LEXICAL ACCESS IN THE VISUAL MODALITY

Diogo Almeida

diogo@nyu.edu | www.diogoalmeida.net

جامعة نيويورك أبوظبي

NYU | ABU DHABI

OUTLINE

1. Introduction
2. Visual word recognition and the structure of the lexicon
3. Interim discussion: Visual word recognition and Language
4. Adaptations of the visual system for visual word recognition
5. Conclusion

INTRODUCTION

INTRODUCTION

Visual word recognition

- Ability to retrieve a word (or lexical item) from one’s mental dictionary (lexicon)
- Mostly written language
- But also sign language
- Mostly words in isolation
- But sometimes also in sentential context

VISUAL WORD RECOGNITION AND THE STRUCTURE OF THE LEXICON

WHAT'S IN A WORD?

Major assumption: knowledge of language is *not* just a list of facts

Because language is productive

- *The stern-faced elephant and his unicorn friends were disappointed to learn that the moon was not made of cheese.*
- You never heard this sentence before
- you can still understand it, and accept it as a possible, albeit weird, sentence of English
  - We don't *memorize* sentences, we *parse* them!

WHAT’S IN A WORD?

We don’t *memorize* sentences, we *parse* them!

- (Virtually) no controversy there!
- But we do memorize some things!
  - *Words* are good candidates for things that can be stored.
- Controversy about words:
  - What exactly is *memorized* and what is *parsed*

WHAT'S IN A WORD?

Major controversy: Storage vs computation in the lexicon

- Everyone agrees that there is storage
- Disagreement is about the amount of computation/parsing in the lexicon
- Two poles of the debate
  - Full decomposition
  - Full storage

DICTIONARY ANALOGY

Physical dictionaries (the books) implement the following architecture

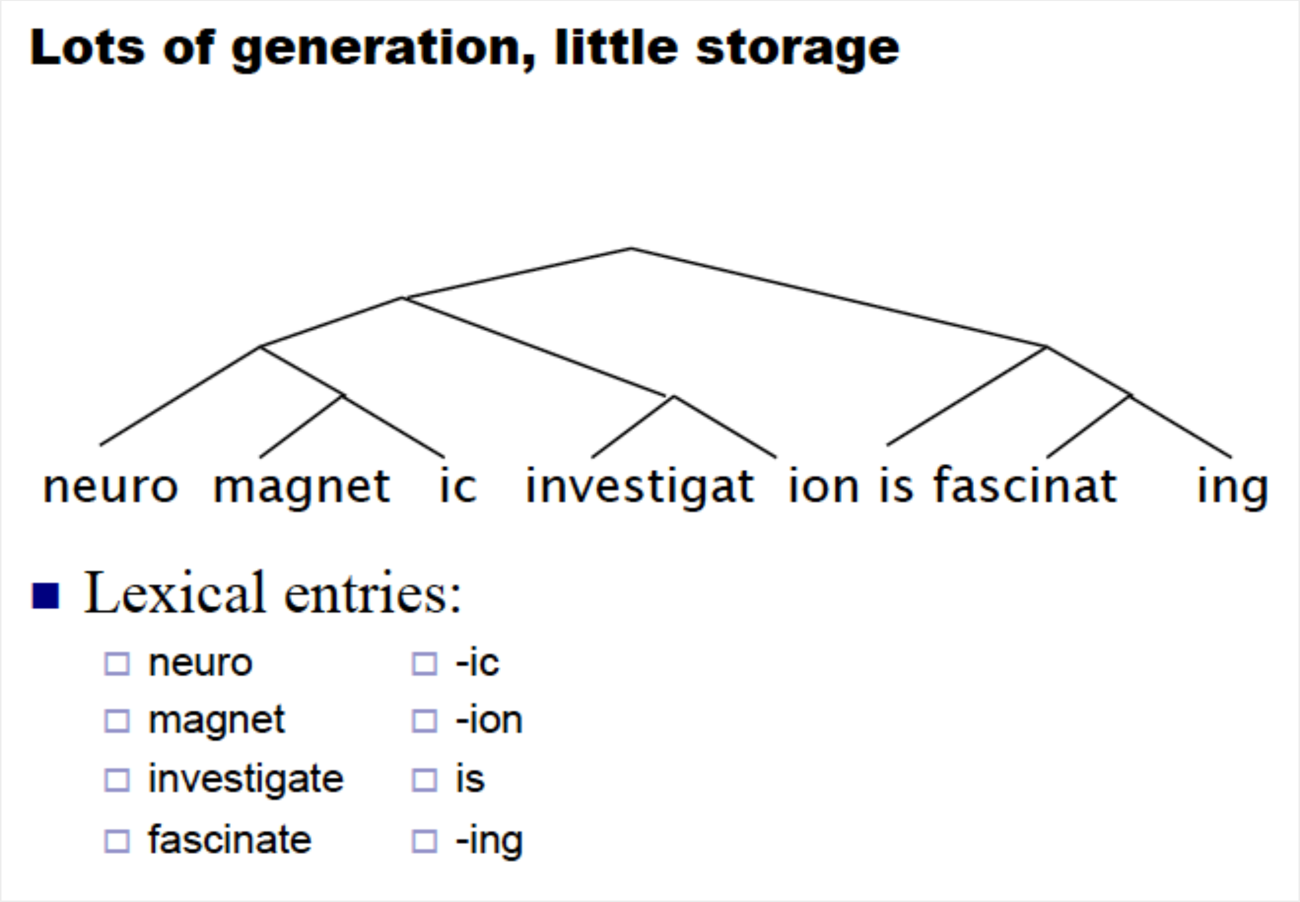
- **Inflected forms:** stored as part of a base form
- **Derived forms:** stored separate from their base form

