the onset of the verb remained at 2.000 ms into the movie while the onset of the disambiguating noun was between 4.225 ms and 5.527 ms into the movie.

To ensure that any effects were not due to specific verbs that were presented in the different conditions, the verbs presented in critical trials were repeated across conditions (strong association, weak association, conflict condition). For instance, children heard the sentence in [1] presented once with an image of a book and cheese (strong association), once with an image of a letter and cheese (weak association), and once with an image of a book and a letter (conflict condition). Furthermore, to ensure that fixations were not driven by other factors such as the salience of the images, children's familiarity with the items, or the frequency of the items, the images were presented in a neutral context for 2 s before the onset of the critical verb.

<sup>1</sup>The verb *klettern* [climb] was strongly associated with *Rutsche* [slide] and less strongly associated with *Berg* [mountain] for children, while the reverse was true for adults. Our assignment of items to conditions, therefore, took this into consideration and differed across the adult and child version of the experiment. Excluding this item does not change the pattern or the significance of the results.

To examine whether there were any effects of repetition, trials were divided into three blocks such that there was no repetition of verbs within a block. Across children, trials were divided into blocks such that a relatively equal number of trials per condition were presented in each block. Due to there being 10 trials per block, this meant that children were presented with four trials from one condition and three trials from the other two conditions in each block. The number of trials per condition within a block was counterbalanced across blocks and across children. Trials were pseudorandomized within blocks. Order of presentation of blocks was counterbalanced such that, across children, the same verb was presented in different conditions in different blocks. Each block also presented children with filler trials, which were inserted into fixed positions in each block such that the first trial of the experiment was always a filler trial, and children were presented with a filler trial every three trials. Side of presentation of images across conditions was counterbalanced across children such that images appeared equally often to the left and to the right in all conditions.

At the end of the experiment, parents of children filled out questionnaires about the associative strength of the item pairs presented to children in the experiment as well as the vocabulary section of German parental communicative inventory reports (*Fragebogen zur Frühkindlichen Sprachentwicklung: FRAKIS*; Szagun, Stumper, & Schramm, 2009—standardized for children between 1;6 and 2;6 years of age).

### Analysis

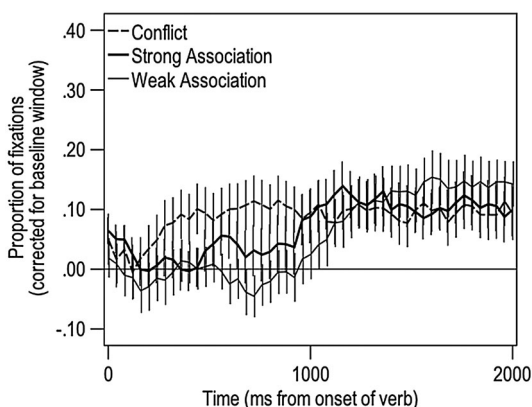
The eye tracker provides an estimate of where children were looking at for each time-stamp during the trial, with one data-point at 60 Hz (every 16 ms). Data from time-stamps were only included when the eye tracker reliably acquired data from one or both eyes of the participant (validity less than 2 on Tobii scale). Averaged data were used for data points where data from both eyes were available. For analysis, this data were further aggregated across two time windows. The first window,

the baseline window, counted all fixations (at least 60 ms) that occurred from the beginning of the trial to the onset of the verb at 2.000 ms. The second window, the verb window, counted all fixations that occurred 240 ms from the onset of the verb to 2.000 ms after the onset of the verb. The delay in the onset of the verb window ensures that only those eye movements that can reliably be attributed to the auditory stimulus are included in analysis presented here.

