



Effects of speech rate and practice on the allocation of visual attention in multiple object naming

Antje S. Meyer^{1*}, Linda Wheeldon², Femke van der Meulen² and Agnieszka Konopka³

¹ Max Planck Institute for Psycholinguistics and Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, Netherlands

² School of Psychology, University of Birmingham, Birmingham, UK

³ Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

Edited by:

Andriy Myachykov, University of Glasgow, UK

Reviewed by:

Robert J. Hartsuiker, University of Ghent, Belgium

Zenzi M. Griffin, University of Texas, USA

*Correspondence:

Antje S. Meyer, Max Planck Institute for Psycholinguistics and Donders Institute for Brain, Cognition and Behavior, Radboud University, Postbus 310, 6500 AH Nijmegen, Netherlands.
e-mail: antje.meyer@mpi.nl

Earlier studies had shown that speakers naming several objects typically look at each object until they have retrieved the phonological form of its name and therefore look longer at objects with long names than at objects with shorter names. We examined whether this tight eye-to-speech coordination was maintained at different speech rates and after increasing amounts of practice. Participants named the same set of objects with monosyllabic or disyllabic names on up to 20 successive trials. In Experiment 1, they spoke as fast as they could, whereas in Experiment 2 they had to maintain a fixed moderate or faster speech rate. In both experiments, the durations of the gazes to the objects decreased with increasing speech rate, indicating that at higher speech rates, the speakers spent less time planning the object names. The eye–speech lag (the time interval between the shift of gaze away from an object and the onset of its name) was independent of the speech rate but became shorter with increasing practice. Consistent word length effects on the durations of the gazes to the objects and the eye–speech lags were only found in Experiment 2. The results indicate that shifts of eye gaze are often linked to the completion of phonological encoding, but that speakers can deviate from this default coordination of eye gaze and speech, for instance when the descriptive task is easy and they aim to speak fast.

Keywords: utterance planning, speech rate, practice, eye movements, visual attention

INTRODUCTION

We can talk in different ways. We can, for instance, use a special register, child directed speech, to talk to a child, and we tend to deliver speeches, and formal lectures in a style that is different from casual dinner table conversations. The psychological processes underlying the implementation of different speech styles have rarely been studied. The present paper concerns one important feature distinguishing different speech styles, i.e., speech rate. It is evident that speakers can control their speech rate, yet little is known about how they do this.

To begin to explore this issue we used a simple speech production task: speakers named sets of pictures in sequences of nouns (e.g., “kite, doll, tap, sock, whale, globe”). Each set was shown on several successive trials. In the first experiment, the speakers were asked to name the pictures as fast as they could. In the second experiment, they had to maintain a fixed moderate or faster speech rate, which allowed us to separate the effects of speech rate and practice. Throughout the experiments, the speakers’ eye movements were recorded along with their spoken utterances. In the next sections, we motivate this approach, discuss related studies, and explain the predictions for the experiments.

SPEECH-TO-GAZE ALIGNMENT IN DESCRIPTIVE UTTERANCES

In many language production studies participants have been asked to name or describe pictures of one or more objects. Though probably not the most common way of using language, picture naming is popular in language production research because it offers good

control of the content of the speakers’ utterances and captures a central component of speech planning, namely the retrieval of words from the mental lexicon.

In some picture naming studies, the speakers’ eye movements were recorded along with their speech. This is useful because a person’s eye gaze reveals where their visual attention is focused, that is, which part of the environment they are processing with priority (e.g., Deubel and Schneider, 1996; Irwin, 2004; Eimer et al., 2007). In picture naming, visual attention, and eye gaze are largely controlled endogenously (i.e., governed by the speaker’s goals and intentions), rather than exogenously (i.e., by environmental stimuli). That is, speakers actively direct their gaze to the objects they wish to focus on. Therefore, eye movements provide not only information about the speaker’s visual processing, but also, albeit more indirectly, about the executive control processes engaged in the task (for discussions of executive control processes see Baddeley, 1986; Posner and Petersen, 1990; Miyake et al., 2000).

The eye movement studies of language production have yielded a number of key findings. First, when speakers name sets of objects they typically look at each of the objects in the order of mention, just before naming it (e.g., Meyer et al., 1998; Griffin, 2001). When speakers describe cartoons of events or actions, rather than naming individual objects, there can be a brief apprehension phase during which speakers gain some understanding of the gist of the scene and during which their eye movements are not related in any obvious way to the structure of the upcoming utterances, but following this, there is again a tight coupling between eye gaze and

speech output, with each part of the display being inspected just before being mentioned (Griffin and Bock, 2000; Bock et al., 2003; but see Gleitman et al., 2007).

A second key result is that the time speakers spend looking at each object (hereafter, gaze duration) depends not only on the time they need for the visual–conceptual processing of the object (e.g., Griffin and Oppenheimer, 2006) but also on the time they require to select a suitable name for the object and to retrieve the corresponding word form. This has been shown in studies where the difficulty of identifying the objects, the difficulty of retrieving their names from the lexicon, or the difficulty of generating the corresponding word forms was systemically varied. All of these manipulations affected how long the participants looked at the objects (e.g., Meyer et al., 1998; Griffin, 2001, 2004; Belke and Meyer, 2007). For the present research, a particularly important finding is that speakers look longer at objects with long names than at objects with shorter names (e.g., Meyer et al., 2003, 2007; Korvorst et al., 2006; but see Griffin, 2003). This indicates that speakers usually only initiate the shift of gaze and attention to a new object after they have retrieved the name of the current object (Roelofs, 2007, 2008a,b). A likely reason for the late shifts of gaze and attention is that attending to an object facilitates not only its identification but also the retrieval of any associated information, including the object name (e.g., Wühr and Waszak, 2003; Wühr and Frings, 2008). This proposal fits in well with results demonstrating that lexical access is not an automatic process, but requires some processing capacity (e.g., Ferreira and Pashler, 2002; Cook and Meyer, 2008; Roelofs, 2008a,b) and would therefore benefit from the allocation of attention. The same should hold for speech-monitoring processes (for reviews and further discussion see Postma, 2000; Hartsuiker et al., 2005; Hartsuiker, 2006; Slevc and Ferreira, 2006), which are capacity demanding and might also benefit from focused visual attention to the objects being described (e.g., Oomen and Postma, 2002).

EMPIRICAL FINDINGS ON EYE–SPEECH COORDINATION AT DIFFERENT SPEECH RATES

The studies reviewed above demonstrated that during naming tasks, the speakers' eye movements are tightly coordinated in time with their speech planning processes, with speakers typically looking at each object until they have planned its name to the level of the phonological form. This coupling of eye gaze and speech planning is not dictated by properties of the visual or the linguistic processing system. Speakers can, of course, choose to coordinate their eye gaze and speech in different ways, moving their eyes from object to object sooner, for instance as soon as they have recognized the object, or much later, for instance after they have produced, rather than just planned, the object's name. In this section, we review studies examining whether the coordination of eye gaze and speech varies with speech rate. One would expect that when speakers aim to talk fast, they should spend less time planning each object name. Given that the planning times for object names have been shown to be reflected in the durations of the gazes to the objects, speakers should show shorter gaze durations at faster speech rates. In addition, the coordination of eye gaze and speech might also change. At higher speech rates, speakers might, for instance, plan further ahead, i.e., initiate the shift of gaze to a new

object earlier relative to the onset of the object name, in order to insure the fluency of their utterances.

Spieler and Griffin (2006) asked young and older speakers (average ages: 20 vs. 75 years, respectively) to describe pictures in utterances such as “The crib and the limousine are above the needle.” They found that the older speakers looked longer at the objects and took longer to initiate and complete their utterances than the younger ones. However, the temporal coordination of gaze with the articulation of the utterances was very similar for the two groups. Before speech onset, both groups looked primarily at the first object and spent similar short amounts of time looking at the second object. Belke and Meyer (2007, Experiment 1) obtained similar results. Older speakers spoke more slowly than younger speakers and inspected the pictures for longer, but the coordination between eye gaze and speech in the two groups was similar.

Mortensen et al. (2008) also found that older speakers spoke more slowly and looked at the objects for longer than younger speakers. However, in this study the older participants had shorter eye–speech lags than younger speakers. Griffin (2003) reported a similar pattern of results. She asked two groups of college students attending schools in different regions of the US to name object pairs in utterances such as “wig, carrot.” For unknown reasons, one group of participants articulated the object names more slowly than the other group. Before speech onset, the slower talkers spent more time looking at the first object and less time looking at the second object than the fast talkers, paralleling the findings obtained by Mortensen and colleagues for older speakers. Thus, compared to the fast talkers, the slower talkers delayed the shift of gaze and carried out more of the phonetic and articulatory planning of the first object name while still attending to that object.