Two poles of the debate

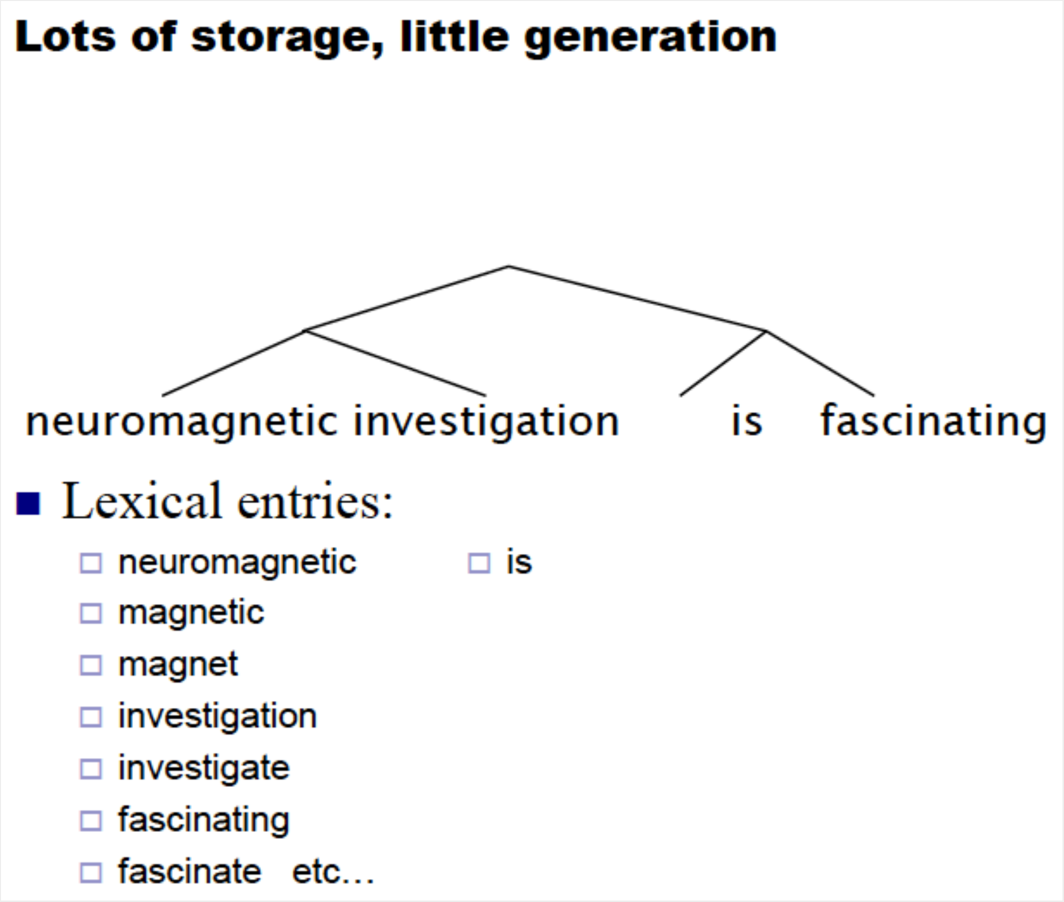
- Full decomposition
  - Traditional dictionaries: *Partial* decomposition
- Full storage

