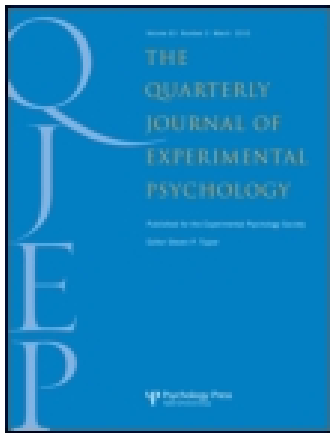


This article was downloaded by: [Universitätsbibliothek Bern]

On: 12 September 2014, At: 09:11

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office:
Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pqje20>

Deconstructing Frith and Snowling's homograph-reading task: Implications for autism spectrum disorders

Jon Brock^{a b} & Samantha Bzishvili^c

^a ARC Centre for Cognition and its Disorders , Australia

^b Department of Cognitive Science , Macquarie University , Sydney , Australia

^c Department of Psychology , Macquarie University , Sydney , Australia

Accepted author version posted online: 17 Jan 2013. Published online: 21 Feb 2013.

To cite this article: Jon Brock & Samantha Bzishvili (2013) Deconstructing Frith and Snowling's homograph-reading task: Implications for autism spectrum disorders, The Quarterly Journal of Experimental Psychology, 66:9, 1764-1773, DOI: [10.1080/17470218.2013.766221](https://doi.org/10.1080/17470218.2013.766221)

To link to this article: <http://dx.doi.org/10.1080/17470218.2013.766221>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or

distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Deconstructing Frith and Snowling's homograph-reading task: Implications for autism spectrum disorders

Jon Brock^{1,2}, and Samantha Bzishvili³

¹ARC Centre for Cognition and its Disorders, Australia

²Department of Cognitive Science, Macquarie University, Sydney, Australia

³Department of Psychology, Macquarie University, Sydney, Australia

The poor performance of autistic individuals on a test of homograph reading is widely interpreted as evidence for a reduction in sensitivity to context termed “weak central coherence”. To better understand the cognitive processes involved in completing the homograph-reading task, we monitored the eye movements of nonautistic adults as they completed the task. Using single trial analysis, we determined that the time between fixating and producing the homograph (eye-to-voice span) increased significantly across the experiment and predicted accuracy of homograph pronunciation, suggesting that participants adapted their reading strategy to minimize pronunciation errors. Additionally, we found evidence for interference from previous trials involving the same homograph. This progressively reduced the initial advantage for dominant homograph pronunciations as the experiment progressed. Our results identify several additional factors that contribute to performance on the homograph reading task and may help to reconcile the findings of poor performance on the test with contradictory findings from other studies using different measures of context sensitivity in autism. The results also undermine some of the broader theoretical inferences that have been drawn from studies of autism using the homograph task. Finally, we suggest that this approach to task deconstruction might have wider applications in experimental psychology.

Keywords: Homographs; Eye movements; Speech production; Autism; Experimental design.

At the heart of the cognitive approach to developmental disorders lies the critical distinction between observable behaviours and the underlying cognitive processes that cause those behaviours (Frith, 2012; Morton, 2004). Within this causal framework, performance on psychological tests is distinguished from the cognitive mechanisms that the test purports to measure. Inferring atypical or impaired cognitive mechanisms on the basis of test performance is thus dependent on a detailed

and accurate understanding of the processes that contribute to measured behaviour.

A case in point is the homograph-reading task, introduced by Frith and Snowling (1983) as a measure of reading comprehension (see Brock & Caruana, in press). The poor performance of autistic individuals on the test is generally construed in terms of a failure to process information in context or “weak central coherence” (Frith, 1989). Indeed, in a rather circular fashion, the

Correspondence should be addressed to Jon Brock, Department of Cognitive Science, Macquarie University, Sydney 2109, Australia. E-mail: jon.brock@mq.edu.au

This research was supported by Australian Research Council Discovery Project DP098466 and a Macquarie University Research Development Grant. We thank Joann Tang for assistance with data processing, Reinhold Kliegl for advice regarding statistical analyses, and Sachiko Kinoshita for comments on an earlier draft.

homograph-reading test is now also considered a good *measure* of central coherence, precisely because autistic individuals perform poorly (Happé & Frith, 2006). However, as we demonstrate in the current study, there are many other factors that contribute to variation in performance, and, as such, conclusions based on the results of studies using the task may have been premature.

