# The role of vision and attention in language processing

Antje S. Meyer

University of Birmingham
and Max Planck Institute for
Psycholinguistics Nijmegen

UNIVERSITY OF BIRMINGHAM

# Introduction

- Psycholinguists (quite sensibly) tend to focus on language processing.

- Exception (also sensible): Reading research
  - Implemented models, e.g. SWIFT (Kliegl et al., 2006), E-Z Reader (Reichle et al., 2003)

# Language is situated

- We see things when we listen and speak.
- We often try to ignore the visual information. This requires cognitive control.
- We often use the visual information. This requires cognitive control, capacity sharing and cross-modal integration.

- Historical background: Modularity

# Plan of the talk

- The visual world paradigm of speech comprehension
- The multiple object naming paradigm of production
- Why do the paradigms work?

# Cooper, 1974

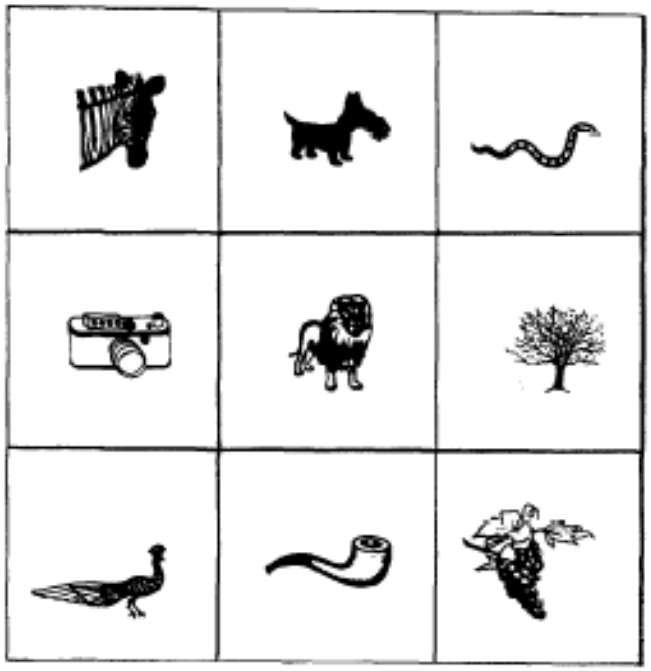
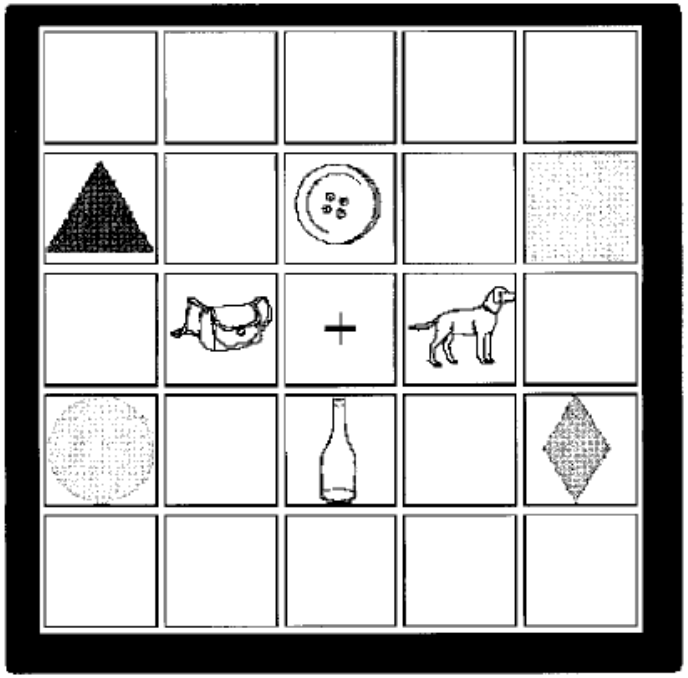


FIG. 1. The slide accompanying Story 3.

“ a practical new research tool for the real-time investigation of perceptual and cognitive processes and, in particular, for the detailed study of speech perception, memory, and language processing" (p. 84).

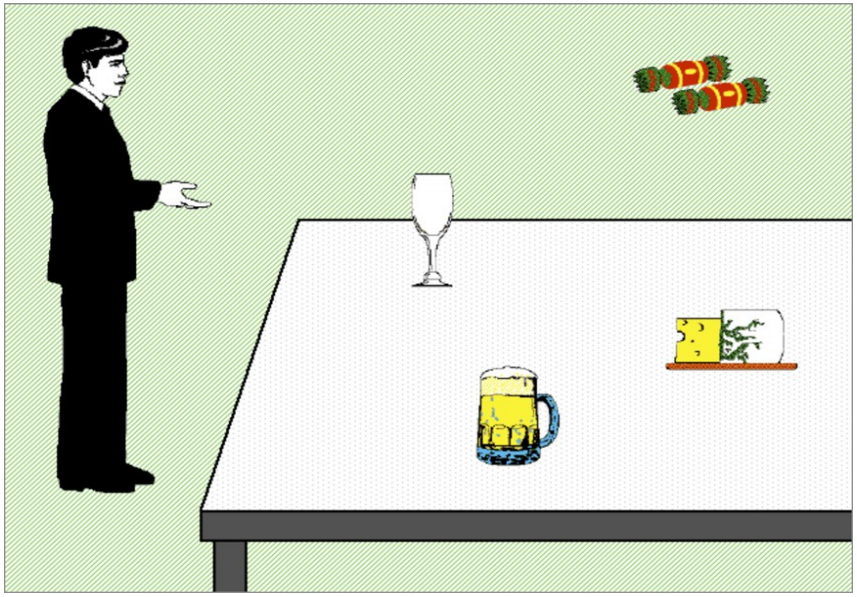
# Modern versions of the visual world paradigm

- Acting out “Look at the dog” (e.g., Dahan & Gaskell, 2007)