FULL DECOMPOSITION VS FULL STORAGE

representation vs computation in the lexicon



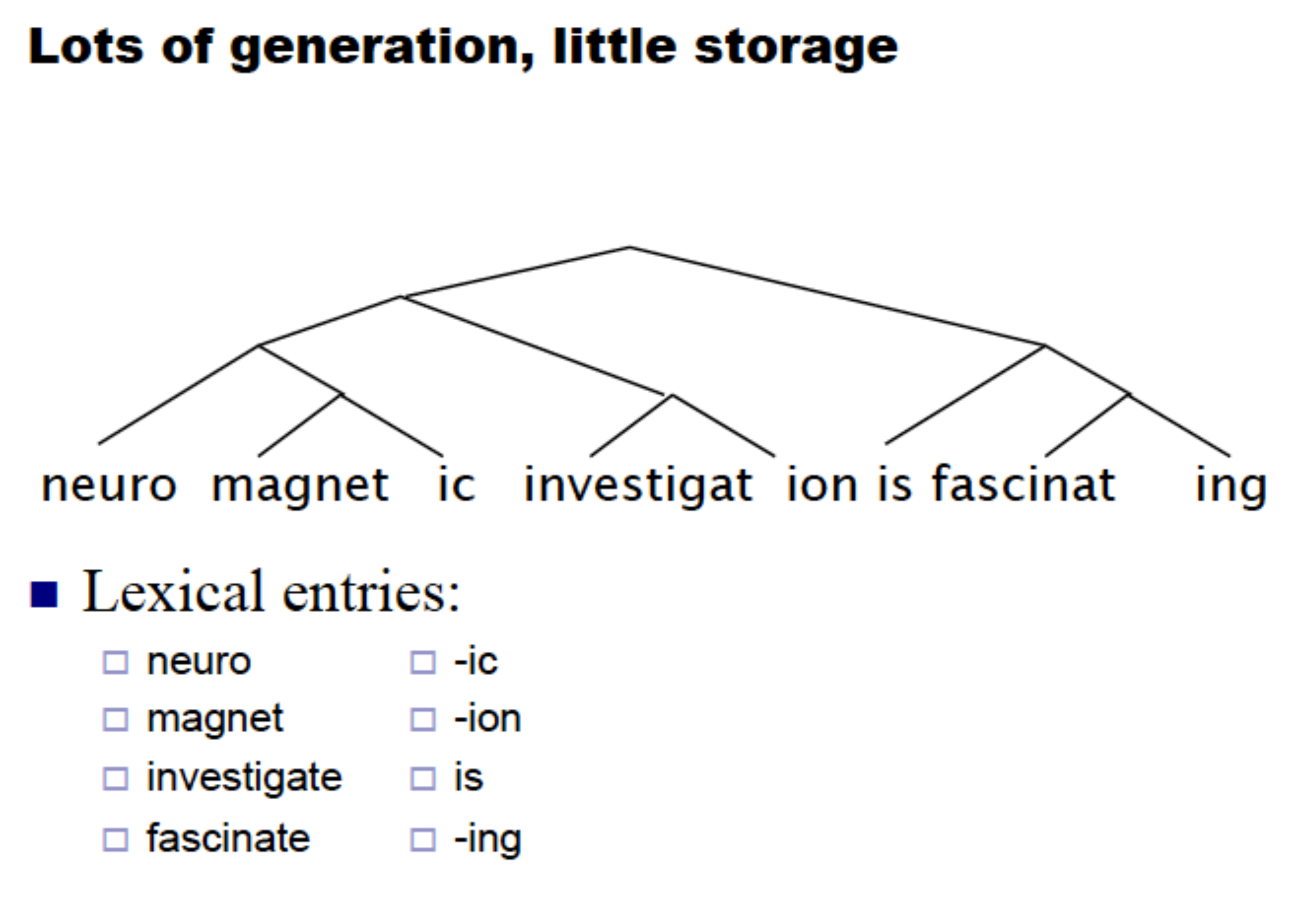
Words are always broken down into constituents