The homograph reading task requires participants to read aloud sentences containing homographs—written words that have multiple meanings associated with the same orthographic form. Crucially, the homographs chosen also have different pronunciations associated with their two meanings. For example, the word “tear” should be pronounced differently in the sentences “In her *eye* was a tear” versus “In her *dress* was a tear”. The pronunciation chosen by the participant indexes their ability to determine the contextually appropriate meanings and thereby provides an incidental measure of their understanding. Strikingly, Frith and Snowling (1983) found that children with autism consistently gave the most common pronunciation of the homographs, regardless of contextual cues.

In a follow-up study, Snowling and Frith (1986) introduced a further manipulation, whereby the homographs appeared before rather than after the critical disambiguating context. They also included a training session in which participants were alerted to the ambiguous nature of some of the words and were given instruction in their alternative meanings. In contrast to their earlier study, Snowling and Frith found no evidence for an autism-specific impairment. However, as Happé (1997) pointed out, the training session meant the task was no longer a spontaneous and open-ended measure of reading comprehension.

Happé’s (1997) study used the same conditions and stimuli as those of Snowling and Frith (1986) but eschewed the training session. When the rare pronunciation was correct (and so participants relied on context to arrive at the correct response), typically developing control children performed better when the context came before rather than after the homograph. In contrast, children with autism showed no significant effect of context

position, despite being older and having higher verbal mental ages than the control children. This Happé interpreted as confirmation of Frith and Snowling’s original conclusion that words are processed out of context in autism.

Three further studies have adopted the same procedures as Happé (1997) but tested adolescents and adults with autism (Burnette et al., 2005; Jolliffe & Baron-Cohen, 1999; Lopez & Leekam, 2003). Like Happé, these studies each found impaired performance relative to typically developing control participants when the rare pronunciation was correct. However, unlike Happé, these studies reported impaired performance across both homograph positions.

Findings from these studies have had a considerable impact upon cognitive accounts of autism. Frith (1989) took autistic participants’ apparent immunity to sentence context as key evidence for her “weak central coherence” account, according to which autistic individuals experience an “inability to draw together information so as to derive coherent and meaningful ideas” (p. 187). The central coherence account has since become one of the most influential cognitive theories of autism, providing a potential explanatory framework for many of the language and communication difficulties affecting autistic individuals (Noens & van Berckelaer-Onnes, 2005), as well as their relative strengths on nonverbal tasks that required them to ignore the meaning or gestalt of the stimuli (Happé & Booth, 2008). Frith (1989) initially suggested that weak central coherence could also explain theory of mind deficits in autism. However, the apparent dissociation between performance on the homograph reading task and understanding of false beliefs (Happé, 1997; Jolliffe & Baron-Cohen, 1999) has led to the current view that central coherence and theory of mind represent independent cognitive deficits within autism (Happé & Frith, 2006).

Yet despite the continued influence of the homograph-reading task on autism research and theory, the inference that autistic individuals process words out of their semantic context has been challenged by a number of studies using other paradigms. Of particular note, Norbury

(2005) devised a test in which participants listened to a sentence ending in a homophone (an ambiguous *spoken* word) and then decided whether a picture relating to one of the homophone meanings matched the sentence. Individual differences in performance were linked to participants' degree of language impairment, with no evidence for an autism-specific deficit once language ability had been controlled for.

The same pattern of results was reported by Brock, Norbury, Einav, and Nation (2008) using a language-mediated eye-movements paradigm. While listening to sentences, participants viewed a computer display containing four objects. In the critical condition, one of the objects in the display was a cohort competitor of a word in the sentence. Participants' tendency to look towards this competitor object was moderated by the extent to which the object was a contextually plausible completion of the sentence. Consistent with Norbury's (2005) findings, the size of this context effect was associated with children's degree of language impairment, but was independent of their autism diagnosis.

Most recently, Henderson, Clarke, and Snowling (2011) employed a cross-modal semantic priming task in which participants listened to either a biased or a neutral sentence ending in a homophone and then named a picture. When the picture was related to the subordinate meaning, participants were quicker to name it if the sentence context was also biased towards the subordinate meaning (e.g., Judy bought/planted the bulb-FLOWER). Importantly, this effect was similar in participants with and without autism, indicating typical sensitivity to sentence context in autism.

Together these recent findings represent a serious challenge for the weak central coherence account, which assumes that context-processing difficulties are a ubiquitous feature of autistic cognition and should, therefore, generalize across paradigms and stimulus modalities. They suggest that the poor performance of autistic individuals on the homograph reading task may not after all be a reflection of reduced sensitivity to context as is widely assumed. However, in order to advance this argument, it is necessary to identify other factors that might contribute to variation in

performance on the task and hence provide an alternative explanation for poor performance of autistic individuals.