# Modern versions of the visual world paradigm

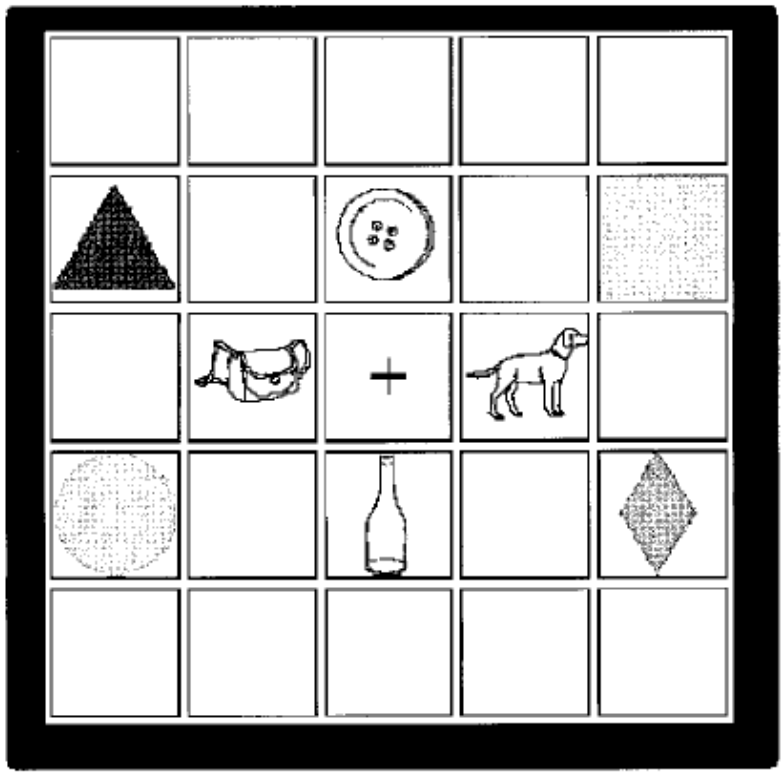
- Passive listening (e.g. Altmann & Kamide, 2007)



“The man will drink the wine”

# SR Eyelink II

# Analysis of fixation proportions



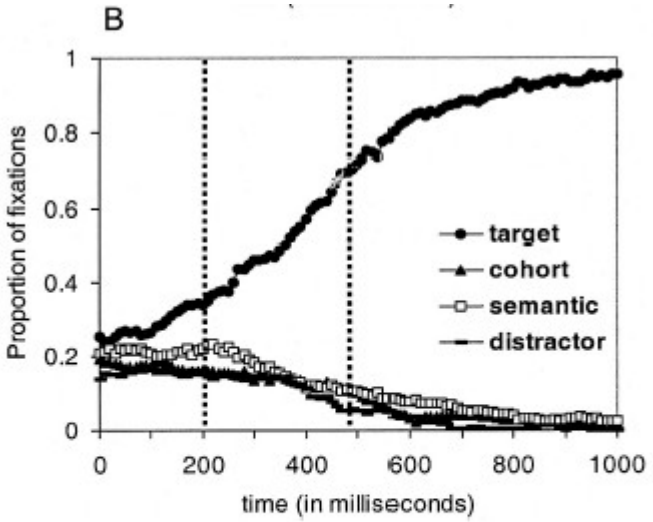



*Figure 1.* Proportion of fixations to the target, the cohort competitor, the semantic competitor, and the distractor over time from the onset of the target word (in milliseconds) in the neutral-verb condition (A) and in the constraining-verb condition (B). The dotted lines indicate the mean time interval over which fixations reflect the processing of the target word (with a 200-ms delay for saccadic latency).

(Dahan & Tanenhaus, 2004)

# Applications of visual world paradigm

- Processing of low-level phonetic cues, e.g., Salverda et al., 2002
  - Utterance “Look at the hamster” triggers earlier looks to target when there is a hamster + three unrelated objects than when there is a hamster in the display as well.
  - Hamster is looked at earlier when first syllable is excised from “hamster” than from “hammer”.

# Applications of visual world paradigm

- Syntactic analysis, e.g. Tanenhaus, 1995
  - Participants acted out “Put the apple on the towel in the box”.
  - Display includes 1 or 2 apples (one on towel)
  - 2-apple display leads to earlier looks to box (= faster understanding) than 1-apple display.
  - In two-apple condition, listeners immediately understand that “on the towel” specifies which apple to move.

# Applications of visual world paradigm

- Predictive listening, e.g. Altmann & Kamide, 2009

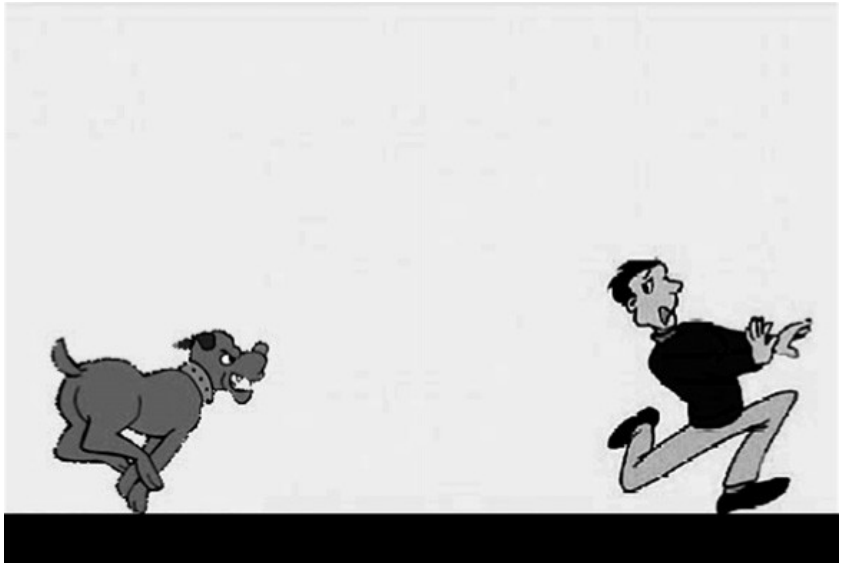
  “The boy will eat the cake” leads to earlier fixations on cake than “The boy will move the cake”. This shows rapid use of semantic knowledge.

  “The girl/man will ride .. “ lead to early looks to age-appropriate vehicles. This shows rapid use of world knowledge.

