

Extrafoveal processing of objects in a naming task: Evidence from word probe experiments

JANE L. MORGAN

Sheffield Hallam University, Sheffield, England

GIJS VAN ELSWIJK

Radboud University Medical Centre, Nijmegen, The Netherlands

AND

ANTJE S. MEYER

University of Birmingham, Birmingham, England

In two experiments, we investigated the processing of extrafoveal objects in a double-object naming task. On most trials, participants named two objects; but on some trials, the objects were replaced shortly after trial onset by a written word probe, which participants had to name instead of the objects. In Experiment 1, the word was presented in the same location as the left object either 150 or 350 msec after trial onset and was either phonologically related or unrelated to that object name. Phonological facilitation was observed at the later but not at the earlier SOA. In Experiment 2, the word was either phonologically related or unrelated to the right object and was presented 150 msec after the speaker had begun to inspect that object. In contrast with Experiment 1, phonological facilitation was found at this early SOA, demonstrating that the speakers had begun to process the right object prior to fixation.

According to current models of speech production, speech is planned incrementally, meaning that speakers generate a plan for the first part of an utterance, produce it, and meanwhile plan the next part of the utterance (e.g., Levelt, 1989). Incrementality is also often taken to imply that different planning processes can be carried out in parallel; for instance, speakers might simultaneously plan the phonological representation of one utterance fragment and the content of the next fragment (Smith & Wheeldon, 1999).

Although the general principle of incrementality is broadly accepted, there is little empirical evidence about the way speakers coordinate different speech planning processes with each other and with the articulation of utterances. One way of studying this is to record the speakers' eye movements while they plan and produce descriptive utterances, such as, "The tree is next to the house" (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998; for a review, see Griffin, 2004). Such research has shown that speakers usually look at each event participant or object they refer to, in the order of mention, and that their gaze typically runs only slightly ahead of the overt speech.

Several studies have shown that the shift of gaze from one object to the next occurs only after the name of the first object has been planned to the level of the phonological form (Meyer, Belke, Häcker, & Mortensen, 2007; Meyer, Roelofs, & Levelt, 2003; Meyer et al., 1998; Meyer & van der Meulen, 2000; Roelofs, 2007). These results suggest that speakers plan descriptive utterances sequentially, with

little temporal overlap in the planning processes for successive objects (e.g., Levelt & Meyer, 2000). However, Morgan and Meyer (2005) demonstrated that speakers could identify objects and begin to retrieve their names before fixating upon them. In their experiments, speakers named triplets of objects in a prespecified order. During the saccade from the first to the second object, the second object seen at trial onset (the interloper) was replaced by a different object (the target), which the participants had to name. Morgan and Meyer found that the speakers spent less time looking at the target when it was identical to the interloper or had the same name (e.g., animal bat/baseball bat) than when it was unrelated to the interloper. This demonstrates that prior to fixation, the right object was not only recognized, but its name was activated as well. In other words, the onset of fixation on the right object did not mark the onset of processing of the object.

Morgan and Meyer (2005) considered two accounts of the origins of the interloper effect. One account is that the speakers processed the objects sequentially, but that the processing of the right object began before it was fixated upon. Eye movements are preceded by corresponding shifts of the focus of visual attention (e.g., Deubel & Schneider, 1996). The attentional shift occurs when the eye movement begins to be planned and involves closely related neural circuits (Awh, Armstrong, & Moore, 2006). It is possible that in Morgan and Meyer's object-naming task, the speakers first exclusively attended to the left object,

and then, when the eye movement was planned, shifted the focus of attention to the right object. The interloper effect could arise during the time period when the speakers' eye gaze was still directed at the left object but the focus of visual attention had already moved to the right object. Alternatively, the speakers' attended area may have included more than one object (e.g., Cave & Bichot, 1999; Goldsmith & Yeari, 2003): Initially, speakers focused on the left object but processed the right object as well, albeit with lower processing priority. Evidence supporting this view comes from a recent study by Meyer, Ouellet, and Häcker (in press). They used the same interloper–target pairs as Morgan and Meyer, but in addition varied the processing difficulty of the left object. This was done by presenting objects that in pretests had been shown to be either easy or difficult to identify and name. Meyer et al. (in press) found that the interloper effects were significantly smaller when the left object was difficult than when it was easy to process. This interaction suggests that the left and right objects were processed in parallel and competed for processing resources.