To this end, in the current study, we gave the homograph reading task to a large group of nonautistic adults. As well as recording spoken responses, we monitored participants' eye movements throughout the task, carefully aligning the record of fixations with the speech recording. Using logistic regression analysis, we then modelled the accuracy data to determine the factors—including eye-tracking measures, trial number, and the content of earlier trials—that would predict accurate pronunciation of the homograph on a given trial.

Method

Participants

Participants were 40 undergraduate psychology students (10 males) aged 18 to 32 years, who completed the study for course credits.

Stimuli

Stimuli were the 20 sentences provided in the appendix of Snowling and Frith (1986) and used in all subsequent studies of homograph reading in autism. In Snowling and Frith's study, some of the sentences were preceded by one or two sentences of story context, but the correct meaning of the homograph was discernable from the final sentence alone (e.g., "The boys played cowboys and Indians. Paul was a cowboy and pretended to have a gun. Tom was an Indian and pretended to have arrows and a *bow*."). For these items, we presented only the final sentence, ensuring that the whole text could be displayed on a single row.

Each sentence contained one of five heterophonic homographs (tear, bow, read, lead, row). There were two orthogonal manipulations: context—whether the context appeared before or after the homograph; and dominance—whether the correct meaning was the dominant or subordinate (less common) meaning.

Procedure

Stimuli were presented on a 35 × 26-cm CRT monitor at 1,024 × 768-pixel resolution.

Movements of the right eye were recorded at 1000 Hz via an EyeLink 1000 tower-mounted eye-tracker (SR Research, Mississauga, Ontario, Canada) with an eye to screen distance of 45 cm. Saccades and fixations were parsed online using the "NORMAL" settings—a saccade was identified if instantaneous velocity exceeded 30° s^{-1} or acceleration exceeded $8,000^\circ \text{ s}^{-2}$. Prior to testing, participants completed a standard three-point calibration routine, with each of the three points placed in a horizontal line at the screen height at which the sentences were to be presented.

Stimulus presentation and voice recording were controlled using Experiment Builder software (SR Research, Mississauga, Ontario, Canada). Each trial began with a circular fixation point placed towards the left edge of the screen. Fixation on the circle triggered the presentation of the sentence, with the first word appearing in the location of the fixation point. Sentences were presented in a single row of text in 18-point black Courier font with a white background. Letters were approximately 6 mm in height. Participants read each sentence aloud, and their speech was recorded at 44.1 kHz. A keypress from the experimenter initiated the next trial. As in the studies of autism, sentences were presented in a different random order for each participant, with no attempt made to avoid repetitions of the same homograph in consecutive trials.

Data analysis

Voice recordings were analysed offline in Adobe Audition to determine the onset time of the first attempt at producing the homograph and whether or not that first attempt was correct. Given concerns about ceiling effects, we ignored any subsequent self-corrections. This is consistent with the coding in Frith and Snowling's (1983) original study of autism and with subsequent studies showing group differences regardless of whether or not self-corrections were allowed (Happé, 1997; López & Leekam, 2003).

Eye movements were analysed using DataViewer software (Version 10.1.165) (SR Research, Mississauga, Ontario, Canada). Trials were only included in the eye-movement analysis

if (a) the first fixation on the homograph was progressive (i.e., words to the right of the homograph had not been fixated before the homograph); (b) the first fixation on the homograph was at least 50 ms (cf. Rayner, 2009); and (c) the eye-to-voice span was at least 100 ms. Eighty of 800 trials were lost to this screening process.

Statistical analyses were conducted in R Version 2.15.0 using the lme4 library (Bates, 2005). All fixed factors were centred, ensuring that the interactions were evaluated at the middle of the data. Accuracy data (coded as 1 or 0) were subjected to generalized mixed random effects analyses, treating subject and item (sentence) as crossed random factors (see Baayen, Davidson, & Bates, 2008). Precise estimation of p -values is not currently possible for these analyses. Nonetheless, given sufficient data, the z -statistics approximate a normal distribution, and z -values outside the range ± 1.96 can be considered to be statistically significant. Gaze duration and speech-onset time data were subjected to ordinary mixed random effects analyses, with p -values estimated via Markov chain Monte Carlo simulation. Quantile-quantile plots were used to check for the normality of residual distributions, with log-transformations applied as appropriate. Model-based trimming of outliers was conducted, with residuals of greater than 2.5 standard deviations being removed before rerunning analyses on the trimmed data. Hierarchical model comparison was employed, whereby fixed factors and interactions were added or removed to the regression equation and retained if the more complex model was a significantly better fit.