# Main results of visual world research

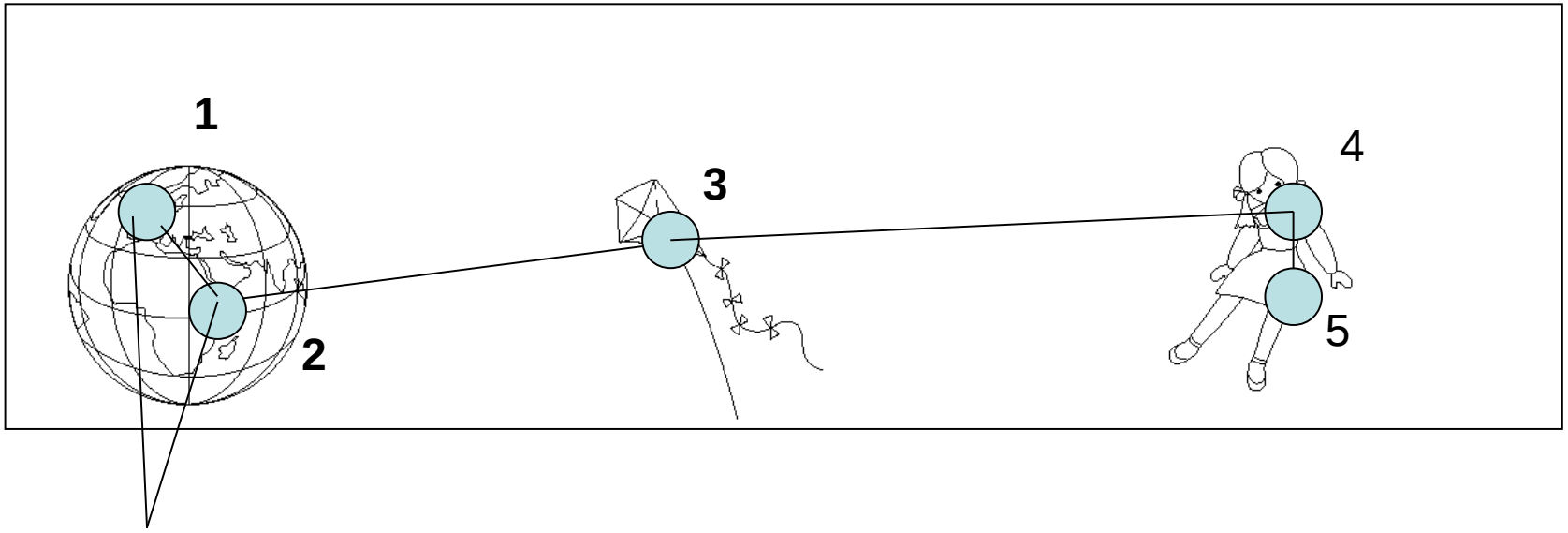
- Listeners use various types of information to understand spoken utterances.
- Listeners use the information as soon as possible.
- The listeners' use of information conveyed by speech and by visual information and their use of semantic and world knowledge is reflected in their eye gaze.

# Use of eye tracking in speech production research



- Description of scenes and events (Bock & Griffin, 1998;Gleitman et al., 2007)
- Can we distinguish between pre-linguistic scene apprehension and linguistic formulation?

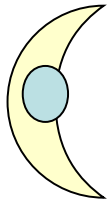
# Multiple-object naming

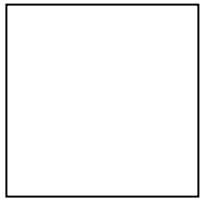


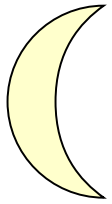

Gaze

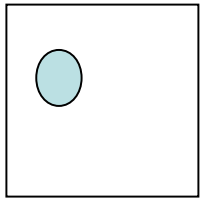
"Globe, kite, doll."

Speakers look at each object they mention.
They look at the objects in the order of mention









The moon…..

The moon next to the square…..

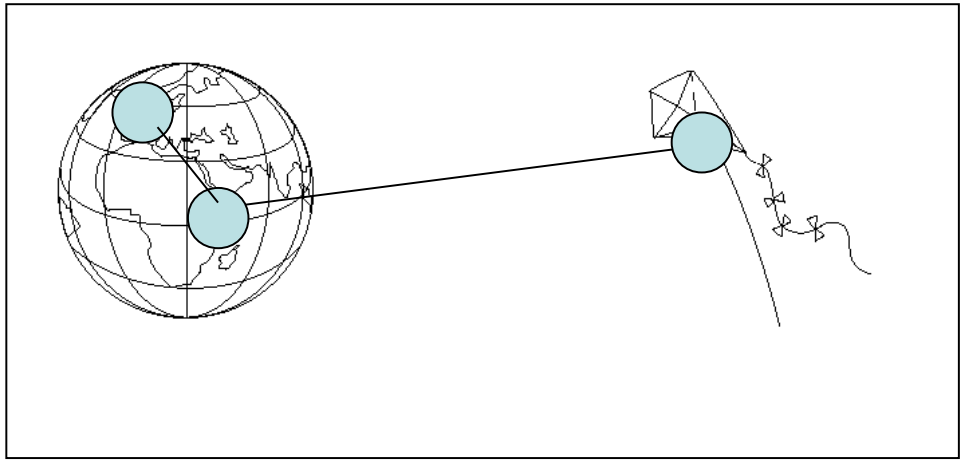
The moon next to the square is pale.

(Meyer, van der Meulen, & Brooks, 2004)