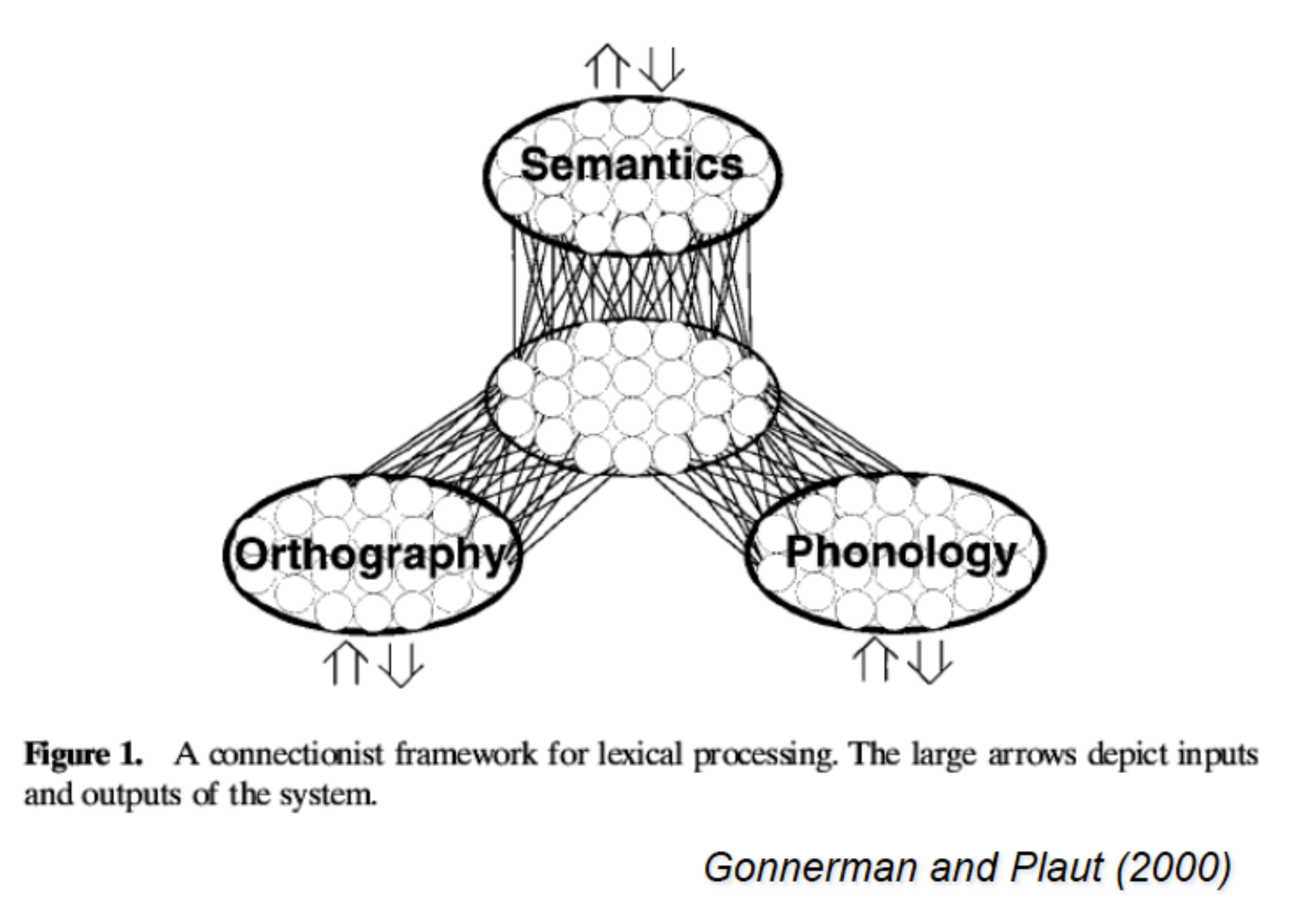
Word structure is an epiphenomenon (similarity to the extreme)