These studies involved comparisons of speakers differing in their habitual speech rates. By contrast, Belke and Meyer (2007, Experiment 2) asked one group of young participants to adopt a speech rate that was slightly higher than the average rate used by the young participants in an earlier experiment (Belke and Meyer, 2007, Experiment 1, see above) or a speech rate that was slightly lower than the rate adopted by older participants in that experiment. As expected, these instructions affected the speakers' speech rates and the durations of their gazes to the objects. In line with the results obtained by Mortensen et al. (2008) and by Griffin (2003), the eye–speech lag was much shorter at the slow than at the fast speech rate.

To sum up, in object naming tasks, faster speech rates are associated with shorter gazes to the objects. Given the strong evidence linking gaze durations to speech planning processes, these findings indicate that when speakers increase their speech rate, they spend less time planning their words (see also Dell et al., 1997). While some studies found no change in the coordination of eye gaze and speech, others found shorter eye–speech lags during slow than during faster speech. Thus, during slow speech, the shift of gaze from the current to the next object occurred later relative to the onset of current object name than during faster speech. It is not clear why this is the case. Perhaps slow speech is often carefully articulated speech and talkers delay the shift of gaze in order to carry out more of the phonetic and articulatory planning processes for an object name while still attending to that object. As Griffin

(2003) pointed out, speakers do not need to look ahead much in slow speech because they have ample time to plan upcoming words during the articulation of the preceding words.

THE PRESENT STUDY

Most of the studies reviewed above concerned comparisons between groups of speakers differing in their habitual speech rate. Interpreting their results is not straightforward because it is not known why the speakers preferred different speech rates. So far, the study by Belke and Meyer (2007) is, to our knowledge, the only one where eye movements were compared when one group of speakers used different speech rates, either a moderate or a very slow rate.

The goal of the present study was to obtain additional evidence about the way speakers coordinate their eye movements with their speech when they adopt different speech rates. Gaze durations indicate when and for how long speakers direct their visual attention to each of the objects they name. By examining the speaker's eye movements at different speech rates, we can determine how their planning strategies – the time spent planning each object name and the temporal coordination of planning and speaking – might change.

Whereas speakers in Belke and Meyer's (2007) study used a moderate or a very slow speech rate, speakers in the first experiment of present study were asked to increase their speech rate beyond their habitual rate and to talk as fast as they could. To the best of our knowledge no other study has used these instructions, though the need to talk fast regularly occurs in everyday conversations.

Participants saw sets of six objects each (see **Figure 1**) and named them as fast as possible. There were eight different sets, four featuring objects with monosyllabic names and four featuring objects with disyllabic names (see Appendix). In Experiment 1, there were two test blocks, in each of which each set was named on eight successive trials. We recorded the participants'

eye movements and speech onset latencies and the durations of the spoken words. We asked the participants to name the same objects on successive trials (rather than presenting new objects on each trial) to make sure that they could substantially increase their speech rate without making too many errors. An obvious drawback of this procedure was that the effects of increasing speech rate and increasing familiarity with the materials on the speech-to-gaze coordination could not be separated. We addressed this issue in Experiment 2.

Based on the results summarized above, we expected that speakers would look at most of the objects before naming them and that the durations of the gazes to the objects would decrease with increasing speech rate. The eye–speech lags should either be unaffected by the speech rate or increase with increasing speech rate. That is, as the speech becomes faster speakers might shift their gaze earlier and carry out more of the planning of the current word without visual guidance.

We compared gaze durations for objects with monosyllabic and disyllabic names. As noted, several earlier eye tracking studies had shown that speakers looked longer at objects with long names than at objects with shorter names (e.g., Meyer et al., 2003, 2007; Korvorst et al., 2006; but see Griffin, 2003). This indicates that the speakers only initiated the shift of gaze to a new object after they had retrieved the phonological form of the name of the current object. In these studies no particular instructions regarding speech rate were given. If speakers consistently time the shifts of gaze to occur after phonological encoding of the current object name has been completed, the word length effect should be seen regardless of the speech rate. By contrast, if at high speech rates, speakers initiate the shifts of gaze from one object to the next earlier, before they have completed phonological encoding of the current object name, no word length effect on gaze durations should be seen.

EXPERIMENT 1

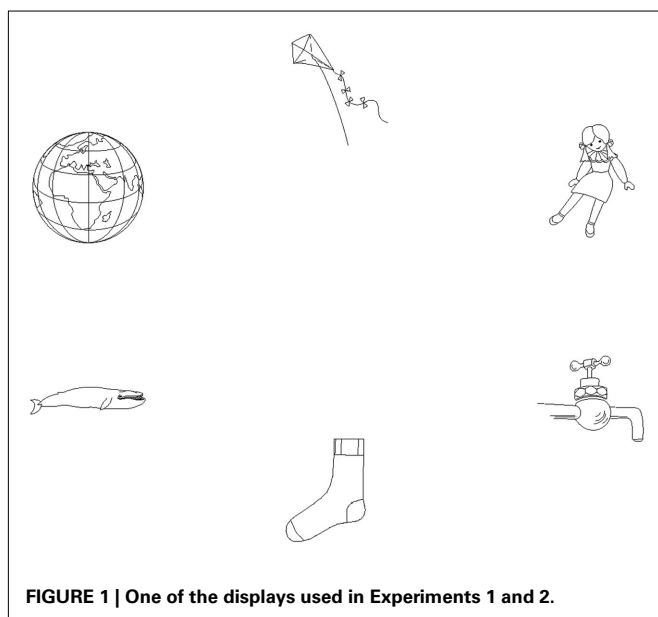
METHOD

Participants

The experiment was carried out with 24 undergraduate students of the University of Birmingham. They were native speakers of British English and had normal or corrected-to-normal vision. They received either payment or course credits for participation. All participants were fully informed about the details of the experimental procedure and gave written consent. Ethical approval for the study had been obtained from the Ethics Board of the School of Psychology at the University of Birmingham.

Materials and design

Forty-eight black-and-white line drawings of common objects were selected from a picture gallery available at the University of Birmingham (see Appendix). The database includes the Snodgrass and Vanderwart (1980) line drawings and others drawn in a similar style. Half of the objects had monosyllabic names and were on average 3.1 phonemes in length. The remaining objects had disyllabic names and were on average 5.1 phonemes in length. The disyllabic names were mono-morphemic and stressed on the first syllable. The monosyllabic and disyllabic object names were matched for frequency (mean CELEX lexical database, 2001, word



form frequencies per million words: 12.1 for monosyllabic words and 9.9 for disyllabic words).

We predicted that the durations of the gazes to the objects should vary in line with the length of the object names because it takes longer to construct the phonological form of long words than of short words. It was therefore important to ensure that the predicted Word Length effect could not be attributed to differences between the two sets in early visual–semantic processing. Therefore, we pre-tested the items in a word–picture matching task (see Jescheniak and Levelt, 1994; Stadthagen–Gonzalez et al., 2009).

The pretest was carried out with 22 undergraduate participants. On each trial, they saw one of the experimental pictures, preceded by its name or an unrelated concrete noun, which was matched to the object name for word frequency and length. Participants indicated by pressing one of two buttons whether or not the word was the name of the object. All objects occurred in the match and mismatch condition. Each participant saw half of the objects in each of the two conditions, and the assignment of objects to conditions was counterbalanced across participants. The error rate was low (2.38%) and did not differ significantly across conditions. Correct latencies between 100 and 1000 ms were analyzed in analyses of variance (ANOVAs) using length (monosyllabic vs. disyllabic) and word–picture match (match vs. mismatch) as fixed effects and either participants or items as random effects (F_1 and F_2 , respectively). There was a significant main effect of word–picture match, favoring the match condition [478 ms (SE = 11 ms, by participants) vs. 503 ms (SE = 9 ms); $F_1(1, 21) = 15.5, p = 0.001$; $F_2(1, 46) = 4.6, p = 0.037$]. There was also a main effect of length, favoring the *longer* names [474 ms (SE = 11 ms) vs. 507 ms (SE = 10 ms), $F_1(1, 21) = 31.1, p < 0.001$; $F_2(1, 46) = 7.5, p = 0.009$]. The interaction of the two variables was not significant (both $F_s < 1$). Note that the difference in picture matching speed between the monosyllabic and disyllabic object sets was in the opposite direction than would be predicted on the basis of word length. If we observe the predicted effects of Word Length in the main experiment, they cannot be attributed to differences between the monosyllabic and disyllabic sets in early visual–conceptual processes.