# Three core findings

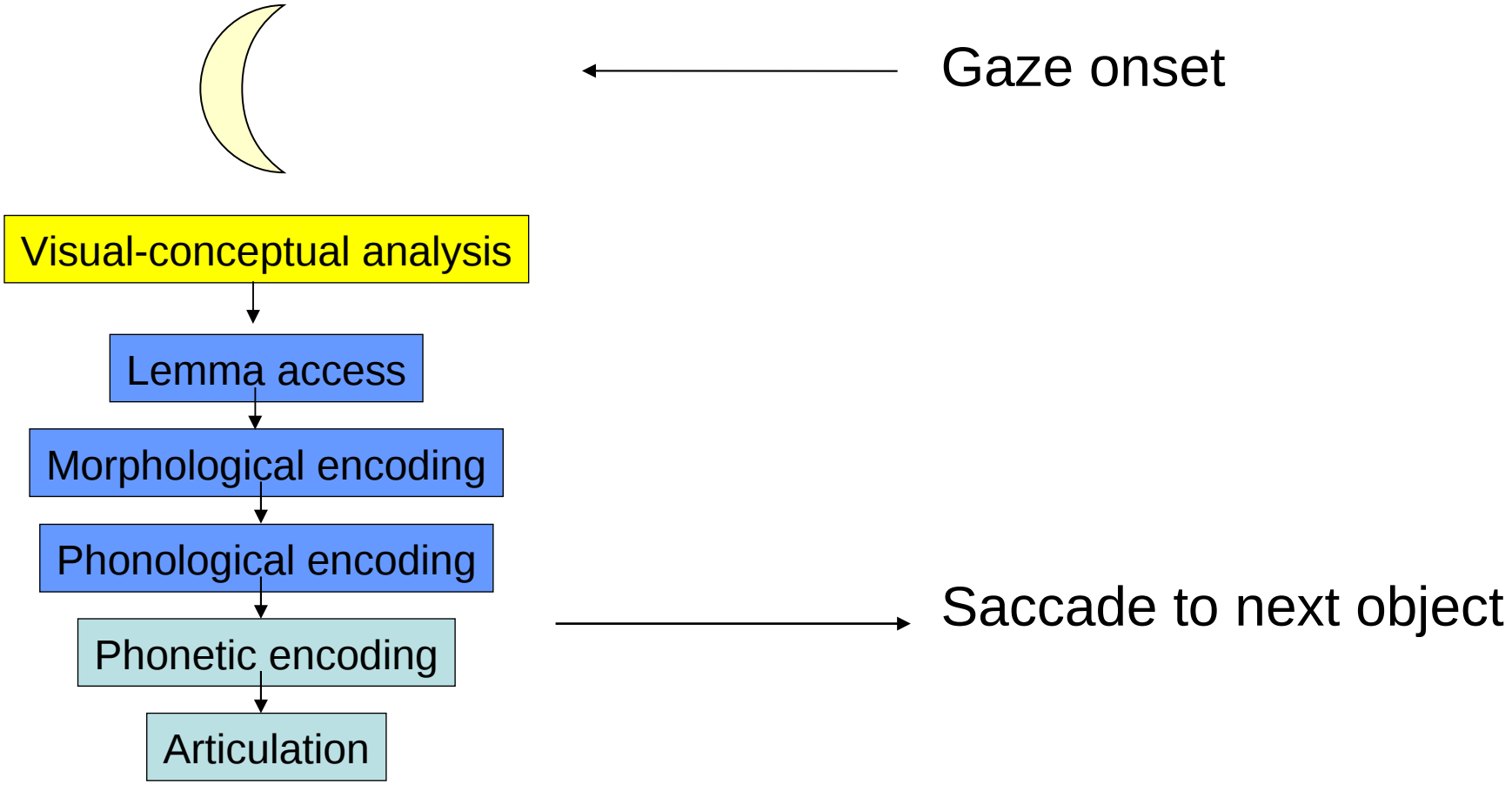
1. Speakers look at each object they mention.
2. They look at the objects in the order of mention.
3. They look at each object until they have generated the phonological representation of the corresponding expression.

# Determinants of gaze durations in object naming



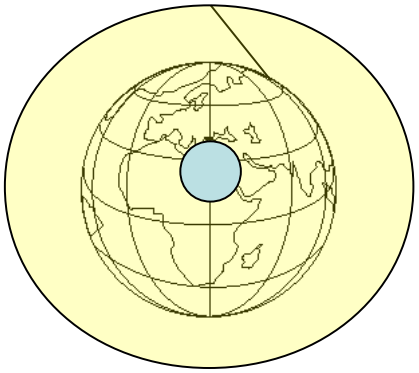
- Visual complexity (Meyer et al., 1998)
- Ease of lexical selection (Belke et al., 2005)
- Name frequency & length (Korvorst et al., 2006; Meyer et al., 2003, 2005)

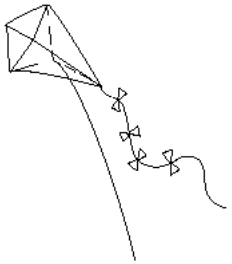
# Coordination of word planning an eye gaze



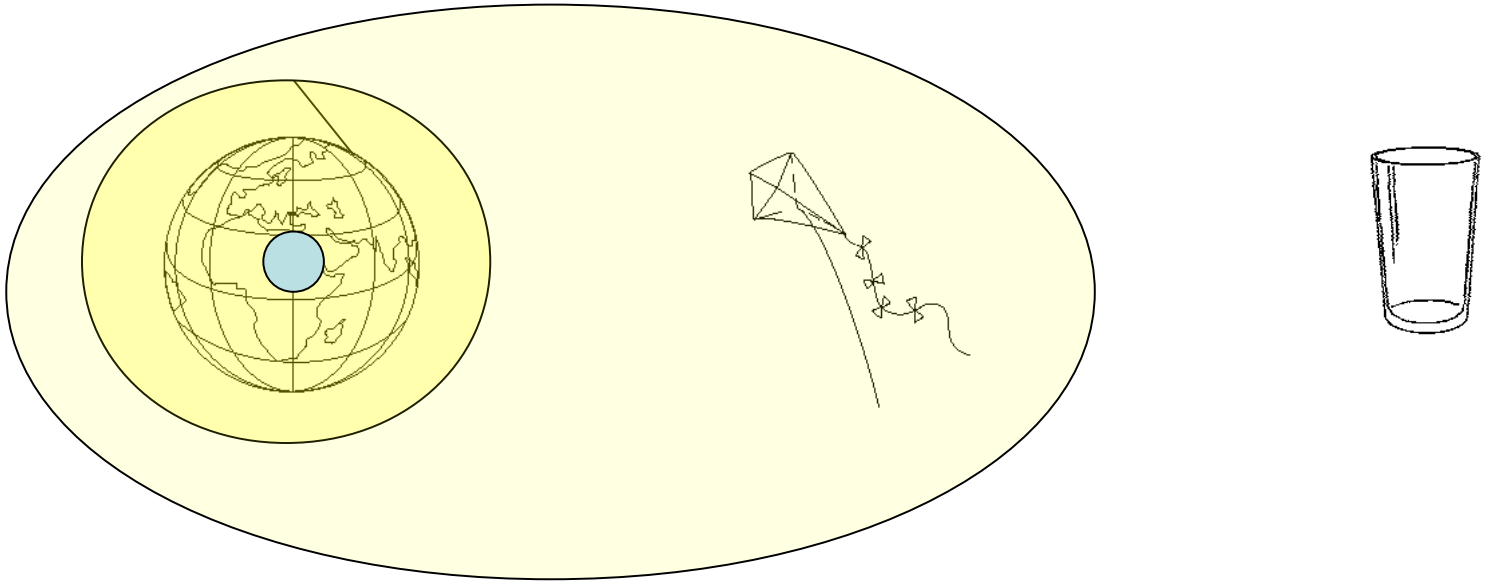

(Levelt, Roelofs, & Meyer, 1999)