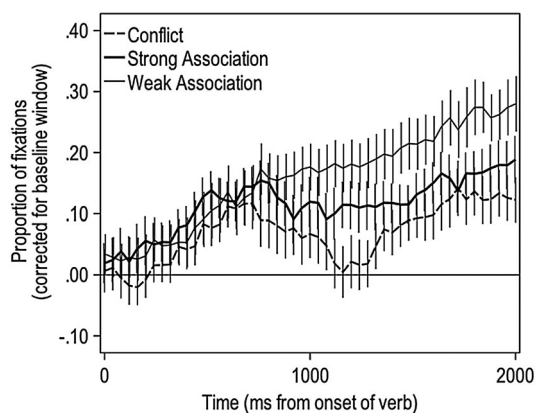
For each window, we determined the amount of time infants fixated the target (T) and distractor (D) images. We counted all fixations that participants made over the target and distractor images ( $\pm 50$  pixels to allow for problems with eye-tracker accuracy). We then calculated the proportion of time,  $T/(T + D)$ , that infants spent looking at the target in each window. This proportional measure was our dependent variable. Note that while there is an unambiguous target in both strongly and weakly associated trials (the other image is unrelated to the verb), there is no unambiguous target in conflict trials (both images are potentially targets). Thus, while we plot proportions of target fixations in strongly and weakly associated trials, we plot proportion of fixations to the strongly associated image in conflict trials. Given this difference, we analyse conflict trials separately from strongly and weakly associated trials.

### Results

Figure 1 plots the proportion of children's fixations to the target for the strongly and weakly associated trials (and fixations to the strongly associated image for conflict trials) for every 40 ms in the verb window (corrected for any preference in looking to this image in the baseline window). Figure 2 plots the same for the adult data. Thus, any data point with a score over 0 indicates a significant increase in looking to the critical image at this point relative to the baseline window. Correction for the baseline window accounts for any inherent preference that participants may have for one image over the other in each condition. Preliminary analyses examined whether there was any effect of repetition across



**Figure 1.** Children: Proportion of fixations to the target in strongly and weakly associated trials and to the strongly associated image in conflict trials throughout the verb window (corrected for fixations in the baseline window). Data points above 0 indicate an increase in fixations to a particular image from the baseline to the verb window.



**Figure 2.** Adults: Proportion of fixations to the target in strongly and weakly associated trials and to the strongly associated image in conflict trials throughout the verb window (corrected for fixations in the baseline window). Data points above 0 indicate an increase in fixations to a particular image from the baseline to the verb window.

blocks and found no statistical interaction between block and any other factors ( $ps > .2$ ). Subsequent analyses, therefore, collapsed the data across blocks.

First we examined whether children and adults showed an increase in looking towards the target—that is, the associated image—from the baseline window to the verb window in strongly and weakly associated trials. Thus, this analysis of variance (ANOVA) examined whether the increase in fixations between the baseline window and the verb window (plotted as a time-course in Figures 1 and 2) differed across the two conditions. A  $2 \times 2$  ANOVA with the factors condition (strong association; weak association) and window (baseline, verb) yielded a main effect of window [children:  $F(1, 29) = 12.56, p = .001, \eta_p^2 = .3$ ; adults:  $F(1, 23) = 34.62, p < .001, \eta_p^2 = .6$ ], a main effect of condition [children:  $F(1, 29) = 27.28, p < .001, \eta_p^2 = .48$ ; adults:  $F(1, 23) = 7.41, p = .012, \eta_p^2 = .24$ ], but no interaction between the two factors [children:  $F(1, 29) = 0.008, p = .9$ ;<sup>2</sup> adults:  $F(1, 23) = 1.82,$

$p = .19$ ], across both groups of participants. Note that we did not analyse performance in the noun window since the noun window does not provide any additional information in critical trials. In keeping with the main effect of window and the absence of an interaction between condition and window, planned paired-samples  $t$  tests found that participants looked longer at the associated target in the verb window than in the baseline window in both the strong association condition—children:  $t(29) = -2.28, p = .03, 95\% \text{ CI } [-.12, -.01]$ ; adults:  $t(23) = -4.69, p < .001, 95\% \text{ CI } [-.17, -.06]$ —and the weak association condition—children:  $t(29) = -2.66, p = .013, 95\% \text{ CI } [-.10, -.01]$ ; adults:  $t(23) = -4.99, p < .001, 95\% \text{ CI } [-.24, -.1]$ .