These findings are important for theories of speech planning because they imply that speakers can divide their visual attention over two or more objects and process them in parallel, and that the low-priority processing of an extrafoveal object can suffice for the name of the object to become activated. These findings are also important for methodological reasons because they demonstrate that the onset of fixation on an object need not correspond to the onset of processing.

The studies by Morgan and Meyer (2005) and by Meyer et al. (in press) used the same small set of interloper–target pairs. In additional unpublished studies, Meyer and colleagues used phonologically related interloper–target pairs (e.g., bed–bell) instead of homophonous pairs. Surprisingly, these studies failed to yield any facilitatory effects of related interlopers. The reasons why the effect was confined to homophonous interlopers are not clear; this may be related to the time course of the activation of the morphological and phonological representations of the picture names or to properties of the transsaccadic memory representations speakers generate (e.g., Polatsek, Rayner, & Collins, 1984).

The goal of the present study was to assess the claim that speakers can process objects they are about to name before fixating upon them. For this study, we chose to use a different paradigm and new materials. We developed a novel paradigm that combined properties of the probe paradigm used by Peterson and Savoy (1998; see also Levelt et al., 1991; Slevc & Ferreira, 2006) with the eye-movement-contingent probe paradigm introduced by Kambe, Duffy, Clifton, and Rayner (2003; see also Rayner, 1975). Participants saw pairs of objects, which they prepared to name. On 75% of the trials, which were filler trials, participants named the objects, starting with the left object on the screen. On the remaining trials, the objects were replaced shortly after trial onset by a single written word probe, which the participants had to name instead of the objects (see Figure 1). The probe was phonologically related or unrelated to the name of one of the objects.

The aim of Experiment 1 was to obtain an estimate of when the phonological form of the name of the left object would be activated. The probes appeared on the left side of the screen (see Figure 1) either 150 or 350 msec after picture onset. The results of earlier studies showed that speakers needed between 250 and 330 msec to retrieve the phonological form of an object name (e.g., Indefrey & Levelt, 2004; Roelofs, 2007). On the basis of these findings, we predicted that 350 msec after picture onset, the phonological form of the name of the left object should be sufficiently activated to prime a related probe. Therefore, related probes should be named faster than unrelated ones. However, we predicted no priming for probes presented 150 msec after trial onset because the phonological form of the picture name should not be activated sufficiently to prime the probe so soon after picture onset.

The aim of Experiment 2 was to determine whether the name of the right object would be activated at the same time—in relation to the onset of fixation on the object—as the name of the left object or earlier. The probes were