FULL DECOMPOSITION VS FULL STORAGE

representation vs computation in the lexicon



Words are always broken down into constituents



Word structure is an epiphenomenon (similarity to the extreme)

TEASING THESE EXPLANATIONS APART

Sources of evidence for the debate:

- Very simple behavioral experiments
- Psycholinguistic experiments
- Neurolinguistic experiments

VERY SIMPLE TRADITIONAL LINGUISTIC EXPERIMENT

Un-lock-able	Un-health-y
ambiguous	**not ambiguous**

TEASING THESE EXPLANATIONS APART

Sources of evidence for the debate:

- Very simple behavioral experiments
  - Intuitions about *possible ambiguity* are actually about *parseability*
  - The full decomposition model *naturally accounts* for the possible ambiguity intuitions.
  - No obvious way for the full storage account to do the same.

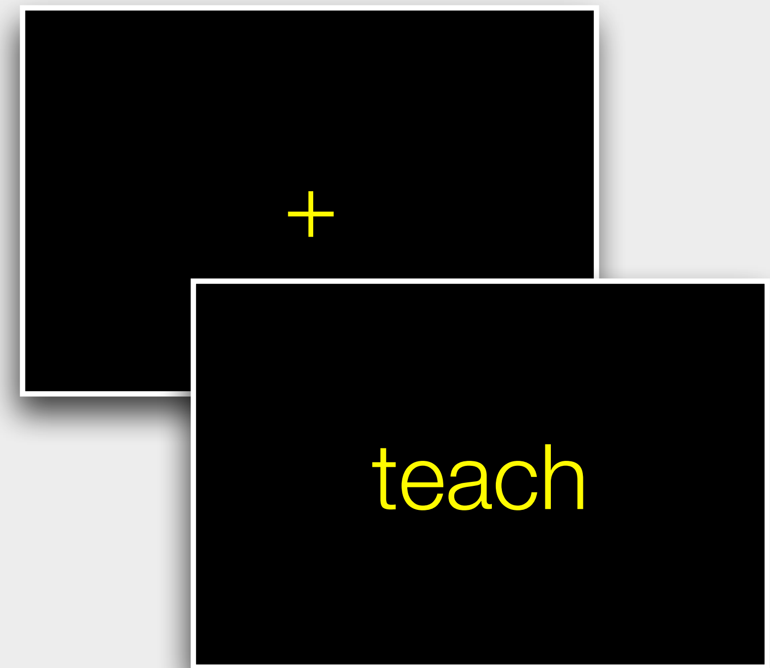
TEASING THESE EXPLANATIONS APART

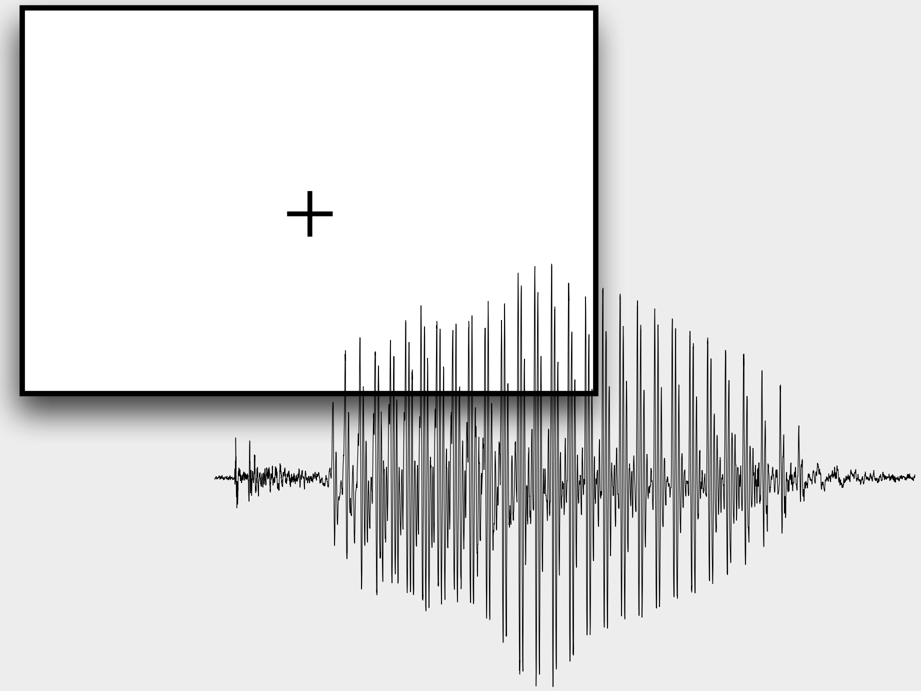
Sources of evidence for the debate:

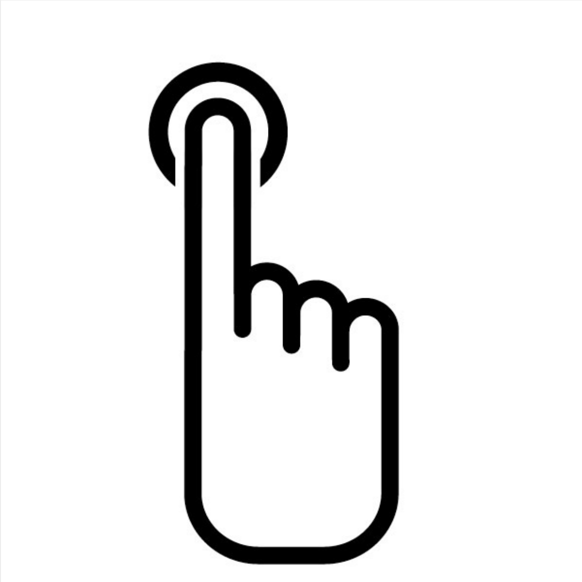
- Very simple behavioral experiments
- Psycholinguistic experiments
- Neurolinguistic experiments

LEXICAL DECISION TASK

(IS THIS A WORD?)







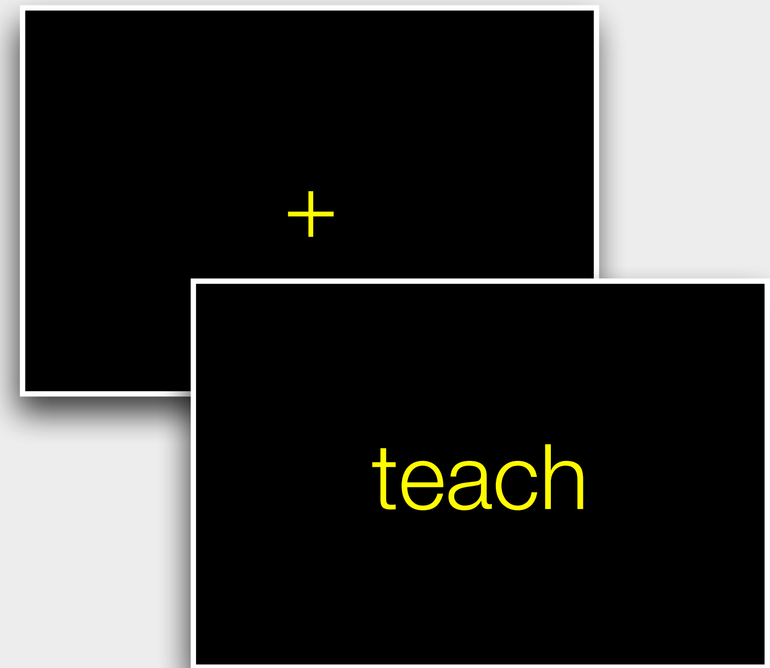
Visual / Auditory processing; Phonetic interpretation of auditory / othographic input; Lexical Search; Monitoring processes; Lexical retrieval; Planning of motor action; Execution of motor action;