The 24 objects with monosyllabic names and the 24 objects with disyllabic names were each combined into 4 sequences of 6 objects. The names in each sequence had different onset consonants, and each sequence included only one complex consonant onset. Care was taken to avoid close repetition of consonants across other word positions. The objects in each sequence belonged to different semantic categories. The pictures were sized to fit into rectangular areas of $3^\circ \times 3^\circ$ visual angle and arranged in an oval with a width of 20° and a height of 15.7° .

Half of the participants named the sequences of objects with monosyllabic names and the other half named the disyllabic sequences. There were two test blocks. In each block, each display was shown on 8 successive trials, creating the total of 64 trials for every participant. The first presentation of each sequence was considered a warm-up trial and was excluded from all statistical analyses.

Apparatus. The experiment was controlled by the experimental software package NESU provided by the Max Planck Institute

for Psycholinguistics, Nijmegen. The pictures were presented on a Samtron 95 Plus 19" screen. Eye movements were monitored using an SMI EyeLink Hispeed 2D eye tracking system. Throughout the experiment, the x - and y -coordinates of the participant's point of gaze for the right eye were estimated every 4 ms. The positions and durations of fixations were computed online using software provided by SMI. Speech was recorded onto the hard disk of a GenuineIntel computer (511 MB, Linux installed) using a Sony ECM-MS907 microphone. Word durations were determined off-line using PRAAT software.

Procedure. Participants were tested individually in a sound-attenuated booth. Before testing commenced, they received written instructions and a booklet showing the experimental objects and their names. After studying these, they were asked to name the objects shown in another booklet where the names were not provided. Any errors were corrected by the experimenter. Then a practice block was run, in which the participants saw the objects on the screen one by one and named them. Then the headband of the eye tracking system was placed on the participant's head and the system was calibrated.

Speakers were told they would see sets of six objects in a circular arrangement, and that they should name them in clockwise order, starting with the object at the top. They were told that on the first presentation of a display, they should name the objects slowly and accurately, and on the seven following presentations of the same display they should aim to name the objects as quickly as possible.

At the beginning of each trial a fixation point was presented in the top position of the screen for 700 ms. Then a picture set was presented until the participant had articulated the sixth object name. The experimenter then pressed a button, thereby recording the speakers' utterance duration and removing the picture from the screen. The mean utterance duration was calculated over the eight repetitions of each set and displayed on the participant's monitor to encourage them to increase their speech rate. (These approximate utterance durations were only used to provide feedback to the participants but not for the statistical analyses of the data.) The experimenter provided additional feedback, informing the participants that their speech rate was good but encouraging them to speak faster on the next set of trials. The same procedure was used in the second block, except that the experimenter provided no further feedback. The inter-trial interval was 1250 ms.

RESULTS

Results from both experiments were analyzed with ANOVAs using subjects as a random factor, followed by linear mixed effects models and mixed logit models (Baayen et al., 2008; Jaeger, 2008). In the latter, all variables were centered before model estimates were computed. All models included participants and items (i.e., the four sequences of objects with monosyllabic names or the four sequences of objects with disyllabic names) as random effects. In Experiment 1, the fixed effects were Word Length (monosyllabic vs. disyllabic words), Block (First vs. Second Block), and Repetition. Repetition was included as a numerical predictor. Variables that did not reliably contribute to model fit were dropped. In models with interactions, only the highest-level interactions are reported below.

Error rates

Errors occurred in 7.5% of the sequences, corresponding to a rate of 1.25% of the words. Of the 115 errors, the majority were hesitations (28 errors) or anticipations of words or sounds (39 errors). The remaining errors were 9 perseverations, 6 exchanges, and 33 non-contextual errors, where object names were produced that did not appear in the experimental materials.

Inspection of the error rates showed no consistent increase or decrease across the repetitions of the picture sets. The ANOVA of the error rates yielded a significant main effect of Block [$F(1, 22) = 5.89, p = 0.024$] and a significant interaction of Block and Word Length [$F(1, 22) = 4.89, p = 0.036$]. This interaction arose because in the first block the error rate was higher for monosyllabic than for disyllabic items [11.90% (SE = 2.2%) vs. 7.74% (SE = 2.30%)], whereas the reverse was the case in the second block [4.46% (SE = 1.40%) vs. 7.74% (SE = 2.08%)]. The interaction of Block, Repetition, and Word Length was also significant [$F(6, 132) = 2.23, p = 0.044$]. No other effects approached significance. The mixed logit analysis of errors also showed an interaction between Block and Word Length ($\beta = 1.05, SE = 0.44, z = 2.41$) as well as an interaction between Word Length and Repetition ($\beta = 0.19, SE = 0.11, z = 1.82$). All trials on which errors occurred were eliminated from the following analyses.

Speech onset latencies

One would expect that the instruction to talk fast might affect not only speech rate, but also speech onset latencies. The average latencies for correct trials are displayed in **Figure 2**. Any latencies below 150 ms or above 1800 ms (1.1% of the data) had been excluded. In the ANOVA the main effect of Block was significant [$F(1, 22) = 87.3, p < 0.001$], as was the main effect of Repetition [$F(6, 132) = 13.1, p < 0.001$; $F(1, 22) = 34.93, p < 0.001$ for the linear trend]. **Figure 2** suggests longer latencies for monosyllabic than for disyllabic items, but this difference was not significant [$F(1, 22) = 1.00, p = 0.33$].

The best-fitting mixed effects model included main effects of Block and Repetition and an interaction between Block and Repetition ($\beta = 9, SE = 3.41, t = 2.67$) reflecting the fact that the effect

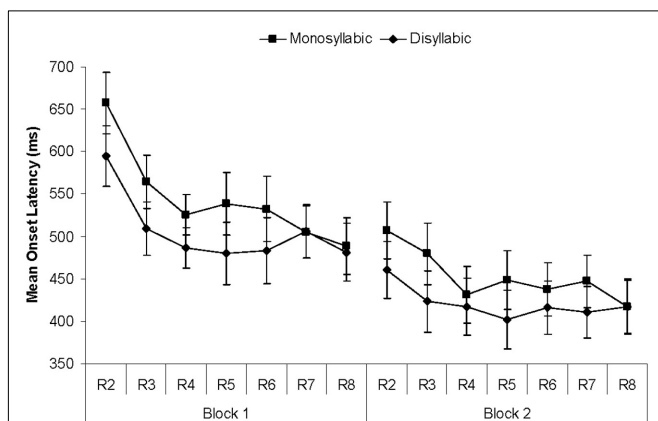


FIGURE 2 | Results of Experiment 1. Mean utterance onset latencies (ms) per block and repetition (R2 through R8) for monosyllabic and disyllabic items. Error bars show SEs.

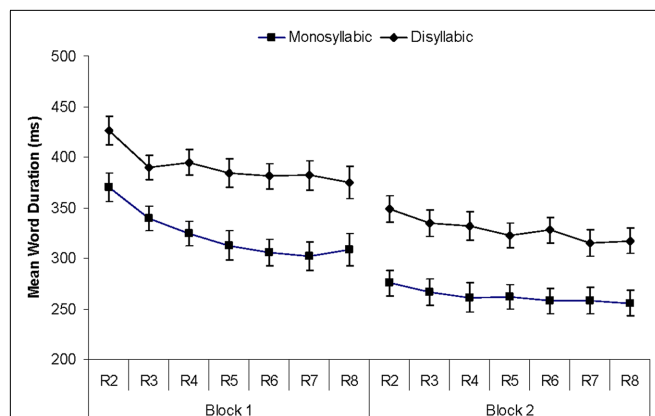


FIGURE 3 | Results of Experiment 1. Mean word durations (ms) per block and repetition (R2 through R8) for monosyllabic and disyllabic items.

of Repetition was stronger in the first than in the second block. There was also an interaction between Word Length and Repetition ($\beta = 9, SE = 3.41, t = 2.58$), as speech onsets declined over time more quickly for monosyllabic than disyllabic words. Model fit was also improved by including by-participant random slopes for Block.