# The Serial Processing Hypothesis



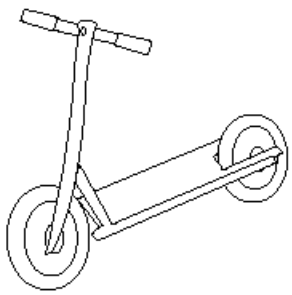


# The Gradient Hypothesis



(e.g., Cave & Bichot, 1999)

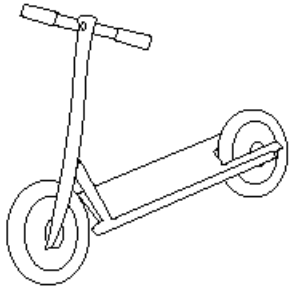
# Extrafoveal object processing



Interloper

(Morgan & Meyer, 2005)

# Extrafoveal object processing



Target

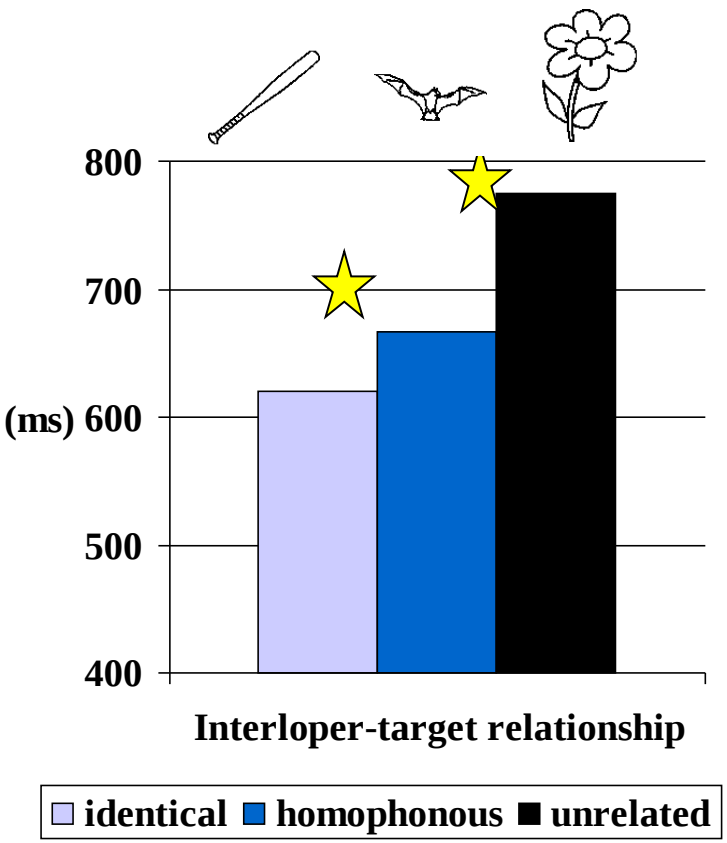
# Interloper-target relationship

| | interloper | target |
|---|---|---|
| Identical (bat-bat) | | |
| Homophonous (bat-bat) | | |
| Unrelated (flower-bat) | | |

Dependent measure:

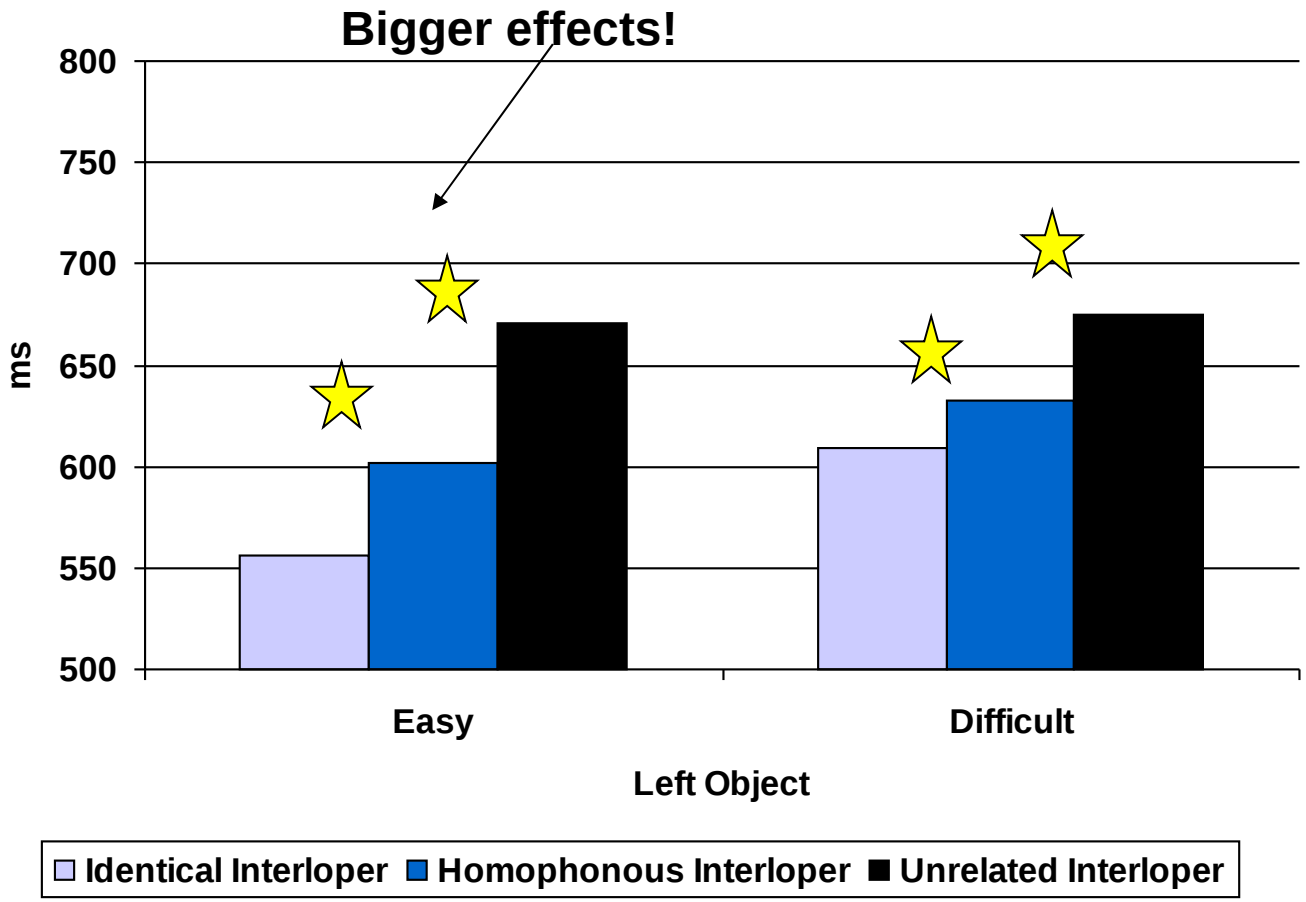
Gaze duration for target

# Target gaze durations



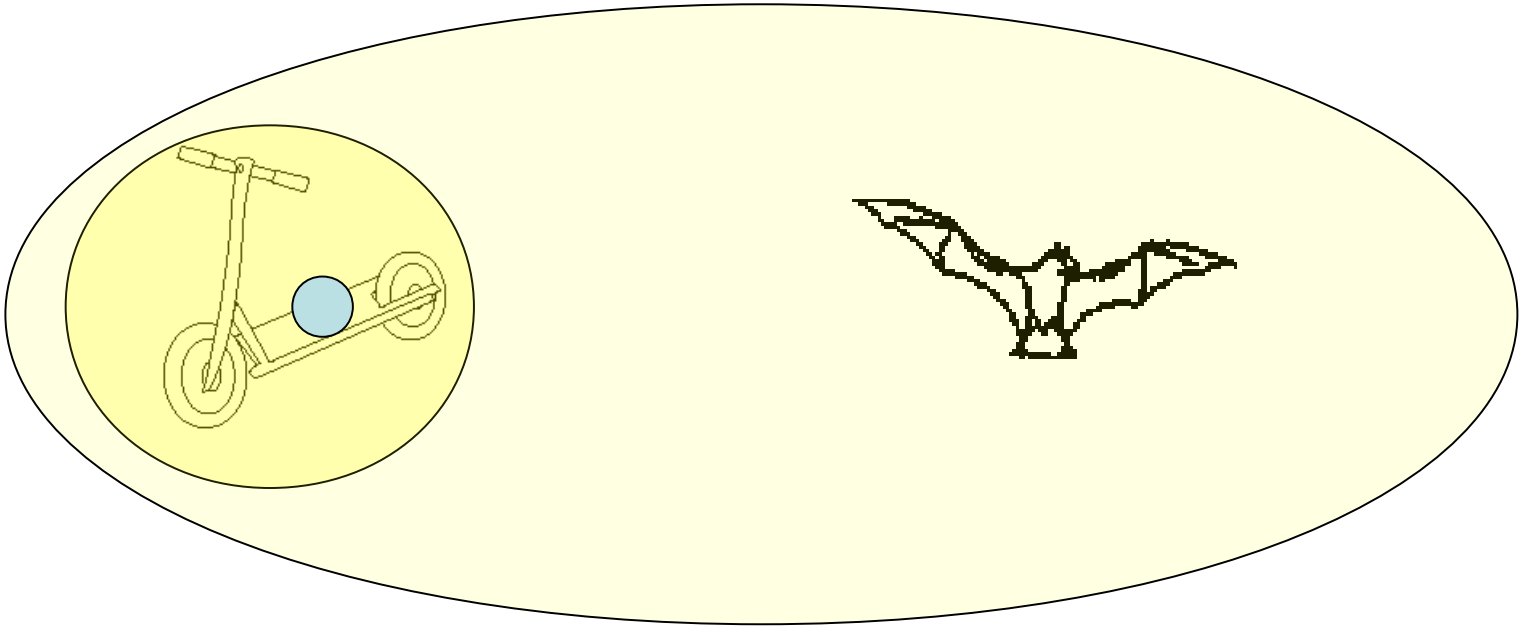

- Interlopers affected the target gaze durations.
- Speakers began to process interlopers prior to fixation.
- Interlopers were processed to the level of word form.

# Effects of left-object difficulty




(Meyer, Ouellet, & Häcker, 2008; see also Meyer, subm.,a,b)

# Gradient Hypothesis



# Conditions for parallel object processing

- The foveated object must be easy to process (Malpass & Meyer, 2010).
- Objects must be close together (Konopka & Meyer, subm.).
- Syntactic structure must be constant (Meyer, subm., a, b).

# Summary of production studies

- Speakers look at the objects they name.
- They fixate upon the objects until they have retrieved a suitable referring expression.
- They sometimes process two objects in parallel.
- Object recognition and speech planning processes are reflected in the speakers' eye gaze.

# Linking eye gaze and speech planning

- Eye gaze reflects on visual attention (e.g. Irwin, 2004).
- Attended objects are processed with priority (fastest, most accurately).
- The easiest way of attending to an objects it to look at it.

# Why do speakers look at and attend to the objects they name?

- Fixation is necessary for identification.
- Retrieval of information from the visual environment is prefered over memory retrieval.
- Visual attention facilitates object recognition and the retrieval of associated information, including the name (e.g. Remington & Folk, 2001;Malpass & Meyer, 2010).