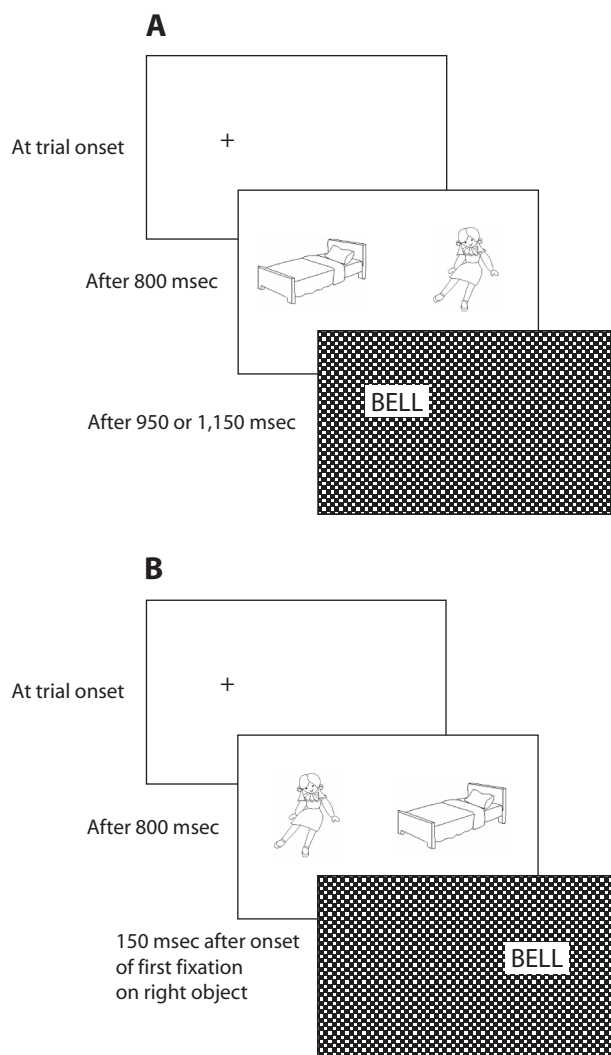


Figure 1. Schematic representation of the trial structure in Experiment 1 (A) and Experiment 2 (B).

presented on the right side of the screen and were related or unrelated to the name of the right object. They appeared 150 msec after the speaker had begun to inspect the right object. If speakers only begin to process the right object after fixation, the results should mirror those expected for the left object, and no priming should be observed for probes presented 150 msec after fixation onset. By contrast, if speakers begin to process the right object prior to fixation, the phonological representation of its name may already be sufficiently activated to prime the processing of the probe.

EXPERIMENT 1

Method

Participants. The experiment was conducted with 40 students from the University of Birmingham. All were native speakers of British English and had normal or corrected-to-normal vision. They received course credits or payment in return for participation.

Materials. Forty pairs of line drawings of common objects with semantically and phonologically unrelated names were selected from a picture database available in the third author's lab. Eight pairs were used on practice trials, 16 on filler trials, and 16 on probe trials. The pictures were fitted into virtual frames that covered a visual angle of $5.5^\circ \times 5.5^\circ$ when viewed from the participant's position and were shown on the horizontal midline of the screen, with a midpoint-to-midpoint distance of 15.4° .

On the probe trials, the left object was combined with a phonologically related or unrelated probe word. The words ranged from three to nine letters in length. The related probes shared from one to three word-initial phonemes with the object name, whereas the unrelated probes shared no phonemes in the same word position with the object name (see the Appendix). Phonologically related and unrelated probes were matched for frequency of occurrence using COBUILD (2000; mean log frequency/million = 1.75; $SD = 0.53$). Four additional words were selected to be used on practice trials. The words were printed in lowercase, 36-point Times New Roman font in a window that covered a visual angle of $1.8^\circ \times 6.3^\circ$. The window was superimposed on a pattern mask that covered the entire screen.

Design. The experiment included eight test blocks containing 32 trials each. In each block, four experimental items were combined with a related probe, four with an unrelated probe, and eight without a probe. Across the eight blocks, each experimental item appeared twice in each of these combinations. In addition, the 16 fillers were shown in each block. Each participant saw the items within the blocks in a different random order, with the constraint that at least 1 filler trial intervened between any 2 probe trials. The order of the blocks was counterbalanced across participants. The first block was preceded by 8 practice trials (4 probe trials and 4 filler trials), and each of the following blocks was preceded by 3 practice trials.

The word probes appeared either 150 or 350 msec after the participant began to inspect the left object. Since the participants were usually looking at the left side of the screen when the object appeared, this means that the word typically appeared about 150 or 350 msec after picture onset. If the participant did not look at the target object within 2,500 msec after picture onset, the word probe was not presented and no naming latency was recorded. The timing of the word probes (stimulus onset asynchrony, SOA) was varied between participants, with 16 participants being tested at the short SOA and 24 at the longer SOA.

Apparatus. We used the experimental software package NESU (Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands). The stimuli were presented on a 19-in. Samtron 95P Plus color monitor. The participants' speech was recorded using a Sony ECM-959DT microphone and a Sony DAT recorder. On probe trials, the participants had to interrupt their picture-naming response and name the probe instead. To determine the probe-naming latencies, the responses were digitized, and the onset of the word-naming response was measured using the speech analysis system PRAAT. Eye

movements were recorded using an SMI Eyelink I head-mounted eyetracking system, which has a spatial accuracy of less than 0.5° . The position of both eyes was estimated every 4 msec.