Word durations

To determine how fast participants produced their utterances, we computed the average word duration for each sequence by dividing the time interval between speech onset and the offset of the last word by six¹. As **Figure 3** shows, word durations were consistently shorter for monosyllabic than for disyllabic items; they were shorter in the second than in the first block, and they decreased across the repetitions of the sequences.

In the ANOVA, we found significant main effects of Word Length [$F(1, 22) = 15.6, p = 0.001$], Block [$F(1, 22) = 143.96, p < 0.001$], and Repetition [$F(6, 132) = 38.02, p < 0.001$; $F(1, 22) = 125.44, p < 0.001$ for the linear trend]. The interaction of Block and Repetition was also significant [$F(6, 132) = 7.22, p < 0.001$], as was the interaction of Word Length, Block, and Repetition [$F(6, 132) = 2.86, p = 0.012$]. The interaction is due to the steeper decrease in word durations in Block 1 for monosyllabic than disyllabic words. The mixed effects model showed an analogous three-way interaction ($\beta = -6, SE = 2.48, t = -2.29$), along with main effects of all three variables. Model fit was also improved by including by-participant random slopes for Block.

Gaze paths

To analyze the speakers' eye movements, we first determined the gaze path for each trial, i.e., established whether all objects were inspected, and in which order they were inspected. On 78.9% of

¹The word durations included any pauses between words. Additional analyses were carried out for word durations measured from word onset to word offset and for the distribution and durations of any pauses, but, with respect to the main question of interest, these analyses provided no additional information. Both word and pause durations decreased across the repetitions of the materials and from the first to the second block.

the trials, the speakers looked at the six objects in the order of mention (simple paths). On 13.2% of the trials they failed to look at one of the objects (skips). As there were six objects in a sequence, this means that 2.2% of the objects were named without being looked at. On 4.5% of trials speakers looked back at an object they had already inspected (regressions). The remaining 3.3% of trials featured multiple skips and/or regressions.

Statistical analyses were carried out for the two most common types of gaze paths, simple paths, and paths with skips. The analysis of the proportion of simple paths yielded no significant effects. The ANOVA of the proportions of paths with skips yielded only a significant main effect of Block [$F(1, 22) = 6.77, p = 0.016$], with participants being less likely to skip one of the six objects of a sequence in the first than in the second block [8.1% (SE = 2.3%) vs. 21.0% (SE = 5.1%)]. The best-fitting mixed logit model included an effect of Block ($\beta = 1.04, SE = 0.42, t = 2.49$) and an effect of Repetition ($\beta = 0.12, SE = 0.05, t = 2.47$). The model also included an interaction between Block and Word Length, but including random by-participant slopes for Block reduced the magnitude of this interaction ($\beta = -0.63, SE = 0.83, t = -0.76$). This suggests that between-speaker differences in word durations across the two blocks largely accounted for the increase of skips on monosyllabic objects in the second block.

Gaze durations

For each trial with a simple gaze path or a single skip we computed the average gaze duration across the first five objects of the sequence. The gazes to the sixth object were excluded as participants tend to look at the last object of a sequence until the end of the trial. Durations of less than 80 ms or more than 1200 ms were excluded from the analysis (1.1% of the trials).

As **Figure 4** shows, gaze durations decreased from the first to the second block and across the repetitions within blocks, as predicted. In the first block, they were consistently longer for disyllabic than for monosyllabic items, but toward the end of the second block the Word Length effect disappeared. The ANOVA of the gaze durations yielded main effects of Block [$F(1, 22) = 21.41, p < 0.001$], and Repetition [$F(6, 132) = 5.39, p < 0.001; F(1, 22) = 7.35, p = 0.013$ for the linear trend]. The interaction of Block and Repetition was

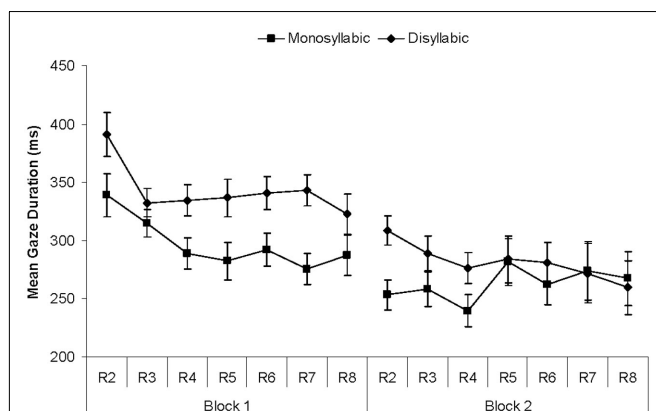


FIGURE 4 | Results of Experiment 1. Mean gaze durations (ms) per block and repetition (R2 through R8) for monosyllabic and disyllabic items.

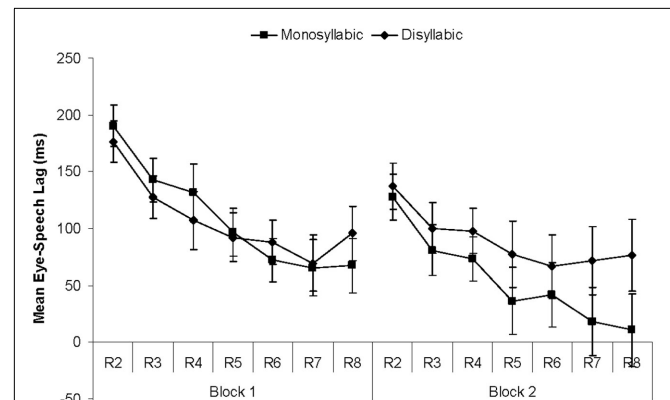


FIGURE 5 | Results of Experiment 1. Mean eye-speech lags (ms) per block and repetition (R2 through R8) for monosyllabic and disyllabic items.

also significant [$F(6, 132) = 3.14, p = 0.007$], as the effect of Repetition was larger in the first than in the second block. The main effect of Word Length was marginally significant [$F(1, 22) = 3.88, p = 0.062$]. Finally, the three-way interaction was also significant [$F(6, 132) = 2.21, p = 0.05$]. Separate ANOVAs for each block showed that in the first block the main effect of Word Length was significant [$F(1, 22) = 6.39, p = 0.019$], as was the effect of Repetition [$F(6, 132) = 11.49, p < 0.001$]. In the second block, neither of the main effects nor their interaction were significant [$F < 1$ for the main effects, $F(6, 132) = 1.67$ for the interaction]. The best-fitting mixed effects model included an interaction between all three factors ($\beta = -10, SE = 3.47, t = -2.96$), along with three significant main effects. Including random by-participant slopes for Block improved model fit.

Eye-speech lags

To determine the coordination of eye gaze and speech we calculated the lag between the offset of gaze to an object and the onset of its spoken name. As **Figure 5** shows, the lags decreased significantly from the first to the second block [$F(1, 22) = 11.56, p = 0.001$] and across the repetitions within blocks [$F(6, 132) = 21.53, p < 0.001; F(1, 22) = 66.17, p < 0.001$ for the linear trend]. The interaction of Block by Repetition was also significant [$F(6, 132) = 2.26, p < 0.05$]. Finally, the interaction of Word Length by Block approached significance [$F(1, 22) = 3.67, p < 0.07$]. As **Figure 5** shows, in the first block the lags for monosyllabic and disyllabic items were quite similar, but in the second block, lags were longer for disyllabic than for monosyllabic items.

The best-fitting mixed effects model included an interaction between Block and Word Length ($\beta = 35, SE = 18, t = 1.96$) and between Repetition and Word Length ($\beta = 7, SE = 3, t = 2.67$), as well as by-participant slopes for Block. Including an interaction between Block and Repetition, however, did not improve model fit [$\chi^2(1) = 1.35$, when comparing models with and without this interaction].

DISCUSSION

In Experiment 1, participants were asked to increase their speech rate across the repetitions of the materials as much as they could

without making too many errors. The analyses of participants' speech and error rates showed that they followed these instructions well: speech onset latencies and spoken word durations decreased from the first to the second block and across the repetitions within each block, while error rates remained low². The speakers' eye gaze remained tightly coordinated with their speech: most of the objects were inspected, just once, shortly before they were named, and the durations of the gazes to the objects decreased along with the spoken word durations. Deviating from earlier findings, we found that the eye–speech lags decreased, rather than increased, as the speech became faster. We return to this finding in the Section “General discussion.”