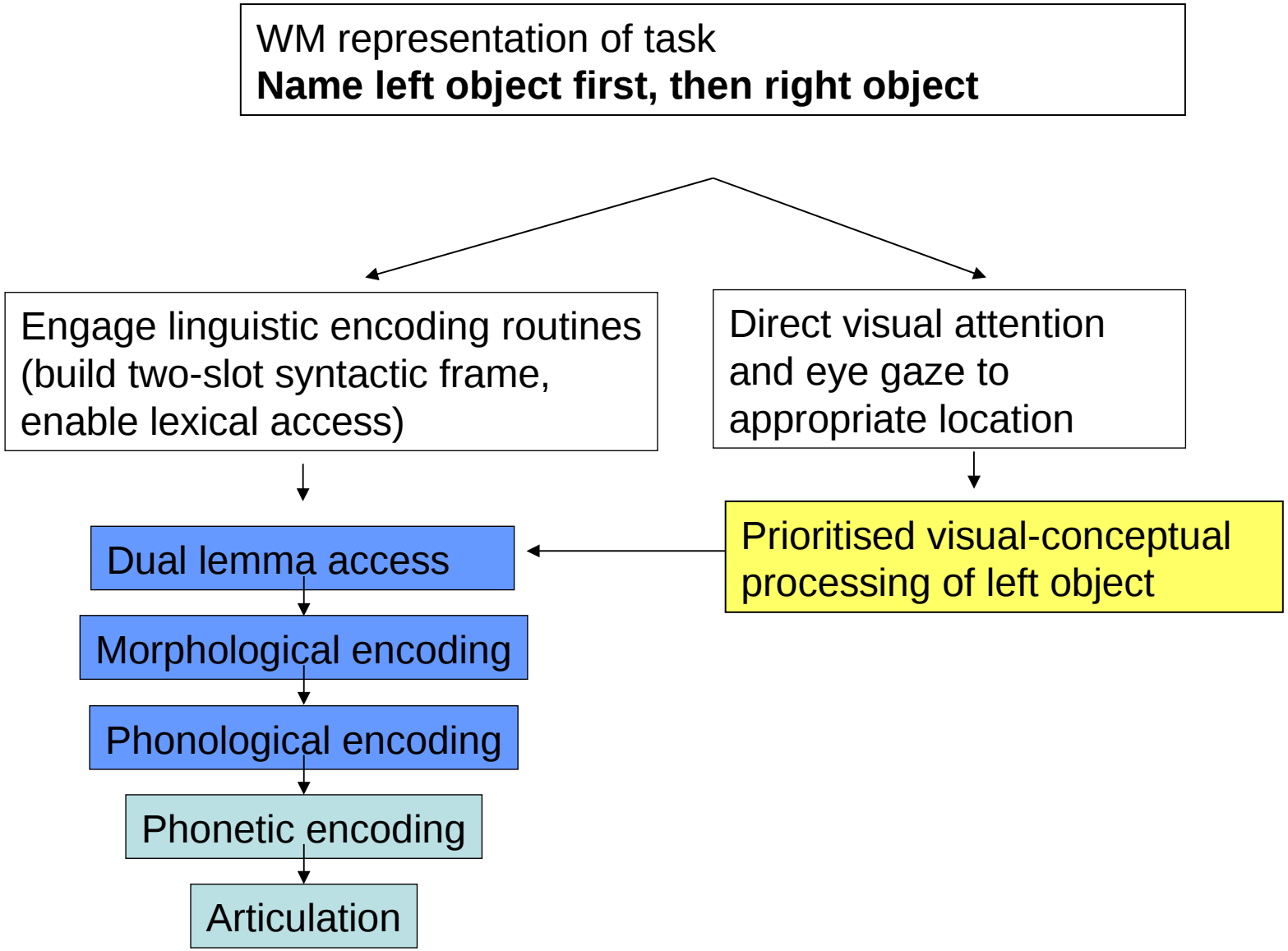
# Why do speakers attend to the objects they name?

- Lexical access requires processing capacity (e.g. Ferreira & Pashler, 2002; Cook & Meyer, 2007).
- Speech planning is constrained by the speed of articulation (Wheeldon & Meyer, subm., Frisson & Meyer, in prep).
  - Eye movements in reading aloud and picture naming are very similar.