Results

Accuracy

As an initial analysis, we considered accuracy of pronunciation as a function of context and dominance (Model 1). Consistent with previous studies, accuracy was greater when the dominant pronunciation was correct, $z = 2.49$, and when the context preceded the homograph, $z = 3.40$. However, the interaction was nonsignificant, $z = 0.90$.

1. Accuracy \sim Context * Dominance + (1|Subject) + (1|Sentence)

As the next step, we added a series of different predictors to this base model, allowing them to interact with both context and dominance (e.g., Model 2).

2. Accuracy \sim Context * Dominance * Trial + (1|Subject) + (1|Sentence)

Adding trial number significantly improved model fit, $\chi^2(4) = 14.6, p = .006$. Accuracy increased as a function of trial number, although this was not significant, $z = 1.52$. However, there was a significant three-way interaction, $z = 2.20$. As can be seen in Figure 1, this interaction arose because accuracy *decreased* across trials for the dominant meaning when the context came before the homograph, but increased for the other three conditions.

Further analyses considered the predictive value of various eye-tracking measures including duration of the first fixation on the homograph, duration of the first run on the homograph (total duration of consecutive fixations on the homograph before saccading to a different word), and gaze duration (total time fixating on the homograph before producing it). None of these significantly improved model

fit. However, adding eye-voice span (time from first fixating on the homograph to producing it) did improve model fit relative to the base model, $\chi^2(4) = 21.7, p < .001$ (see Model 3). Longer eye-voice spans were associated with greater accuracy, $z = 4.18$, with no significant interactions.

3. Accuracy \sim Context * Dominance * log_Eye_Voice_Span + (1|Subject) + (1|Sentence)

Given the significant three-way interaction between context, dominance, and trial, we also considered the effects of interference from previous trials. Specifically, we hypothesized that exposure to the subordinate meaning and pronunciation of the homograph on earlier trials might act to annul the initial dominance effect on later trials. We coded each trial to determine whether the subordinate meaning of the homograph had been encountered on a previous trial and added this to the model (Model 4). This significantly improved fit relative to the base model, $\chi^2(4) = 23.9, p < .001$. There was a significant two-way interaction with dominance, $z = -2.01$, and a significant three-way interaction, $z = 2.19$. In contrast, coding trials according to whether or not the *dominant* meaning of the homograph had been encountered on a previous trial did not improve model fit, $\chi^2(4) = 4.15, p = .386$.

4. Accuracy \sim Context * Dominance * Trial * Subordinate_Encountered + (1|Subject) + (1|Sentence)

Having established that accuracy was predicted by context, dominance, trial, eye-voice span, and previous trials featuring the subordinate meaning, we combined the different predictors via a process of model comparison, adding predictors if they significantly enhanced the fit of the model and removing them at a later stage if they became redundant (i.e., their removal did not significantly reduce model fit). Importantly, we arrived at the same final model (Model 5), regardless of the order in which different fixed factors were introduced. Note in particular that trial was now found to be redundant, with changes in accuracy over trial being explained in terms of previous encounters with the subordinate meaning of the homograph and eye-voice span. Table 1 provides the full

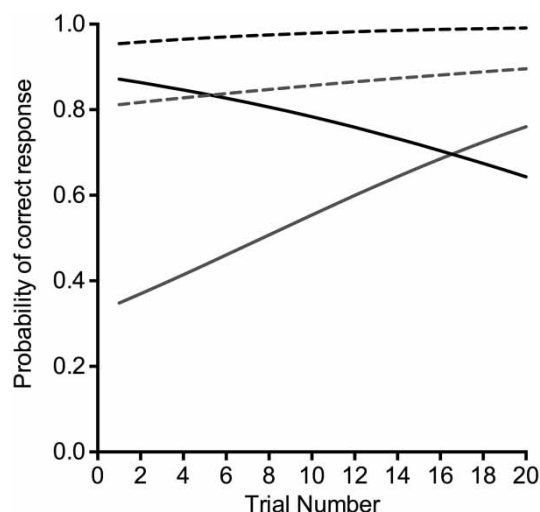


Figure 1. Accuracy as a function of trial and condition. Black lines indicate dominant pronunciation; grey lines represent subordinate pronunciation. Solid lines represent context after; broken lines represent context before.