In addition, we observed subtle changes in the coordination of eye gaze and speech: in the second block, the objects were more likely than in the first block to be named without being fixated first, and there was a Word Length effect on gaze durations in the first but not in the second block. This indicates that in the first block the participants typically looked at each object until they had retrieved the phonological form of its name, as participants in earlier studies had done (e.g., Korvorst et al., 2006; Meyer et al., 2007), but did not do this consistently in the second block. As **Figures 4** and **5** show, in the second half of the second block, the durations of the gazes to monosyllabic and disyllabic items were almost identical, but the eye–speech lag was much longer for disyllabic than monosyllabic items. Apparently, participants disengaged their gaze from monosyllabic and disyllabic items at about the same time, perhaps as soon as the object had been recognized, but then needed more time after the shift of gaze to plan the disyllabic words and initiate production of these names.

The goal of this experiment was to explore how speakers would coordinate their eye gaze and speech when they tried to speak as fast as possible. In order to facilitate the use of a high speech rate, we presented the same pictures on several successive trials. This meant that the effects of increasing speech rate and practice were confounded. Either of those effects might be responsible for the change of the eye–speech coordination from the first to the second block. To separate the effects of practice and speech rate, a second experiment was conducted, where participants were first trained to produce the object names either at a fast or more moderate pace, and then named each display on 20 successive trials at that pace.

EXPERIMENT 2

METHOD

Participants

The experiments were carried out with 20 undergraduate students of the University of Birmingham. They had normal or corrected-to-normal vision and received either payment or course credits for participation. The participants received detailed information about the experimental procedure and gave written consent to participate.

²The speech planning model proposed by Dell et al. (1997) predicts that the proportion of anticipatory errors in a set of movement errors should increase as the overall error rate decreases. In our experiment, the total number of movement errors in the first and second block was almost the same (25 vs. 23 errors), but the proportion of anticipatory errors was much higher in the second than in the first block [96% (1 out of 23) vs. 68% (8 out of 25)]. This finding is in line with the assumption of the model that practice strengthens the links between the plan and upcoming units.

Materials and design

The same experimental sequences were used as in Experiment 1. In addition, four training sequences were constructed, two consisting of objects with monosyllabic names and two consisting of objects with disyllabic names. All participants named the monosyllabic and the disyllabic sequences. Ten participants each were randomly assigned to the Moderate Pace and the Fast Pace group. Five participants in each of these groups were first tested on the monosyllabic sequences and then on the disyllabic sequences. For the remaining participants the order of the sequences was reversed.

Procedure

To encourage participants to adopt different speech rates, we used different presentation times for the pictures. The presentation time for the Fast Pace condition was the average speech rate over the last four repetitions of the sets in Experiment 1. This was 2150 ms for monosyllabic sequences and 2550 ms for disyllabic sequences. This corresponded to speech rates of 3.5 words/s and 2.8 words/s for the monosyllabic and disyllabic words, respectively. The moderate speech rates were 1/3 slower than the fast rates, resulting in 2850 ms of presentation time for the monosyllabic sequences (2.7 words/s) and 3350 ms for the disyllabic ones (2.1 words/s).

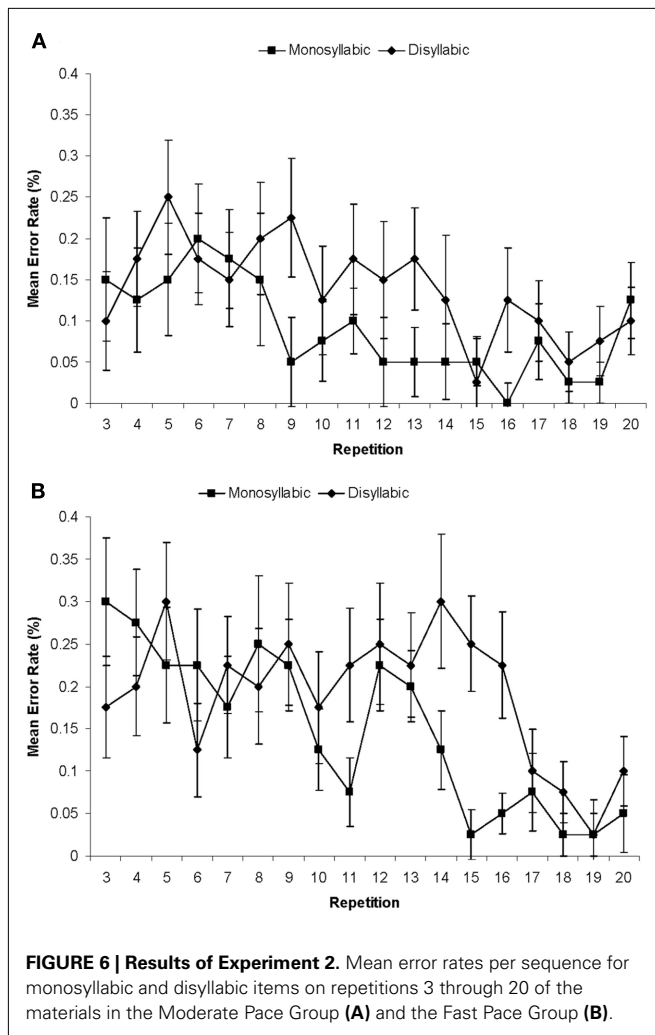
After receiving instructions, speakers named the objects individually, as in Experiment 1. The headband of the eye tracking system was placed on the participant's head and the system was calibrated. Speakers were instructed that they would have to maintain a constant speech rate. A tone was played coinciding with the end of the display time, which was also the time by which all six objects should have been named.

Speakers were trained on a particular pace using the training sequences. The first training sequence was presented four times while a metronome indicated the speech rate for the upcoming block. Then the metronome was switched off and the same training sequence was named eight more times. A second training sequence was then named on 12 repetitions. If the experimenter then felt that more training was required, the training sequences were repeated. When the speaker was successful in naming the training sequences at the required speed, the first four target sequences were presented. Each sequence was shown on 20 successive trials. After a short break the training and testing procedure was repeated for the second set of sequences.

RESULTS

Error rates

The first two presentations of each set were considered warm-up trials and were excluded from all analyses. 11.28% of the remaining trials included one more error. Of the 333 errors, most were hesitations (131 errors), anticipations (112 errors), or incorrect object names (67 errors), which were nouns that were not names of any of the objects in the current display. The remaining errors were 5 perseverations, 10 exchanges, and 8 combinations of errors. The error rates for the different experimental conditions are shown in **Figure 6**. In the ANOVA only the main effect of Repetition was significant [$F(17, 306) = 3.38, p < 0.001$] with errors becoming less frequent across the repetitions $F(1, 18) = 23.14, p < 0.001$ for the linear trend). The mixed logit analysis of errors

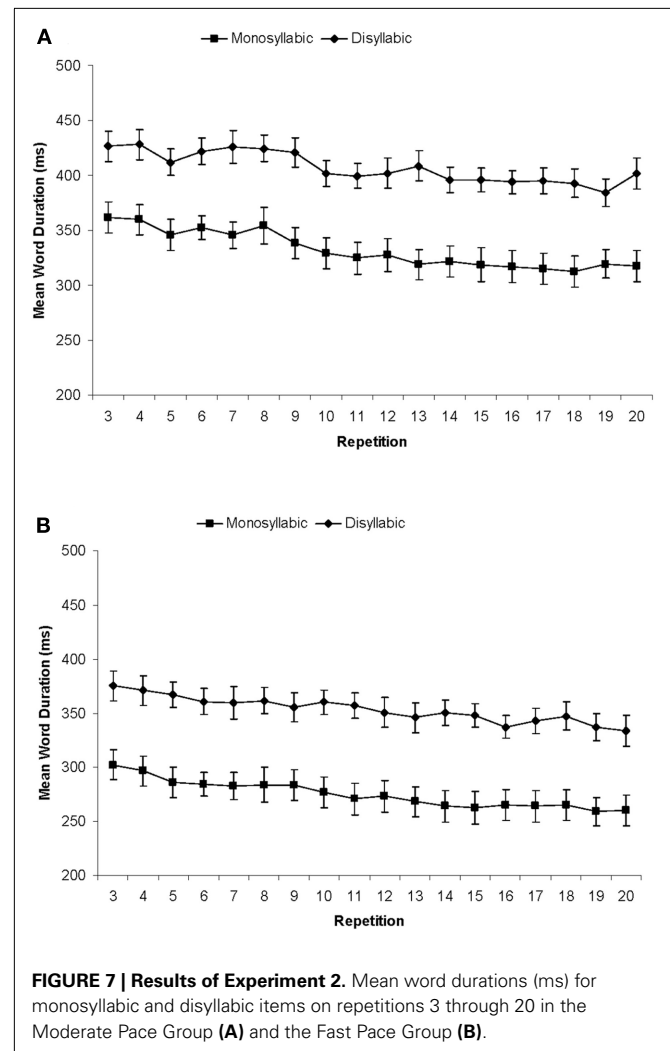


showed an interaction between Repetition and Word Length ($\beta = 0.06$, $SE = 0.02$, $z = 2.30$), as the number of errors on disyllabic sequences dropped to the level of errors on monosyllabic sequences over time. Model fit was improved by including random by-participant slopes for Word Length.