Next, we analysed performance in the conflict condition to examine whether there is an increase in children's preference for any one image from the baseline to the verb window. The data are

<sup>2</sup>Not all children were reported to know all the words presented to them in the experiment, at least according to parental reports. While parental reports have been shown to underestimate word knowledge (Houston-Price, Mather, & Sakkalou, 2007), we, nevertheless, carried out a reduced analysis removing any trials where children were reported not to comprehend either the target or distractor image or the verb. This reduced analysis yielded very similar results compared to the main analyses. In particular, there was a significant main effect of window in the  $2 \times 2$  ANOVA comparing the strong and the weak association condition,  $F(27) = 9.35, p = .005, \eta_p^2 = .26$ , with no significant interaction between condition and window,  $p > .8$ . There was also a significant increase in fixations to the strongly associated image from the baseline to the verb window in the conflict condition,  $t(27) = -2.92, p = .007$ .



analysed with regard to the strongly associated image, although we note that both images are potential targets in the verb window. Thus, any data point with a score over 0 indicates a significant increase in looking to strongly associated image at this point—in preference to the weakly associated image—relative to the baseline window. Despite no differences in participants' responding in strong association trials and weak association trials, when presented with the two associated images simultaneously, both adult and child participants show an increase in preference for the strongly associated image (in comparison to the weakly associated image) from the baseline window to the verb window—children:  $t(29) = -3.27$ ,  $p = .003$ , 95% CI  $[-.13, -.03]$ ; adults:  $t(23) = -2.92$ ,  $p = .008$ , 95% CI  $[-.13, -.02]$ .<sup>3</sup>

Finally, we examined the extent to which the prediction effects in each condition correlate with children's production vocabulary size as has been shown to be the case in previous studies examining prediction of linguistic input in children (Mani & Huettig, 2012).<sup>4</sup> Preliminary analysis showed that children's comprehension vocabulary size correlated strongly with their production vocabulary size. To examine the reality of a comprehension-independent influence of children's production abilities on their prediction skills, we analysed whether toddlers' language production skills correlate with their language prediction skills, once the shared variation between production and comprehension scores has been partialled out. The residualized production score refers to the amount of variability in PRODUCTION VOCABULARY SIZE that is left over after accounting for the variability shared with

children's COMPREHENSION VOCABULARY SIZE. Using the residuals output by regressing PRODUCTION VOCABULARY SIZE with COMPREHENSION VOCABULARY SIZE, we found a positive correlation between residualized production scores and children's looking times in the verb window in weakly associated trials ( $z = 0.42$ ;  $p = .026$ ) and in strongly associated trials ( $z = 0.38$ ,  $p = .045$ ), but not in conflict trials ( $z = -0.14$ ,  $p = .47$ ). None of the correlations between residualized comprehension scores and children's prediction ability was significant ( $p > .05$ ).

We further analysed performance separately in strongly associated, weakly associated, and conflict trials for evidence of subtle modulation of prediction effects by the associative strength of the relationship between the verb and the label for the target image. Such modulation would provide especially strong evidence for the conclusion that prediction performance is influenced by the associative strength of the relationship between the verb and the label of the target image. In particular, this analysis examined the average ratings of associative strength provided by parents of infants taking part in the experiment and the proportion of looking to the target in the verb window. While there were no significant correlations between the (averaged) associative strength and prediction performance in strongly or weakly associated trials ( $p > .3$ ), we found that fixations to the strongly associated image in conflict trials (in the verb window) negatively correlated with the rating of the associative strength of the weakly associated image and the verb (as indicated by parents of infants taking part in the study). Thus, the more

<sup>3</sup>We suggest that the data from conflict trials and the strongly and weakly associated trials ought not to be analysed together due to differences in the relationship of "target" and "distractor" images across these trials. However, one could have an a priori hypothesis about the strongly associated image being more appropriate than the weakly associated image and, therefore, more suitable as a target. Therefore, we analysed the data in a  $3 \times 2$  ANOVA with condition (conflict, strongly associated, weakly associated) and window (baseline, verb) as within-subjects factors. This analysis yielded highly similar results to the main analyses reported for both children and adults. In particular, there was no significant interaction between window and condition,  $p = .8$ . Nevertheless, we maintain that this is a less preferable method of analysing the data due to the a priori assumption that the strongly associated image is the target.

<sup>4</sup>We could not obtain vocabulary data from two children who are, consequently not included in these analyses. We also excluded one child whose vocabulary scores were more than 2 standard deviations below the mean—that is, this child was reported to comprehend 39 words and produce 34 words. Further analysis of the data from this child suggested that the caregiver predominantly marked the child as comprehending a word only when the child also produced this word. Note that the pattern remained highly similar for both strongly and weakly associated trials even when the data from this child were included.

strongly the label for the weakly associated image was associated with the verb, the less infants looked towards the strongly associated image in conflict trials ( $z = -0.47$ ,  $p = .036$ ). There was no correlation between the ratings for the strongly associated image and performance in conflict trials ( $p = .8$ ).