Table 1. *Parameter estimates for final model predicting accuracy of pronunciation*

<i>Estimate</i>	<i>Estimate</i>	<i>Standard error</i>	<i>z-value</i>
(Intercept)	1.9324	0.3288	5.878*
Context	2.5054	0.6427	3.898*
Dominance	1.5101	0.6307	2.394*
Subordinate encountered	0.4647	0.2832	1.641
Eye-to-voice span	2.2677	0.5123	4.426*
Context by dominance	1.0428	1.2613	0.827
Context by subordinate encountered	0.7526	0.5678	1.325
Dominance by subordinate encountered	-1.1567	0.5679	-2.037*
Three-way interaction	2.6442	1.1427	2.314*

* indicates $|z| > 1.96$.

breakdown of the final model. As shown in Figure 2, performance improved with increasing eye-to-voice span across all conditions. Notably, in the context-after condition, the advantage for the dominant meaning was entirely eliminated if the subordinate meaning had been encountered on a previous trial.

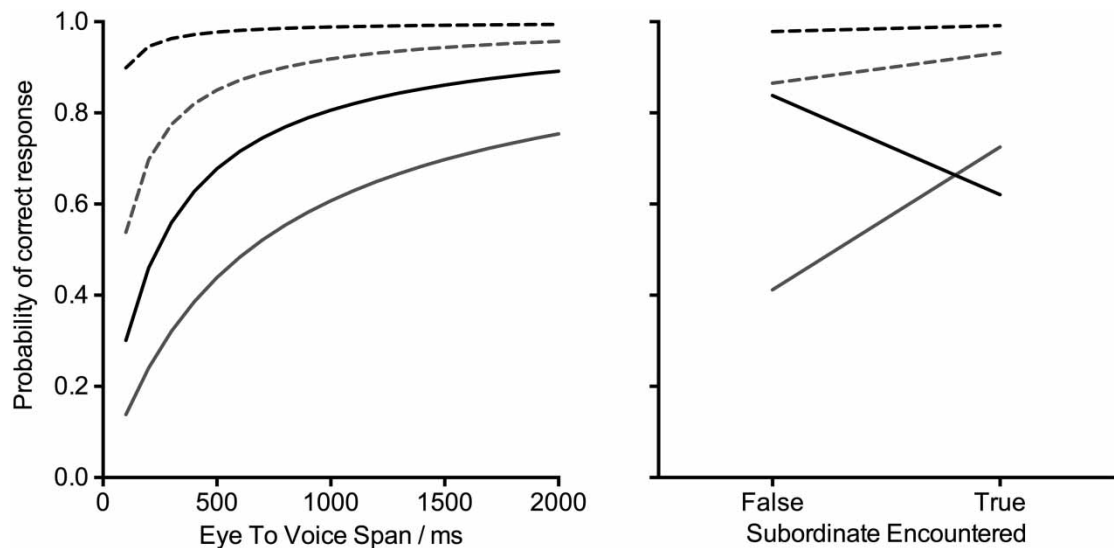


Figure 2. *Final model of pronunciation accuracy as a function of (a) eye-voice span and (b) previous encounter with the subordinate meaning. Black lines indicate dominant pronunciation; grey lines represent subordinate pronunciation. Solid lines represent context after; broken lines represent context before.*

5. Accuracy \sim Context * Dominance * Subordinate
_Encountered + log_Eye_To_Voice_Span + (1|
Subject) + (1|Sentence)

At this point, we inspected the random effects to identify potential outliers. There were no outliers for the random effect of subject, with all participants within 2 standard deviations of the group mean. With respect to the item random effect, there was one clear outlier—the sentence “I read a story now and then I do some maths”, which attracted a disproportionate number of errors. Model 5 was reapplied, excluding this item from the data, but the pattern of significance was unaltered.

We also conducted a series of reanalyses that maintained the fixed effect structure in Model 5 but changed the random effect structure. These reanalyses included (a) removing one or other of the random effects; (b) treating homograph rather than sentence as by-items random factor; (c) allowing context and dominance to interact with the random effects. In each case, the changes had only a minimal impact on the estimates of the regression parameters and did not affect the patterns of statistical significance. In sum, the

conclusions based on Model 5 are not contingent upon the random effects structure adopted in the model.

Eye-voice span

Given the importance of eye-voice span in predicting accuracy, we also investigated the effect of context, dominance, and trial number on the (log-transformed) duration of the eye-voice span (Model 6). Eye-voice span was significantly longer when the context came after the homograph, $t = -4.29$, $p_{\text{MCMC}} < .001$, but there was no effect of dominance, $t = -0.42$. Of note, eye-voice span duration increased significantly with trial number, $t = 4.20$, $p_{\text{MCMC}} < .001$. No interactions approached significance, $p_{\text{MCMC}} > .50$, indicating a uniform effect of trial on eye-voice span across conditions.