Speech onset latencies

The participants were instructed to maintain specific speech rates across the repetitions of the materials. The Fast Pace speakers timed their utterances well and completed the monosyllabic sequences on average after 2017 ms (target: 2150 ms) and the disyllabic sequences after 2506 ms (target: 2550 ms). The Moderate Pace completed the monosyllabic sequences after 2425 ms (target: 2850 ms) and the disyllabic sequences after 2908 ms (target: 3350 ms). This was faster than expected, but there was still a considerable difference in utterance completion time to the Fast Pace group.

The analyses of the speech onset latencies (excluding all errors and 1.1% of the trials with latencies below 150 ms) only yielded a main effect of Word Length, with sequences of monosyllabic words being initiated faster than sequences



of disyllabic words [means: 456 ms ($SE = 19$ ms) vs. 490 ms ($SE = 22$ ms); $F(1, 18) = 7.66$; $p = 0.013$]. The best-fitting mixed effects model included a marginal effect of Word Length ($\beta = 34.12$, $SE = 18.69$, $t = 1.83$) and by-participant random slopes for this factor.

Word durations

Analyses of variance of word durations showed the expected effects of Word Length [$F(1, 18) = 641.31$; $p < 0.001$] and Pace [$F(1, 18) = 10.64$; $p < 0.001$]. The main effect of Repetition was also significant [$F(17, 306) = 22.06$, $p < 0.001$]. As **Figure 7** shows, word durations decreased across the repetitions of the materials, yielding a significant linear trend [$F(1, 18) = 72.41$, $p < 0.001$]. There were no significant interactions.

In the mixed effects linear model, all three factors contributed to model fit as additive effects (Word Length: $\beta = 83$, $SE = 9$, $t = 8.82$; Pace: $\beta = -51$, $SE = 15$, $t = -3.29$; and Repetition: $\beta = -2$, $SE = 0.28$, $t = -8.54$). Model fit was also improved by including random by-participant slopes for Word Length and Repetition, as well as random by-item slopes for Pace.

Gaze paths

As in Experiment 1, the participants usually (on 78.51% of the trials) looked at all objects in the order of mention (simple gaze paths). On 17.87% of trials they skipped one of the six objects. They produced regressions (looking back at an object they had already inspected) on only 0.2% of trials. The remaining 3.42% of trials included multiple skips and/or regressions.

In the analysis of the proportions of simple paths, only the main effect of Pace was significant [$F(1, 18) = 7.50, p = 0.013$] with the speakers using the Moderate Pace being more likely to fixate upon all objects than the speakers using the Fast Pace [means: 85.51% (SE = 0.62%) vs. 71.30% (SE = 0.62%)]. The best-fitting mixed effects model included a three-way interaction ($\beta = -0.09, SE = 0.04, z = -2.17$), with a marginally reliable main effect of Pace and interaction between Pace and Repetition. This pattern arose because speakers using the Moderate Pace, but not the speakers using the Fast Pace, were more likely to fixate objects with disyllabic names than objects with monosyllabic names at later repetitions of the picture sequences. Model fit was also improved by including random by-participant slopes for Word Length and Repetition as well as random by-item slopes for Pace and Repetition.

The analysis of the proportion of trials with skips yielded a complementary pattern: there was only a significant main effect of Pace [$F(1, 18) = 7.74, p = 0.012$], with speakers using the fast pace skipping one of the objects more often (on 23.77% of the trials) than speakers using the moderate pace (10.18% of the trials, SE = 0.45% for both groups).

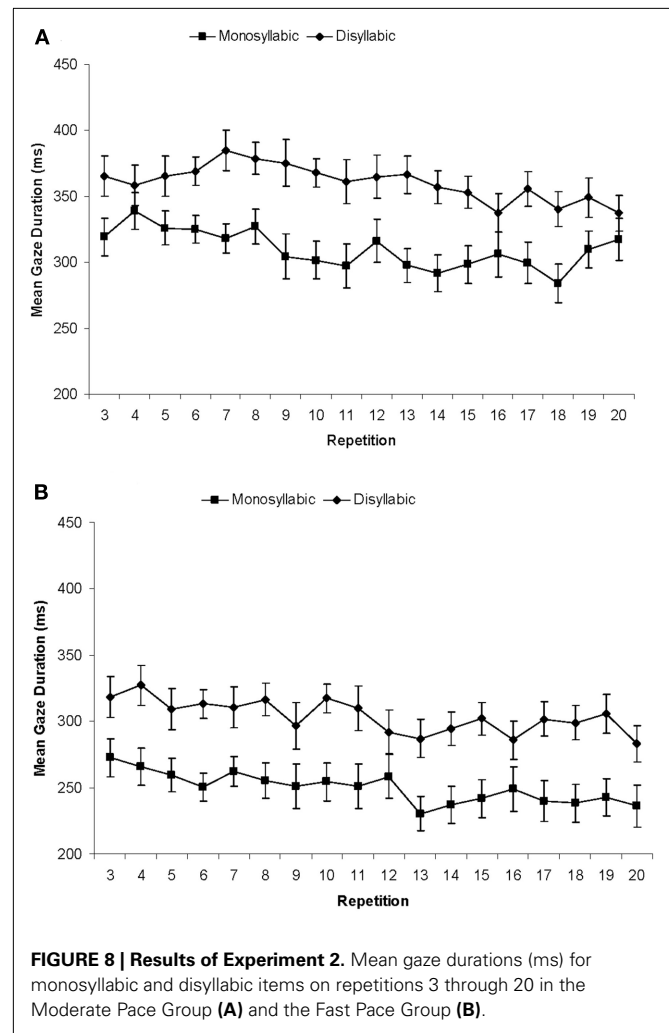
The best-fitting mixed effects model included a main effect of Pace, a two-way interaction between Word Length and Repetition, as well as a three-way interaction ($\beta = 0.12, SE = 0.05, z = 2.38$). The rate at which speakers skipped pictures increased over time, but speakers using the Moderate Pace were less likely to skip pictures with disyllabic than with monosyllabic names. Model fit was also improved by including random by-participant slopes for Word Length and Repetition.

Gaze durations

Gaze durations were calculated in the same way as for Experiment 1. The statistical analysis showed a significant main effect of Pace [$F(1, 18) = 15.47, p = 0.001$], as gazes were shorter at the Fast than the Moderate pace (see **Figure 8**). The main effect of Repetition was also significant [$F(17, 306) = 4.19, p < 0.002$], with gaze durations decreasing across the repetitions of the sequences [$F(1, 18) = 11.85, p = 0.003$ for the linear trend]. Finally, the main effect of Word Length was significant [$F(1, 21) = 123.87, p < 0.001$], with gaze durations being longer, by 53 ms, in the disyllabic than in the monosyllabic sets. There were no significant interactions. The best-fitting mixed effects model included all three factors as additive effects (Word Length: $\beta = 48, SE = 11, t = 4.42$; Pace: $\beta = -55, SE = 13, t = -4.12$; and Repetition: $\beta = -1.24, SE = 0.51, t = -2.42$). Model fit was also improved by including random by-participant slopes for Word Length and Repetition.