Finally, we note that strongly and weakly associated images in conflict trials tended to also vary with regard to how frequent these labels were likely to be in children's input (according to analysis of caregiver child interaction in the German Szagun corpus<sup>5</sup>). Thus, for six of the 10 item sets used, the strongly associated image was also more frequent, for two trials there was no or a very small difference in the frequency of the two labels, while for two further trials the weakly associated trials were more frequent than the strongly associated trials. One possibility is that performance in the conflict trials could be attributed to the relative frequency of the strongly and weakly associated labels as opposed to the associative strength of the relationship between the verb and the label for these images.<sup>6</sup> We, therefore, examined the extent to which fixations to the strongly associated image in the verb window in conflict trials correlated with (a) frequency of the strongly associated label, (b) the frequency of the weakly associated label, and (c) the difference in frequency of the strongly and weakly associated label. Neither of these factors modulated performance in conflict trials. We also then divided trials into two conditions, one where the frequency of the strongly associated label was greater than the frequency of the weakly associated label ( $n = 12$ ), counterbalancing for side of presentation of image, and one where the frequency of the weakly associated label was either greater than (or very similar to) the frequency of the strongly associated label ( $n = 8$ ), and

examined whether there were any differences in performance across the baseline phase and the verb phase of the trial. Independent-samples  $t$ -tests found no significant difference between the former and latter conditions across the baseline phase, verb phase, or the change in fixations from the baseline to the verb phase ( $ps > .3$ ). Indeed, when separately analysing whether there was a change in fixations to the strongly associated image across the two conditions, we found that both conditions tended towards a significant difference in fixations to the strongly associated image from the baseline phase to the verb phase [frequent strongly associated label:  $t(11) = -1.91$ ,  $p = .083$ ; frequent weakly associated label:  $t(7) = -1.91$ ,  $p = .099$ ].

Finally, we ran a stepwise multiple regression to examine the extent to which performance in conflict trials was modulated by the associative strength of the relationship between the verb and the labels for the images and the frequency of the labels for the images, with the following factors entered into the regression model: (a) associative strength of the relationship between the verb and the strongly associated label; (b) associative strength of the relationship between the verb and the weakly associated label; (c) difference in ratings of the associative strength of the relationship between strongly and weakly associated labels; (d) frequency of the strongly associated label; (e) frequency of the weakly associated label; (f) difference in frequency of strongly and weakly associated labels, with proportion of fixations to the strongly associated label in the verb window as the dependent variable in the analysis. As suggested by the correlations reported above, only the associative strength of the relationship between the verb and the weakly associated label was retained as a predictor of performance in conflict trials, with a change in one unit of this factor leading to a change in  $(-).47$  units in fixations in the verb window,  $F(19) = 5.16$ ,  $p = .036$ ,  $R^2 = .23$ .

<sup>5</sup>This corpus includes data from 22 normal hearing children whose interactions were recorded at five points between 1;4 and 2;10 years of life. We included all mentions of the label across all data points at 1;4, 1;8, 2;1, 2;5, and 2;10 years of age.

<sup>6</sup>We note, however, that this is precisely the reason why we included a baseline phase in the experiment since were children to merely fixate the image of the more frequent label (in six out of 10 trials), then this preference should also be present in the baseline phase. Any increase from the baseline to the verb phase, therefore, excludes fixations driven by a putative preference for the more frequent image.

## Discussion

The findings of the current study support the following conclusions. First, children and adults are more likely to predict upcoming linguistic input based on the relative strength of association of these linguistic stimuli with one another in their past experience. Thus, given the relative strength of the associative relationship between *read* and *book* and *read* and *letter* in our experience, we are more likely to predict *book* as a thematically appropriate continuation of the verb *read*. Second, while we replicate Mani and Huettig's (2012) finding of an influence of children's production skills on their prediction of upcoming linguistic input, we note that this influence of production on children's prediction was only present in conditions where there was a uniquely associated image presented on screen—that is, in strongly associated and weakly associated trials but not in conflict trials. We interpret these findings to suggest that multiple factors contribute simultaneously to predictive language processing (cf., Huettig, 2015; Mani & Huettig, 2013). In what follows, we discuss these conclusions in more detail.