Procedure. Participants were instructed that they would see object pairs that they should name using bare nouns, starting with the left object. They were also told that sometimes the object pairs would only be shown briefly before being replaced by a single word. On these trials, they should not name the objects, but should read the word aloud as quickly and accurately as possible.

The participants were then shown a booklet containing pictures of all the objects used in the experiment, with the objects' names printed beneath. A naming session followed, during which the experimenter corrected any erroneous responses. The eyetracking system was then set up and calibrated, and the first experimental block began.

Each trial began with the presentation of a fixation point for 800 msec in the center of the region where the left object would appear. Then a two-object display was shown. On filler trials, the display remained in view for 2,500 msec, and the participants named both objects. On probe trials, the objects were removed 150 or 350 msec after the onset of the first fixation on the left object and were replaced by a single-word probe, which the participant had to name. Word probes remained on the screen for 2,500 msec.

Results and Discussion

On 10% of the filler trials, participants named one or both objects incorrectly, corrected themselves, or committed speech errors. The mean speech onset latency for the remaining filler trials was 809 msec at SOA 150 and 789 msec at SOA 350. On virtually all of these trials, the participants first looked at the left object (for, on average, 582 msec), and then at the right object. The error rate, naming latency, and gaze duration were similar to those obtained in other experiments requiring the naming of two or more objects (e.g., Meyer et al., 1998; Meyer & van der Meulen, 2000), which indicates that the participants prepared their utterances in a similar fashion, despite knowing that they might have to switch to a word-naming task.

Three percent of the probe trials were excluded from the analyses because participants had made speech errors or because their latencies exceeded 1,200 msec. Two percent of these trials occurred in the related condition and 1% in the unrelated condition. A further 3% of trials were lost because participants failed to look at the left object within 10 msec of the probe onset or because of technical difficulties.

At SOA 150, the naming latencies for related and unrelated probes were virtually identical (see Table 1). However, at SOA 350, the mean naming latency was shorter by 16 msec in the related than in the unrelated condition. The interaction of relatedness and SOA was significant in the subject analysis [$F_1(1,38) = 4.47, p < .05$] and approached significance in the item analysis [$F_2(1,15) = 3.28, p < .10$]. Planned comparisons showed that the effect at SOA 350 was significant [$t_1(23) = 2.86, p < .01$; $t_2(15) = 2.25, p < .05$], whereas the 1-msec difference at SOA 150 was not.

In sum, Experiment 1 yielded the expected pattern of results: Phonologically related probes were named faster than unrelated ones when they were presented 350 msec after the participants had begun to process the relevant object, but not when they were presented 200 msec earlier. This reflects the notion that the phonological form of an object name takes time to become activated (e.g., Indefrey & Levelt, 2004; Schriefers, Meyer, & Levelt, 1990).

Table 1
Mean Probe-Naming Latencies (in Milliseconds)
and Priming Effects

SOA	Related Probes		Unrelated Probes		Priming Effect
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	
Experiment 1					
150	641	22	640	20	−1
350	620	20	636	20	16
Experiment 2					
150*	554	22	602	19	48

Note—SOA, stimulus onset asynchrony, in milliseconds. *SOA measured from onset of fixation on right object.

EXPERIMENT 2

In Experiment 2, the participants again named object pairs, starting with the left object on the screen. However, the probes were now presented on the right side of the screen 150 msec after the participants had begun inspecting the right object, and were either phonologically related or unrelated to that object. Were speakers to begin processing the right object prior to fixation, phonological facilitation could be observed at this early SOA.

Method

Participants. The experiment was conducted with 16 participants recruited from the same pool as were the participants in Experiment 1.

Materials and Design. The same objects and words were used as in Experiment 1. However, the left and right object of each experimental pair were exchanged, and the probes were now centered around the midpoint of the right object and appeared 150 msec after the onset of the first fixation to that object. As in Experiment 1, participants were instructed to name the left object first, and then the right object.