Eye–speech lags

In contrast to gaze durations, eye–speech lags were not affected by the Pace, $F < 1$ (see **Figure 9**). There was a significant main



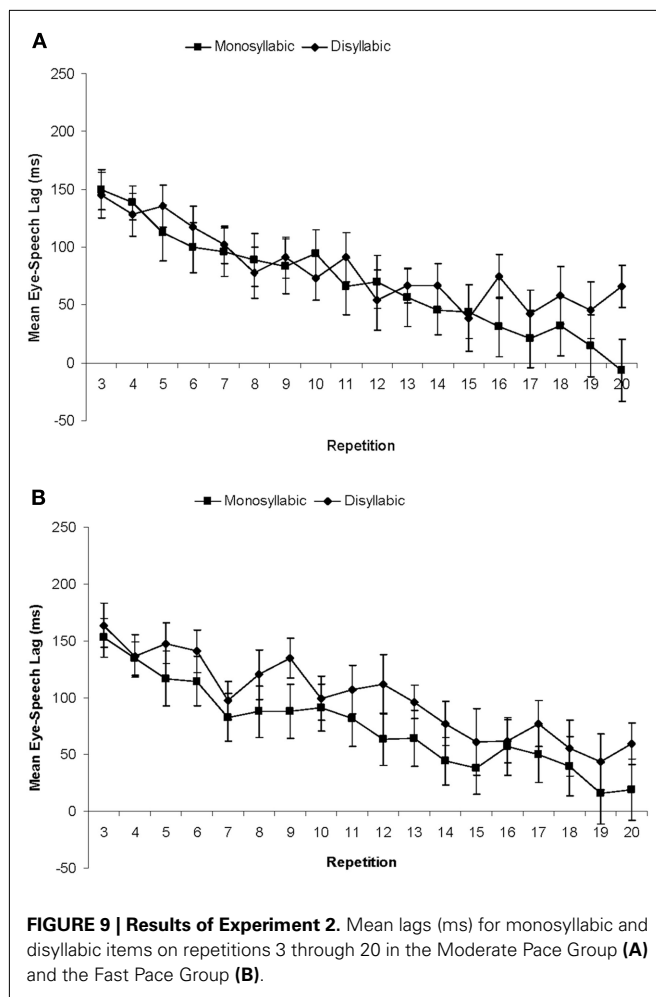
effect of Word Length for the lags [$F(1, 18) = 5.44, p = 0.032$], which were longer by 19 ms for disyllabic than for monosyllabic sets. Lags decreased significantly across the repetitions of the materials [$F(17, 306) = 26.94, p < 0.001$ for the main effect; $F(1, 18) = 99.38, p < 0.001$ for the linear trend].

The best-fitting mixed effects model included Repetition as an additive effect as well as an interaction between Repetition and Word Length ($\beta = 2, SE = 0.67, t = 3.40$), as lags for monosyllabic object names decreased more quickly than for disyllabic object names. Model fit was also improved by including random by-participant slopes for Word Length and Repetition.

DISCUSSION

In Experiment 1, the effects of practice and speech rate on speech-to-gaze alignment were confounded, as participants practiced naming the object sets at increasing speed. In Experiment 2 we aimed to separate these effects by asking participants to name the object sequences repeatedly at a fixed moderate or faster pace.

The comparisons of the Fast vs. Moderate Pace group showed that the difference in speech rate was accompanied by differences in gaze durations, while there was no difference in the eye–speech



lags. Thus, speakers adopting different speech rates differed in how much time they spent looking at and attending to each object, but they did not differ in their planning span.

Evaluating the effects of practice was complicated by the fact that participants, contrary to the instructions, slightly increased their speech rate across the repetitions of the sequences. This increase in speech rate was accompanied by a small decrease in gaze durations and a more substantial decrease in the lags. Apparently those processes carried out before the shift of gaze from one object to the next as well as those carried out after the shift of gaze became faster as participants became more familiar with the materials, and this resulted in the unintended increase in speech rate. The reasons why the lags decreased more than the gaze durations are not clear. One possibility is that focused attention was required to counteract practice effects arising from the repetition of the materials; as soon as attention turned from the current object to the next object, the remaining planning processes were completed at a default pace which increased as the materials became more familiar.

In Experiment 1, we had observed an effect of Word Length on gaze durations in the first but not in the second test block. By contrast, in Experiment 2 the Word Length effect on gaze durations was maintained across the entire experiment, demonstrating that

participants consistently fixated upon each object until they had retrieved the phonological form of its name. Moreover, we found a significant Word Length effect for the eye–speech lag. This effect reflects the fact that the processes the speakers carried out after the shift of gaze – i.e., phonetic and articulatory encoding – took more time for the longer words. No length effect on the lags had been seen in Experiment 1. A possible account of this difference between the experiments is that in Experiment 2, participants typically planned both syllables of the disyllabic items before speech onset, whereas in Experiment 1 they often planned only one syllable before the onset of articulation (Meyer et al., 2003; Damian et al., 2010). In line with this account, Experiment 2 also yielded an effect of Word Length on utterance onset latencies, which had been absent in Experiment 1. This might reflect that in Experiment 2 the participants usually planned the full phonological and phonetic form of the first object name before beginning to speak, whereas in Experiment 1 they often initiated the utterances earlier (for further discussions of word length effects on utterance onset latencies see Bachoud-Lévi et al., 1998; Griffin, 2003, and Meyer et al., 2003). Thus, the different instructions affected not only the speech rates, but also led the participants of Experiment 2 to adopt a more careful, deliberate speech style.

GENERAL DISCUSSION

The goal of these studies was to explore how speakers' gaze-to-speech alignment would change when they used different speech rates. We found that, at all speech rates, the speakers' eye movements were systematically coordinated with their overt speech: they fixated upon most of the objects before naming them, and shorter spoken word durations were accompanied by shorter gazes to the objects. As explained in the Introduction, there is strong evidence that gaze durations reflect on the times speakers take to recognize the objects and to plan their names. Therefore the decreasing gaze durations observed with increasing speech rates show that speakers spend less time attending to each of the objects and preparing their names when they speak fast than when they speak more slowly.

In our study, the eye–speech lag, the time between the shift of gaze from one object to the next and the onset of the name of the first object, was not systematically affected by speech rate. There was no evidence that speakers would plan their utterances further, or less far, ahead when they used different speech rates. This result is consistent with findings reported by Spieler and Griffin (2006) and Belke and Meyer, 2007, Experiment 1). It contrasts with results obtained by Griffin (2003), Mortensen et al. (2008), and Belke and Meyer (2007), who found that slower speech rates were associated with shorter eye–speech lags than faster speech rates. The reasons for this difference in the results are not clear. Griffin (2003) and Mortensen et al. (2008) compared the eye–speech coordination of speakers differing in their habitual speech rate. It is not known why the speakers differed in speech rate. One possibility is that the slower speakers initiated the shift of gaze to a new object later, relative to the speech planning processes, than the faster speakers: whereas the faster speakers directed their eye gaze and attention to a new object as soon as they had retrieved the phonological form of the name of the current object, the slower speakers only initiated the shift a little later, after completing part of the phonetic

or articulatory encoding as well. This would yield longer gazes and shorter lags in slower compared to faster speakers. The slower speakers might have used such a more cautious processing strategy because they monitored their speech more carefully or tried to minimize the interference from new visual input (Griffin, 2003, 2004; Belke and Meyer, 2007). Similarly, when the speakers in Belke and Meyer's study were asked to use a very slow, rather than moderate speech rate, they may have altered the criterion for initiating the shifts of gaze from one object to the next because maintaining a very slow speech rate might require close attention to the phonetic and articulatory planning processes.

In the present study we did not observe systematic changes of the eye–speech lags with different speech rates, but we did see that the gaze-to-speech coordination was much tighter in Experiment 2 than in Experiment 1. This is evidenced clearly by the effects of name length on gaze durations and eye–speech lags found only in Experiment 2. The speakers of Experiment 2 were explicitly asked to pay attention to an aspect of the form of their utterances, the speech rate. Maintaining the prescribed speech rate was difficult, as shown by the fact that the Moderate Pace group consistently produced their utterances too fast, and that both groups increased their speech rate across the repetitions of the materials. As the names of the objects became available more and more quickly with increasing practice, it would probably have been easier for participants to increase their speech rate than to counteract the practice effects and maintain a constant speech rate. The systematic alignment of the shifts of eye gaze and attention with the completion of phonological planning may have been a way of supporting and monitoring the regular timing of the utterances. By contrast, the participants of Experiment 1 could produce the object names as soon as they had been planned, and no monitoring of speech rate was required. Since the production of the sequences of object names became less demanding with practice, attending to the objects until their names had been planned became less

beneficial, and therefore speakers often moved their eyes to the next object before they had retrieved the phonological form of the present object.