6. log Eye-Voice Span ~ Context * Dominance * Trial + (1|Subject) + (1|Sentence)

Further investigations revealed that, despite the progressive increase in eye-voice span, there was no effect of trial on the time at which participants began pronouncing the homograph relative to the onset of presentation of the sentence, $t = 0.36$,

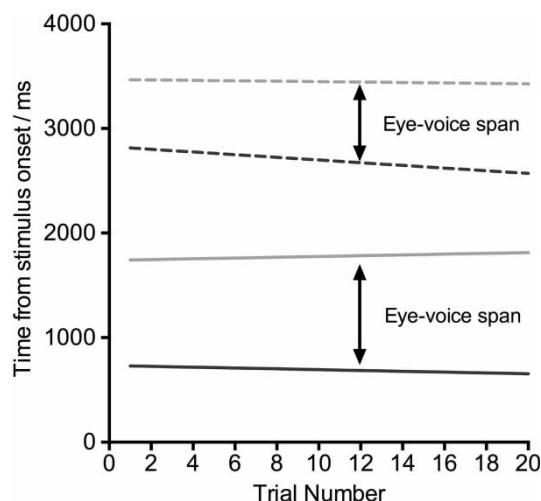


Figure 3. Time of first fixation on the homograph (dark grey) and onset of production (light grey) as a function of trial and context position. Solid lines represent context after; broken lines represent context before.

$p_{\text{MCMC}} = .707$. Instead, eye-voice spans increased because participants were quicker to fixate on the homograph in later trials, $t = -3.29$, $p_{\text{MCMC}} = .002$ (Figure 3).

Finally, we considered the distribution of eye movements during the eye-voice span. The proportion of the eye-voice span spent fixating on the homograph was greater when the context came before the homograph, $t = 6.47$, $p_{\text{MCMC}} < .001$. This may reflect clause “wrap-up” effects (cf. Rayner, Kambe, & Duffy, 2000) or simply the fact that the homograph was the last word on the line of text in 8 of the 10 sentences, leaving nowhere else for the eyes to go. There was no effect of dominance, $t = 0.27$, but a borderline reduction with trial number, $t = 1.95$, $p_{\text{MCMC}} = .051$, indicating that with increased eye-voice span in later trials, participants predominantly fixated on words other than the homograph. Figure 4 shows the distribution of these fixations for each of the sentences in the subordinate condition (similar patterns are apparent in the dominant condition). In the context-before condition, the homograph was the last word, so fixations off the homograph were necessarily regressions. Notably, in the context-after condition, the eye-voice span afforded participants the opportunity to fixate four or five words beyond the homograph and potentially disambiguate it before they began producing it.

Discussion

The poor performance of autistic individuals on the homograph reading task is one of the most influential findings in research on autistic cognition and remains a cornerstone of the weak central coherence account of the disorder. However, as the current study demonstrates, the task is deceptively complex. Even in healthy, nonautistic adults, performance reflects the interaction of a number of factors.

Previous studies using the homograph-reading task have considered only the total number of correct responses in each condition for each participant. Here we instead modelled the likelihood of an accurate response at the level of single trials. This