Results and Discussion

On 10% of the filler trials, participants produced speech errors or did not look at the objects in the expected order. The mean speech-onset latency for the remaining filler trials was 796 msec. Six percent of the probe trials were excluded from the analyses due to subject error (3% in the related condition and 3% in the unrelated condition). A further 7% of trials were omitted due to technical problems. The saccade from the first to the second object ended, on average, 607 msec after trial onset, and the probe was presented 150 msec later. The mean probe-naming latency was significantly shorter (by 48 msec) in the related than in the unrelated condition [$t_1(15) = 3.51$, $p < .01$; $t_2(15) = 3.66$, $p < .01$; see Table 1], which indicates that the phonological form of the object name was sufficiently activated to prime the naming of the probe.

GENERAL DISCUSSION

In Experiment 1, the probes were related or unrelated to the name of the left object. Phonological facilitation was seen when the probes were presented 350 msec after the speaker had begun to inspect the object, but not when they were presented 200 msec earlier. By contrast, in Experiment 2, where the probes were related or unrelated to the name of the right object, phonological facilitation

was observed when the probes were presented 150 msec after the right object was first fixated. An ANOVA that compared naming latencies at SOA 150 in Experiments 1 and 2 yielded a significant interaction between relatedness and experiment [$F_1(1,30) = 11.38$, $F_2(1,15) = 11.50$, both $ps < .01$]. Taken together, the results of the two experiments demonstrate that relative to the onset of gaze to the object, the name of the right object was available earlier than was the name of the left object. We conclude that the right object began to be processed before it was fixated, which was not possible for the left object.

Our results are consistent with findings by Morgan and Meyer (2005) and Meyer et al. (in press), who also suggested that speakers can begin to process objects they are about to name prior to fixation. As explained earlier, it was important to obtain converging evidence for this conclusion because these two studies used the same paradigm and the same small set of materials, and a third study had yielded conflicting evidence.

Meyer et al. (in press) proposed that speakers planning to name two objects processed them in parallel but initially prioritized the object to be named first, such that its name became available before the name of the second object. Experiment 1 of the present study showed that 350 msec after the onset of the display, the phonological form of the name of the left object was sufficiently activated to prime the processing of the probe. However, an additional experiment showed that at the same moment in time, the phonological form of the name of the right object was not yet sufficiently activated to prime the processing of a related probe. This pattern supports the view that speakers had prioritized the first object, rather than processing the first and second objects in exactly the same way.

As mentioned in the Introduction, it is generally assumed that speakers plan utterances incrementally. The present results, together with the findings of our earlier studies, indicate that incrementality should not be equated with strict sequentiality. When speakers name several objects, they direct the focus of their visual attention to each of the objects in the order of mention; but while they are focusing on one object, the next object may already be processed “in the background.” This conclusion fits well with theories of speech planning that propose that speakers generate successive utterance fragments in parallel, always giving processing priority to one fragment, but processing the next fragment as well (e.g., Dell, Burger, & Svec, 1997). An important question for further research is whether speakers generally distribute their visual attention across several objects or whether they adopt a more sequential strategy when, for instance, the individual objects are difficult to identify or name.

Finally, the results presented here are also important for methodological reasons: On the positive side, they show that eye tracking can be used in speech-planning tasks to control the presentation of stimuli. Eye movements can be used to indicate when a shift of visual attention from one region of a display to another has occurred, and display changes can be linked to this event (see also Kambe et al., 2003). On the more negative side, however, the results show that the onset of gaze to an object need not correspond to the onset of processing. This needs to be taken into account

when variables such as fixation duration or gaze duration are used to estimate the time required to process objects.

AUTHOR NOTE

This research was supported by Economic and Social Research Council Grant R000239391, awarded to A.S.M. Correspondence concerning this article should be addressed to J. L. Morgan, Psychology Group, Faculty of Development and Society, Sheffield Hallam University, Sheffield S10 2BP, England (e-mail: j.l.morgan@shu.ac.uk).