CONCLUSION

What do speakers do when they use different speech rates? Our study showed that eye gaze and speech output remained well coordinated across a range of speech rates: regardless of the speech rate, speakers look at most of the objects they name, and when they spoke faster the durations of the gazes to the objects decreased. This indicates that speakers spent less time planning each object name. When they were required to maintain a fixed speech rate, their eye gaze-to-speech coordination was very tight, with the shift of gaze being initiated after the name of each object had been phonologically encoded. This might be because maintaining a fixed speech rate is difficult, especially for well-practiced utterances, and requires careful monitoring of the speech output. When speakers aimed to speak as fast as they could, they initially still looked at, and attended to, each object until they had retrieved the phonological form of its name, but later moved their eyes earlier, perhaps as soon as they had identified the object. Together, the results suggest that looking at each object until its name has been planned to the level of the phonological form is a well established default strategy. Speakers use it because attending to the objects facilitates the recognition of the objects and the retrieval of their names. However, speakers can deviate from this strategy when they aim to monitor their speech very carefully, or when the utterances they produce are highly familiar and close monitoring is not required.

ACKNOWLEDGMENTS

This research was partly funded by ESRC grant R000239659 to Antje S. Meyer and Linda R. Wheeldon. It was carried out under the auspices of the University of Birmingham, UK.

REFERENCES

- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 390–412.
- Bachoud-Lévi, A. C., Dupoux, E., Cohen, L., and Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *J. Mem. Lang.* 39, 331–346.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Belke, E., and Meyer, A. S. (2007). Single and multiple object naming in healthy aging. *Lang. Cogn. Process.* 22, 1178–1210.
- Bock, K., Irwin, D. E., Davidson, D. J., and Levelt, W. J. M. (2003). Minding the clock. *J. Mem. Lang.* 48, 653–685.
- CELEX lexical database. (2001). Available at: <http://celex.mpi.nl>
- Cook, A. E., and Meyer, A. S. (2008). Capacity demands of word production: new evidence from dual-task experiments. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 886–899.
- Damian, M. F., Bowers, J. S., Stadthagen-Gonzalez, H., and Spalek, K. (2010). Does word length affect speech onset latencies in single word production? *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 892–905.
- Dell, G. S., Burger, L. K., and Svec, W. R. (1997). Language production and serial order: a functional analysis and a model. *Psychol. Rev.* 104, 123–147.
- Deubel, H., and Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.* 36, 1827–1837.
- Eimer, M., van Velzen, J., Gherri, E., and Press, C. (2007). ERP correlates of shared control mechanisms involved in saccade preparation and in covert attention. *Brain Res.* 1135, 145–166.
- Ferreira, V. S., and Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *J. Exp. Psychol. Learn. Mem. Cogn.* 28, 1187–1199.
- Gleitman, L. R., January, D., Nappa, R., and Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *J. Mem. Lang.* 57, 544–569.
- Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition.* 82, B1–B14.
- Griffin, Z. M. (2003). A reversed length effect in coordinating the preparation and articulation of words in speaking. *Psychon. Bull. Rev.* 10, 603–609.
- Griffin, Z. M. (2004). “Why look? Reasons for eye movements related to language production,” in *The Interface of Language, Vision, and Action: What We Can Learn from Free-Viewing Eye Tracking*, eds J. M. Henderson and F. Ferreira (New York: Psychology Press), 213–247.
- Griffin, Z. M., and Bock, K. (2000). What the eyes say about speaking. *Psychol. Sci.* 11, 274–279.
- Griffin, Z. M., and Oppenheimer, D. M. (2006). Speakers gaze at objects while preparing intentionally inaccurate names for them. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 943–948.
- Hartsuiker, R. J. (2006). Are speech error patterns affected by a monitoring bias? *Lang. Cogn. Process.* 21, 856–891.
- Hartsuiker, R. J., Pickering, M. J., and De Jong, N. (2005). Semantic and phonological context effects in speech error repair. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 921–932.
- Irwin, D. E. (2004). “Fixation location and fixation duration as indices of cognitive processes,” in *The Interface of Language, Vision, and Action: What We Can Learn from Free-Viewing Eye Tracking*, eds J. M. Henderson and F. Ferreira (New York: Psychology Press), 105–133.

- Jaeger, T. F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards mixed logit models. *J. Mem. Lang.* 59, 434–446.
- Jescheniak, J. D., and Levelt, W. J. M. (1994). Word frequency effects in speech production: retrieval of syntactic information and of phonological form. *J. Exp. Psychol. Learn. Mem. Cogn.* 20, 824–843.
- Korvorst, M., Roelofs, A., and Levelt, W. J. M. (2006). Incrementality in naming and reading complex numerals: evidence from eyetracking. *Q. J. Exp. Psychol.* 59, 296–311.
- Meyer, A. S., Belke, E., Häcker, C., and Mortensen, L. (2007). Regular and reversed word length effects in picture naming. *J. Mem. Lang.* 57, 210–231.
- Meyer, A. S., Roelofs, A., and Levelt, W. J. M. (2003). Word length effects in picture naming: the role of a response criterion. *J. Mem. Lang.* 47, 131–147.
- Meyer, A. S., Sleiderink, A. M., and Levelt, W. J. M. (1998). Viewing and naming objects: eye movements during noun phrase production. *Cognition* 66, B25–B33.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wegner, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks. A latent variable analyses. *Cogn. Psychol.* 41, 49–100.
- Mortensen, L., Meyer, A. S., and Humphreys, G. W. (2008). Speech planning during multiple-object naming: effects of aging. *Q. J. Exp. Psychol.* 61, 1217–1238.
- Oomen, C. C. E., and Postma, A. (2002). Limitations in processing resources and speech monitoring. *Lang. Cogn. Process.* 17, 163–184.
- Posner, M. I., and Petersen, S. E. (1990). The attention system of the human brain. *Annu. Rev. Neurosci.* 13, 25–42.
- Postma, A. (2000). Detection of errors during speech production: a review of speech monitoring models. *Cognition* 77, 97–131.
- Roelofs, A. (2007). Attention and gaze control in picture naming, word reading, and word categorization. *J. Mem. Lang.* 57, 232–251.
- Roelofs, A. (2008a). Attention, gaze shifting, and dual-task interference from phonological encoding in spoken word planning. *J. Exp. Psychol. Hum. Percept. Perform.* 34, 1580–1598.
- Roelofs, A. (2008b). Tracing attention and the activation flow in spoken word planning using eye movements. *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 353–367.
- Slevc, L. R., and Ferreira, V. S. (2006). Halting in single-word production: a test of the perceptual-loop theory of speech monitoring. *J. Mem. Lang.* 54, 515–540.
- Snodgrass, J. G., and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *J. Exp. Psychol. Hum. Learn. Mem.* 6, 174–215.
- Spieler, D. H., and Griffin, Z. M. (2006). The influence of age on the time course of word preparation in multiword utterances. *Lang. Cogn. Process.* 21, 291–321.
- Stadthagen-Gonzalez, H., Damian, M. F., Pérez, M. A., Bowers, J. S., and Marin, J. (2009). Name-picture verification as a control measure for object naming: a task analysis and norms for a large set of pictures. *Q. J. Exp. Psychol.* 62, 1581–1597.
- Wühr, P., and Frings, C. (2008). A case for inhibition. Visual attention suppresses the processing of irrelevant objects. *J. Exp. Psychol. Gen.* 137, 116–130.
- Wühr, P., and Waszak, F. (2003). Object-based attentional selection can modulate the Stroop effect. *Mem. Cognit.* 31, 983–994.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 29 July 2011; accepted: 02 February 2012; published online: 20 February 2012.

Citation: Meyer AS, Wheeldon L, van der Meulen F and Konopka A (2012) Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Front. Psychology* 3:39. doi: 10.3389/fpsyg.2012.00039

This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Meyer, Wheeldon, van der Meulen and Konopka. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.

APPENDIX**MONOSYLLABIC SETS**

Lamp coin rope bat straw pie; pin toe spoon leaf bow rat; owl mask
web corn sword brush; kite doll tap sock whale globe.

DISYLLABIC SETS

Lemon toilet spider pencil coffin basket; saddle bucket penguin
ladder whistle carrot; barrel wardrobe monkey statue rabbit garlic;
sausage dragon robot tortoise candle orange.