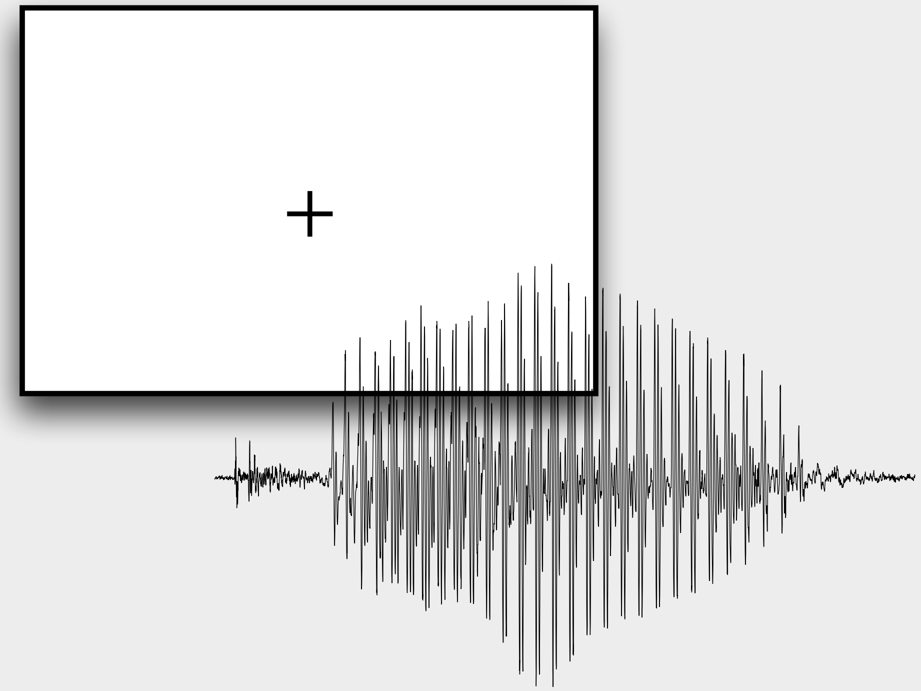
time

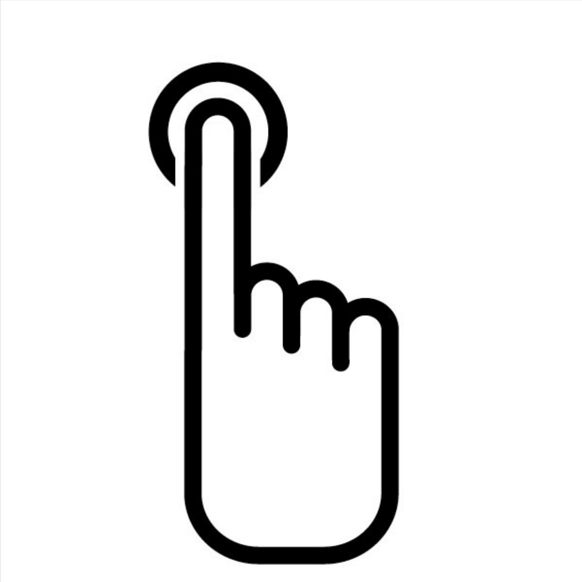
word presentation: 0ms

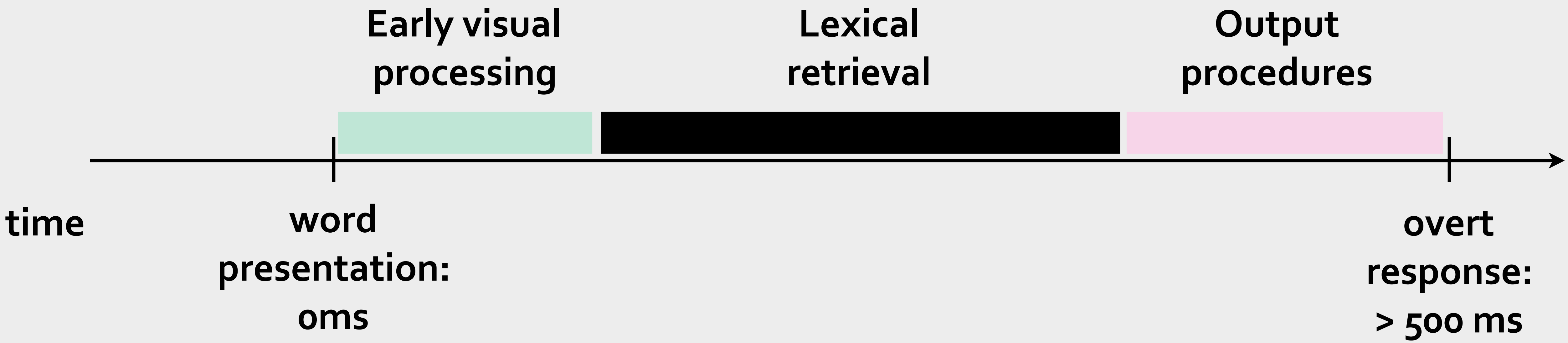
overt response: > 500 ms

LEXICAL DECISION TASK
(IS THIS A WORD?)