In the conflict condition presented to participants in the current study, children and adults heard a phrase such as “The boy reads...” and saw two images, one of which was a strongly associated with the verb (e.g., *book*), while the other was weakly associated with the verb (e.g., *letter*). Children and adults oriented towards the strongly associated image in preference to the weakly associated image soon after hearing the verb *reads*, suggesting that they were more likely to predict *book* as a thematically appropriate continuation of the verb *reads*, as opposed to *letter*. Thus, when presented with two thematically appropriate targets for a particular image, children and adults orient towards a more strongly associated target rather than a weakly associated target.

Furthermore, we found that fixations towards the strongly associated image were modulated by the associative strength of the relationship between the verb and the weakly associated label. Thus, the stronger the relationship between the weakly associated label and the verb, the more

children were also likely to fixate the weakly associated image. Indeed, fixations in the verb window in conflict trials were modulated by only this factor compared to other factors such as the strength of the association between the strongly associated label and the verb or the frequency of the labels for both images. This finding rules out alternative explanations for performance in conflict trials, based, for instance, on the frequency of the items presented. Especially taken together with the preference for strongly associated images in conflict trials, the modulation of performance by the associative strength ratings provide especially strong evidence for an important role for associations in our prediction of upcoming linguistic input as suggested by some models of predictive processing to date (e.g., Bar, 2007, 2009). However, as we argue below, the results raise questions about some of the assumptions raised in these models.

First, we note that only variability in the associative strength of the weakly associated label and the verb modulated fixations in conflict trials, while the associative strength of the strongly associated labels did not. One strong possibility is that there was increased variability in the ratings of associative strength in the weakly associated condition (range: 1.89–4.06) relative to the strongly associated condition (range: 3.6–4.97). It would be interesting to examine whether similar findings are reported when there is increased variability across both conditions potentially in experiments involving adults, since such controls are difficult to obtain in data with young children given their limited vocabularies. Indeed, this finding could have important implications for the mechanisms underlying the influence of associative strength on predictive language processing. We currently interpret this finding tentatively as indicative of the competition between fixations to the strongly and weakly associated images. Taken together, it appears that children's eye movements are influenced by the strength of the association between the verb and the image; the stronger this association is, the more likely the child will look at the respective image. Thus, in cases where the a priori assumed weakly associated image was strongly associated with the verb, children fixated this

image longer than other weakly associated images that were less strongly associated with the verb.

Second, we note that when strong and weak associations were presented separately—that is, alongside a thematically inappropriate distractor—there were no differences in the prediction effect for either continuation. Indeed, it appears to be the case that only when the visual context provides both strongly and weakly associated images in parallel, thereby putting them in direct competition with one another, that we are able to tap into the role of associative strength in predictive language processing. This raises questions as to the extent to which past associations are involved in our prediction of upcoming linguistic input in daily use, where such visual context may not be immediately available. Does prediction continue unaided by associative influences in the absence of such direct competitive contexts? It would be of interest to explore whether other paradigms are able to tap into differences in the time-course of prediction of strongly and weakly associated continuations (e.g., event-related potential, ERP, paradigms, cf. Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007).

Finally, the associative strength of the relationship between the verbs and the labels for the presented images did not uniquely modulate performance in the current study. As in previous work (Mani & Huettig, 2012), we found that children with larger production vocabularies were better predictors than children with smaller production vocabularies. However, we note that this result was restricted to performance in the strong and weak association condition. No such correlation was found in conflict trials. This intriguing dissociation between conflict trials and single target trials lends itself to a discussion of the factors influencing predictive language processing.