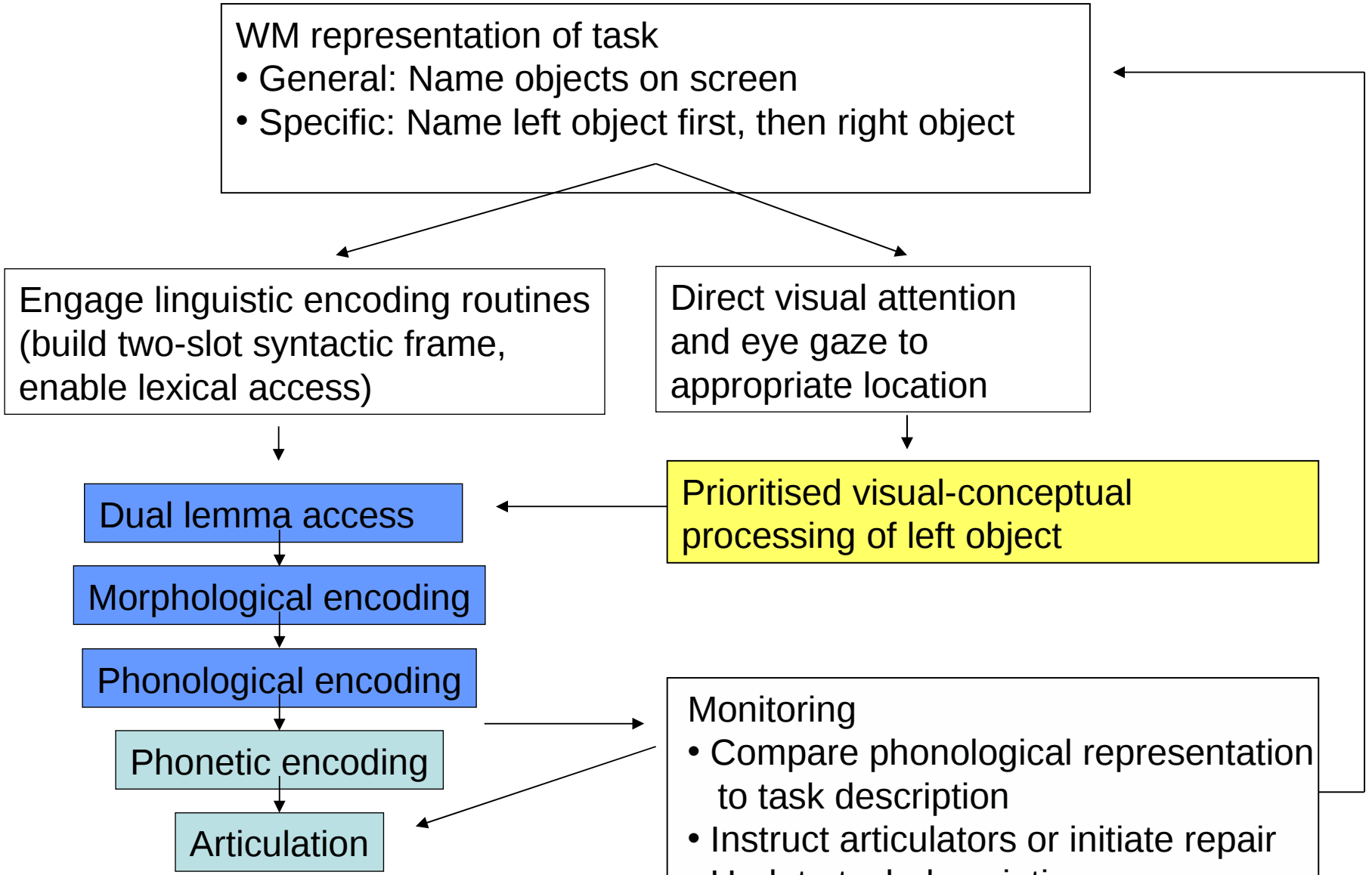
# Eye gaze reflects speech plan

- Speakers process the objects largely sequentially.
- This fits well with other results showing that speech planning is highly incremental (e.g., Meyer, 1996; Dell et al., 1997).
- Visual attention and speech planning are governed by a common plan. Executive control processes govern the implementation of the plan.

# A working model of multiple object naming



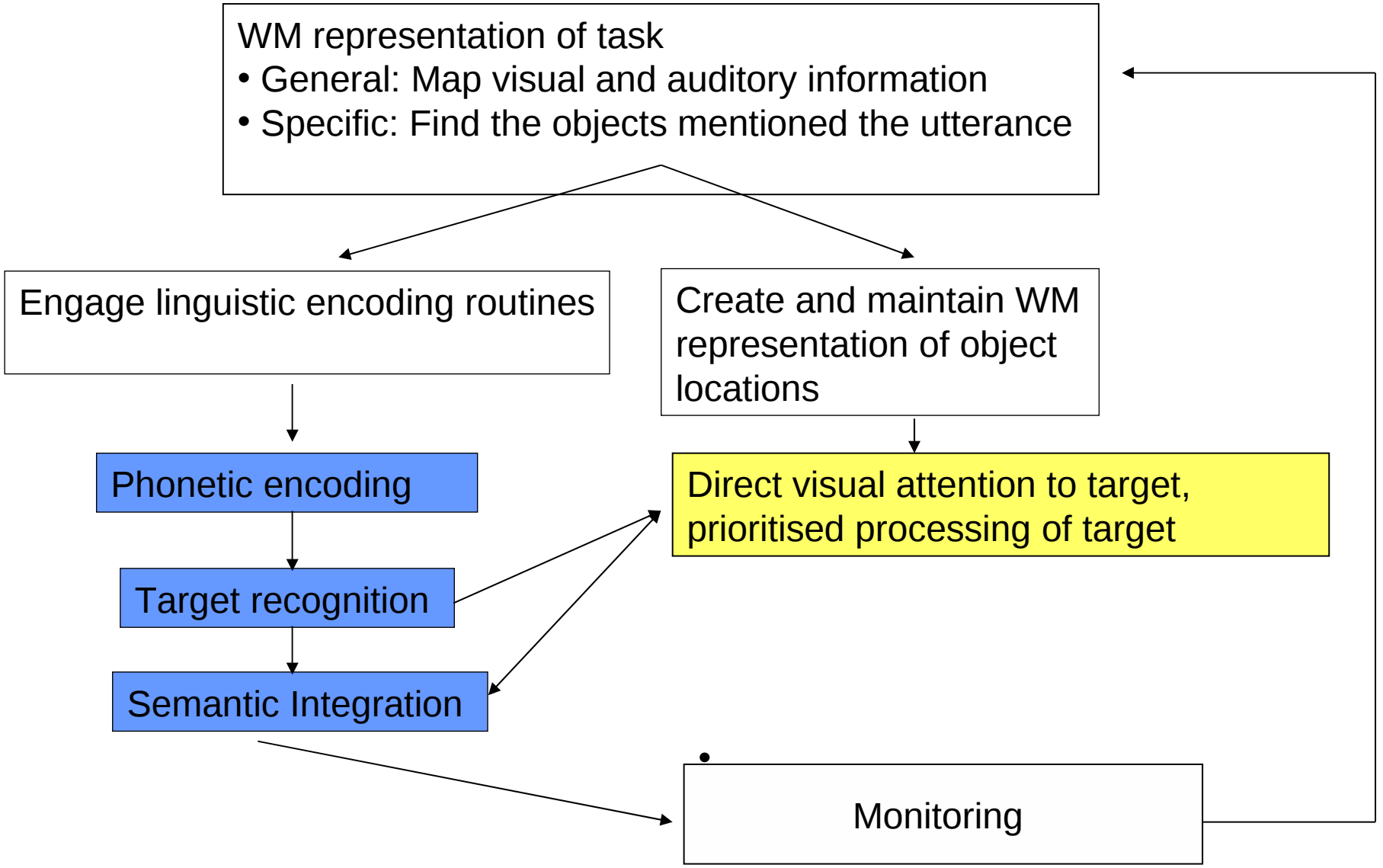
# A working model of multiple object naming



# Linking eye gaze to speech comprehension

- Listeners look at task-relevant objects because this facilitates carrying out the task.
- The task can be explicit (“look at the tree”) or implicit, mapping visual and auditory representations (see Altmann & Kamide, 2009).
- Objects become relevant when they are mentioned or implied.

# A working model of situated listening



# Looks to incorrect locations

- Listeners look at
  - a hammer instead of a hamster (Salverda et al., 2003)
  - a rope instead of a snake (Huettig & Altmann, 2007)
  - a strawberry instead of lips (Huettig & Altmann, in press).
- Active working memory representations of objects attract attention (usually to task-relevant object, sometimes to a distractor)

# Summary

- A key issue in current psycholinguistics is how language processing is related to visual processing, attention and executive control.
- This can be studied in eye tracking experiments.
- The results suggest straightforward eye-speech links.
- However, these links are mediated by attentional processes
  - visual attention governs eye movements
  - executive control processes govern visual attention
- We are working towards comprehensive yet detailed models of the language-attention-vision interplay.

# Collaborators and sponsors


Collaborators:

Eva Belke Steven Frisson Christine Häcker

Falk Huettig Glyn Humphreys Agnieskza Konopka

Debra Malpass Jane Morgan Linda Mortensen

Marc Ouellet Joost Rommers Anna Telling

Linda Wheeldon

Sponsors:

British Academy BBSRC ESRC

Nuffield Foundation