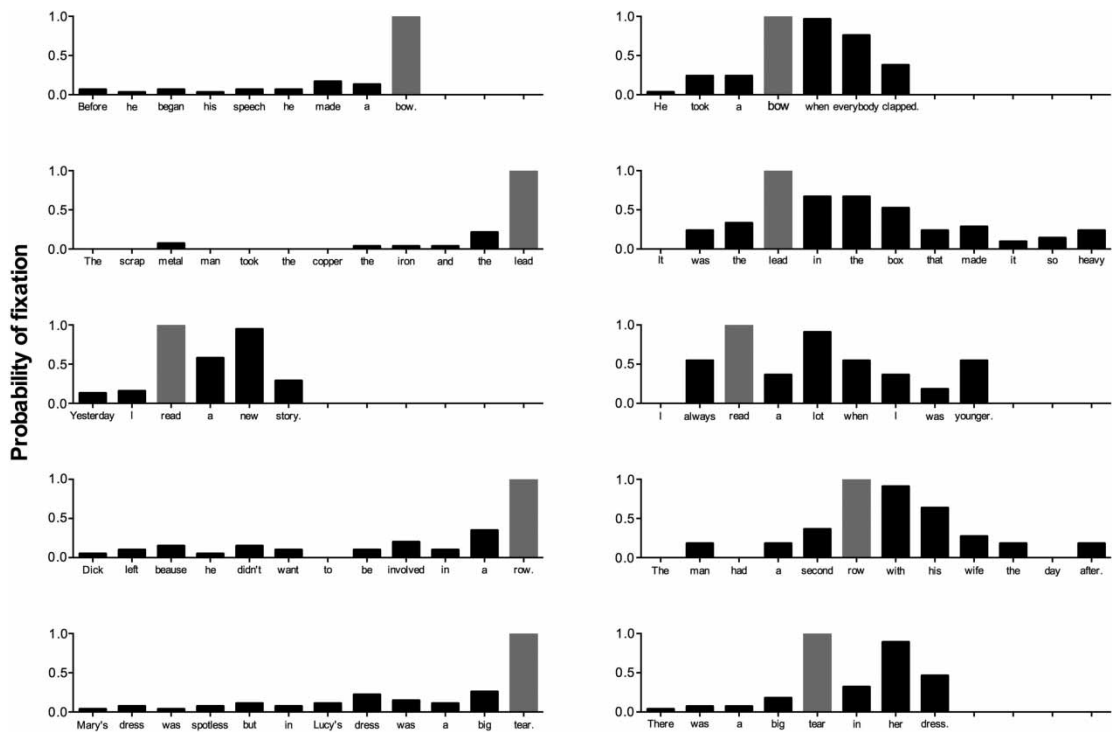


Figure 4. Probability of fixation on each word in the sentence during the eye-voice span. Data are plotted for each sentence in the subordinate condition. Left column shows the context-before condition. Right column shows the context-after condition.

revealed at least two distinct contributions to performance that have not previously been considered.

First, as trial number increased, participants' pattern of eye movements changed such that they arrived earlier at the homograph. This engendered an increased eye-to-voice span, which in turn was a significant predictor of accuracy. One possible explanation for these findings is that, as the experiment proceeded, and the ambiguous nature of some of the words became apparent, participants changed their reading strategy, giving themselves more time between fixating and pronouncing the homograph in order to minimize the number of overt errors.

Second, accuracy was a function of a three-way interaction between context, ambiguity, and trial number. Over the course of the experiment, accuracy decreased for the dominant condition when the context appeared after the homograph in the sentence, but increased in the other three

conditions. This interaction appeared to reflect interference from previous trials. Specifically, if participants had previously encountered the subordinate pronunciation of the homograph, they were more likely to produce it again, at least in the context-after condition. This entailed an improvement in accuracy for the subordinate meaning and a decrement for the dominant meaning. In essence, the dominance effect was abolished by the end of the experiment, by which time a previous encounter with the subordinate meaning was guaranteed.

Although we did not test individuals with autism in this study, the results have clear implications for autism research. In particular, they demonstrate that performance on the homograph reading task cannot simply be assumed to measure context processing or central coherence. Homograph reading accuracy is also influenced by interference from previous trials, the length of the participant's eye-to-voice span and, it would appear, their ability to

detect errors and strategically alter eye-to-voice span to minimize future errors. Our data do not necessarily refute the central coherence account of impaired homograph reading. Nonetheless, variation in these confounding effects must be accounted for prior to attributing group differences to “weak central coherence”. Indeed, our findings may go some way to explaining the discrepant results of studies using homograph reading task studies using alternative tests of sensitivity to sentence context that are not affected by these confounds.

The eye-tracking analysis also highlights the considerable lag between the participant fixating on the homograph and actually pronouncing it. Frith and Snowling's (1983) task is widely assumed to measure the immediate and spontaneous use of context, but it is clear that the eye-to-voice span affords the opportunity to reflect and perhaps reinterpret the homograph before producing it out loud. By contrast, in Brock et al.'s (2008) study, context effects were measured in terms of the influence of a preceding verb on eye movements that occurred during the acoustic lifetime of the target word. This, we argue, is a much purer measure of immediate context effects.

Our observations also undermine some of the far-reaching theoretical inferences that have been drawn from previous studies using the homographs task. For example, Happé (1997; see also Happé & Frith, 2006) concluded that theory of mind impairments in autism are entirely independent of context-processing difficulties (contra Frith, 1989) on the grounds that autistic children in her sample who passed false-belief tasks nevertheless failed to demonstrate a significant effect of context *position* in the homograph reading task.

Happé's analyses rested on the assumption that homographs appearing before the contextual information were effectively processed out of context so acted as a baseline condition. Our study undermines that assumption, showing that skilled readers have time to read well beyond the homograph and take in the disambiguating context before producing the homograph. Indeed, the absence of a context position effect in Happé's subgroup with “intact” theory of mind actually reflected strong performance regardless of context

position. A detailed discussion of the relationship between theory of mind and central coherence is beyond the scope of this paper. Nonetheless, this example demonstrates the more general point that the theoretical implications of a study can be easily influenced by what turn out to be incorrect assumptions about the nature of the task.