We interpret the pattern of results found in the current study as indicative of multiple factors that simultaneously contribute to predictive language processing. In trials where the target (strongly or weakly associated label) is the only image that is consistent with the thematic constraints imposed by the verb, the child's productive language skills modulate their prediction skills. Here, since only one image is thematically appropriate, the strength

of the relationship between the label for this image and the verb does not influence responding—prediction here appears to be dominated by the child's language competence, in particular the child's productive competence. These results support models that afford an influential role for production in listeners' anticipation of upcoming language input, such that listeners' implicit production of what they think the speaker will say next drives predictive language processing (Chang et al., 2006; Dell & Chang, 2013; Pickering & Garrod, 2013). In contrast, in trials where both objects are consistent with the thematic constraints imposed by the verb—that is, conflict trials—performance appears to be more strongly modulated by the associative strength of the relationship between the verb and the labels for the images. The separable influences of associations and production-based representations suggest, therefore, that associations are not just used by the production system (or forward model) to predict upcoming language input. In contrary, it appears that a more dynamic pluralistic approach is called for.

Note that this is not to suggest that past associations do not influence such production-based predictive language processing. Indeed, current models of production-based predictions both include the possibility of past associations influencing the production component to predictive language processing (Chang et al., 2006; Dell & Chang, 2013; Huettig, 2015; Pickering & Garrod, 2013). However, we suggest that the dissociation in the influences of associations and production skills noted in the current study suggests that there is more to production-based prediction than merely past associations and, furthermore, more to prediction than merely the generation of experience-independent implicit productions of what the speaker is likely to say next.

This points, again, to multiple factors influencing predictive language processing, with factors such as associations (the current study), production-based representations (Gampe & Daum, 2014; Mani & Huettig, 2012; the current study), general vocabulary-based influences (Borovsky et al., 2012; Rommers, Meyer, & Huettig, 2015), the availability of prosodic information (Keitel et al., 2013), and

participants' reading skills (Huettig & Brouwer, 2015; Mani & Huettig, 2014; Mishra, Singh, Pandey, & Huettig, 2012) influencing predictive language processing. We suggest, therefore, that what is required is a dynamic view of predictive language processing where different factors may have more or less of an influence in different contexts, allowing for some to take precedence over others based (a) on the context in which the listener finds herself anticipating upcoming linguistic input, and, indeed, (b) individual differences of the listeners themselves.

## REFERENCES

- Bar, M. (2007). The pro-active brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, 11, 280–289.
- Bar, M. (2009). The pro-active brain: Memory for predictions. *Philosophical Transactions of the Royal Society B*, 364, 1235–1243.
- Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory 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(2), 234–272.
- Dell, G. S., & Chang, F. (2013). The P-Chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society B*, 369, 20120394. doi:10.1098/rstb.2012.0394
- Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505.
- Federmeier, K. D., Wlotko, E., De Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, 1146, 75–84.
- Gampe, A., & Daum, M. M. (2014). Productive verbs facilitate action prediction in toddlers. *Infancy*, 19, 301–325.
- Houston-Price, C., Mather, E., & Sakkalou, E. (2007). Discrepancy between parental reports of infants' receptive vocabulary and infants' behaviour in a preferential looking task. *Journal of Child Language*, 34, 701–724.
- Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*, 1626, 118–135.
- Huettig, F., & Brouwer, S. (2015). Delayed anticipatory spoken language processing in adults with dyslexia - Evidence from eye-tracking. *Dyslexia*, 21(2), 97–122.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). Prediction and thematic information in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49, 133–156.
- Keitel, A., Prinz, W., Friederici, A. D., von Hofsten, C., & Daum, M. M. (2013). Perception of conversations - The importance of semantics and intonation in children's development. *Journal of Experimental Child Psychology*, 116, 264–277.
- Kukona, A., Fang, S., Aicher, K. A., Chen, H., & Magnuson, J. S. (2011). The time course of anticipatory constraint integration. *Cognition*, 119, 23–42.
- Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, 1146, 23–49.
- 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.
- 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.
- Melinger, A., & Weber, A. (2006). Database of noun associations for German. Retrieved from <http://www.coli.uni-saarland.de/projects/nag/>
- 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, 1–10.
- Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, 86, 314–329.
- Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11, 105–110.

- Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, 36, 329–347.
- Rommers, J., Meyer, A. S., & Huettig, F. (2015). Verbal and nonverbal predictors of language-mediated anticipatory eye movements. *Attention, Perception, & Psychophysics*, 77(3), 720–730.
- Szagan, G., Stumper, B., & Schramm, A. S. (2009). *Fragebogen zur frühkindlichen Sprachentwicklung (FRAKIS) und FRAKIS-K (Kurzform)*. Frankfurt: Pearson Assessment.