REFERENCES

- AWH, E., ARMSTRONG, K. M., & MOORE, T. (2006). Visual and oculomotor selection: Links, causes and implications for spatial attention. *Trends in Cognitive Sciences*, **10**, 124-130.
- CAVE, K. R., & BICHOT, N. P. (1999). Visuospatial attention: Beyond a spotlight model. *Psychonomic Bulletin & Review*, **6**, 204-223.
- COBUILD Corpus of English Sentences [CD]. (2000). Birmingham, U.K.: University of Birmingham.
- DELL, G. S., BURGER, L. K., & SVEC, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, **104**, 123-147.
- DEUBEL, H., & SCHNEIDER, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, **36**, 1827-1837.
- GOLDSMITH, M., & YEARI, M. (2003). Modulation of object-based attention by spatial focus under endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 897-918.
- GRIFFIN, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 213-247). New York: Psychology Press.
- GRIFFIN, Z. M., & BOCK, K. (2000). What the eyes say about speaking. *Psychological Science*, **11**, 274-279.
- INDEFREY, P., & LEVELT, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, **92**, 101-144.
- KAMBE, G., DUFFY, S. A., CLIFTON, C., JR., & RAYNER, K. (2003). An eye-movement-contingent probe paradigm. *Psychonomic Bulletin & Review*, **10**, 661-666.
- LEVELT, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- LEVELT, W. J. M., & MEYER, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, **12**, 433-452.
- LEVELT, W. J. M., SCHRIEFERS, H., VORBERG, D., MEYER, A. S., PECHMANN, T., & HAVINGA, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, **98**, 122-142.
- MEYER, A. S., BELKE, E., HÄCKER, C., & MORTENSEN, L. (2007). Use of word length information in utterance planning. *Journal of Memory & Language*, **57**, 210-231.
- MEYER, A. S., OUELLET, M., & HÄCKER, C. (in press). Parallel processing of objects in a naming task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*.
- MEYER, A. S., ROELOFS, A., & LEVELT, W. J. M. (2003). Word length effects in picture naming: The role of a response criterion. *Journal of Memory & Language*, **47**, 131-147.
- MEYER, A. S., SLEIDERINK, A. M., & LEVELT, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, **66**, B25-B33.
- MEYER, A. S., & VAN DER MEULEN, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review*, **7**, 314-319.
- MORGAN, J. L., & MEYER, A. S. (2005). Processing of extrafoveal objects during multiple-object naming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 428-442.
- PETERSON, R. R., & SAVOY, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 539-557.
- POLLATSEK, A., RAYNER, K., & COLLINS, W. E. (1984). Integrating pictorial information across eye movements. *Journal of Experimental Psychology: General*, **113**, 426-442.
- RAYNER, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, **7**, 65-81.
- ROELOFS, A. (2007). Attention and gaze control in picture naming, word reading, and word categorizing. *Journal of Memory & Language*, **57**, 232-251.
- SCHRIEFERS, H., MEYER, A. S., & LEVELT, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory & Language*, **29**, 86-102.
- SLEVC, L. R., & FERREIRA, V. S. (2006). Halting in single word production: A test of the perceptual loop theory of speech monitoring. *Journal of Memory & Language*, **54**, 515-540.
- SMITH, M., & WHEELDON, L. R. (1999). High level processing scope in spoken sentence production. *Cognition*, **73**, 205-246.

APPENDIX

Object and Word Combinations Presented on the Probe Trials

Object Name		Word Probe	
Target	Flanker	Phonologically Related	Unrelated
apple	foot	appetite	dock
boot	candle	boost	penny
bread	tie	breath	idea
eye	ashtray	idea	kennel
dog	scissors	dock	appetite
kettle	ruler	kennel	soft
pencil	squirrel	penny	boost
sock	flower	soft	breath
shoe	train	shoot	art
carrot	belt	carriage	bottom
arm	pepper	art	butterfly
tree	spoon	treat	cloth
button	violin	butterfly	carriage
bottle	stool	bottom	keen
key	tomato	keen	shoot
clock	fish	cloth	treat