Given the above considerations, an obvious next step would be to conduct a similar analysis of predictors of homograph reading accuracy in autistic individuals. However, it is clear from previous studies of homograph reading in autism that significant differences in group averages mask considerable within-group variation in performance (Brock & Caruana, in press). Even amongst those who perform poorly, there may be different underlying reasons for their difficulty. With only five trials per condition and significant order effects, the homograph task is clearly ill-suited to investigation of reliable individual differences. Ultimately, the aim should be to rule out the confounding factors that we have identified, and, while the current study suggests that this can be done statistically, a better approach is to improve the task itself to avoid these confounds. For example, filler sentences should be added so that the nature of the task does not become obvious, and sentences involving the same homograph (or other critical word) should be presented in different blocks, preferably in separate test sessions.

In conclusion, the current study has succeeded in teasing apart a number of distinct factors that contribute to performance on Frith and Snowling's (1983) influential and widely used test of homograph reading. In doing so, we have demonstrated that homograph reading is not the pure measure of context processing that many autism researchers have implicitly assumed it to be. This is not to devalue the unique contribution of Frith and Snowling's ground-breaking study, but rather to point towards alternative explanations for their results. Indeed, these insights go some way towards explaining the discrepancies between studies of autism using the homograph task and those using alternative measures of context processing. More generally, we have

highlighted again the pitfalls inherent in conflating performance on a particular task with the underlying cognitive construct that it is assumed to measure (Morton, 2004), and we have demonstrated the potential of both eye-tracking and trial-level analyses for deconstructing performance on such tasks.

Original manuscript received 1 August 2012

Accepted revision received 6 January 2013

First published online 20 February 2013

REFERENCES

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Bates, D. M. (2005). Fitting linear mixed models in R. *R News*, 5, 27–30.
- Brock, J., & Caruana, N. (in press). Reading for sound and reading for meaning in autism: Frith and Snowling (1983) revisited. In J. Arciuli & J. Brock (Eds.), *Communication in autism: Trends in Language Acquisition Research (TiLAR) Series*. Amsterdam, The Netherlands: John Benjamin.
- Brock, J., Norbury, C. F., Einav, S., & Nation, K. (2008). Do individuals with autism process words in context? Evidence from language-mediated eye-movements. *Cognition*, 108, 896–904.
- Burnette, C. P., Mundy, P. C., Meyer, J. A., Sutton, S. K., Vaughan, A. E., & Charak, D. (2005). Weak central coherence and its relations to theory of mind and anxiety in autism. *Journal of Autism and Developmental Disorders*, 35, 63–73.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.
- Frith, U. (2012). Why we need cognitive explanations of autism. *Quarterly Journal of Experimental Psychology*, 65, 2073–2092.
- Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Journal of Developmental Psychology*, 1, 329–342.
- Happé, F. G. E. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, 15, 1–12.
- Happé, F. G. E., & Booth, R. D. L. (2008). The power of the positive: Revisiting weak central coherence in autism spectrum disorders. *Quarterly Journal of Experimental Psychology*, 61, 50–63.
- Happé, F. G. E., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36, 5–25.
- Henderson, L. M., Clarke, P. J., & Snowling, M. J. (2011). Accessing and selecting word meaning in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, 52, 964–973.
- Jolliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory; linguistic processing in high-functioning adults with autism or Asperger's syndrome: Is local coherence impaired? *Cognition*, 71, 149–185.
- López, B., & Leekam, S. R. (2003). Do children with autism fail to process information in context? *Journal of Child Psychology and Psychiatry*, 44, 285–300.
- Morton, J. (2004). *Understanding developmental disorders: A causal modeling approach*. Oxford: Blackwell.
- Noens, I. L., & van Berckelaer-Onnes, I. A. (2005). Captured by details: Sense-making, language and communication in autism. *Journal of Communication Disorders*, 38, 123–141.
- Norbury, C. F. (2005). Barking up the wrong tree? Lexical ambiguity resolution in children with language impairments and autistic spectrum disorders. *Journal of Experimental Child Psychology*, 90, 142–171.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62, 1457–1506.
- Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology*, 53, 1061–1080.
- Snowling, M., & Frith, U. (1986). Comprehension in “hyperlexic” readers. *Journal of Experimental Child Psychology*, 42, 392–415.