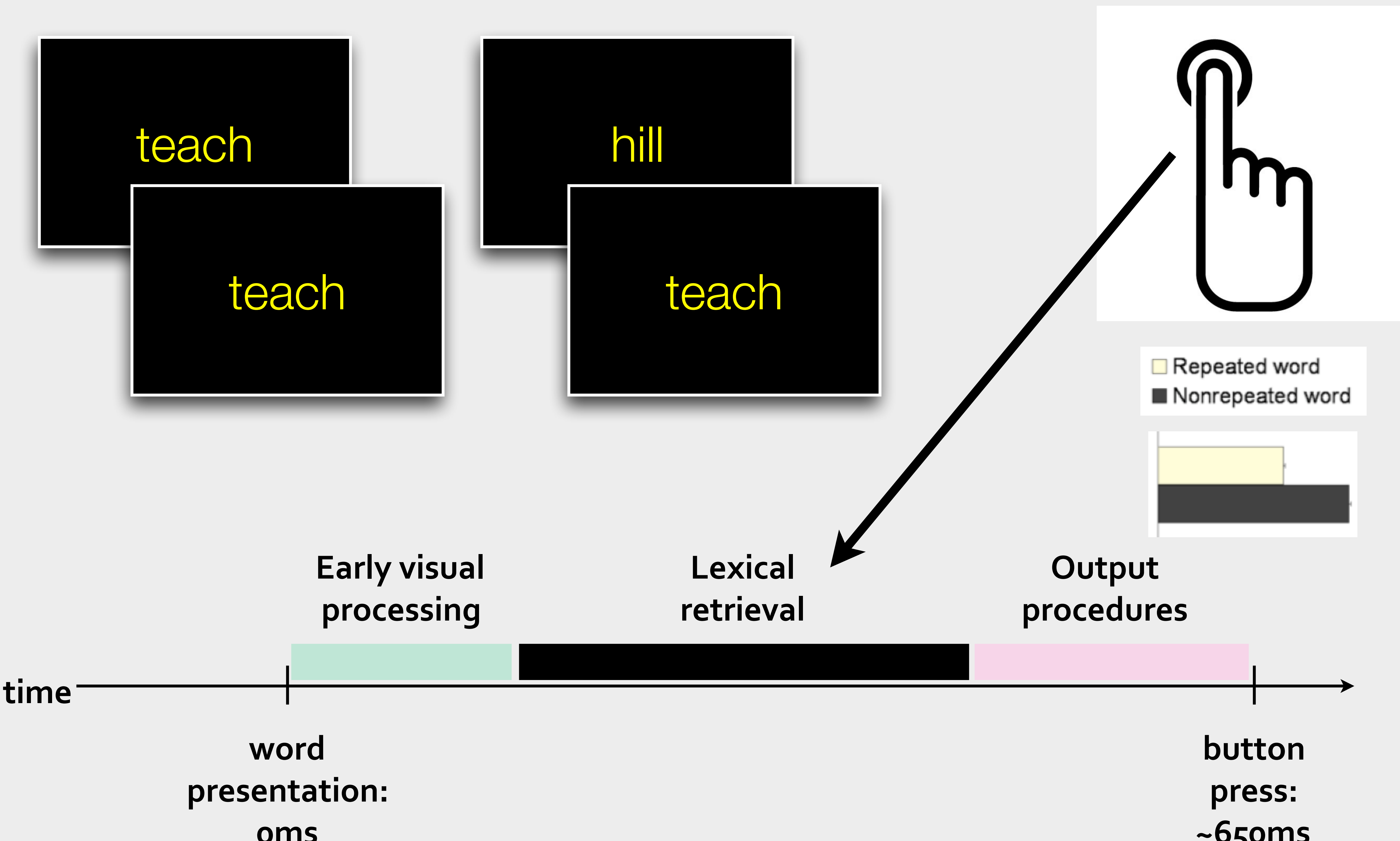




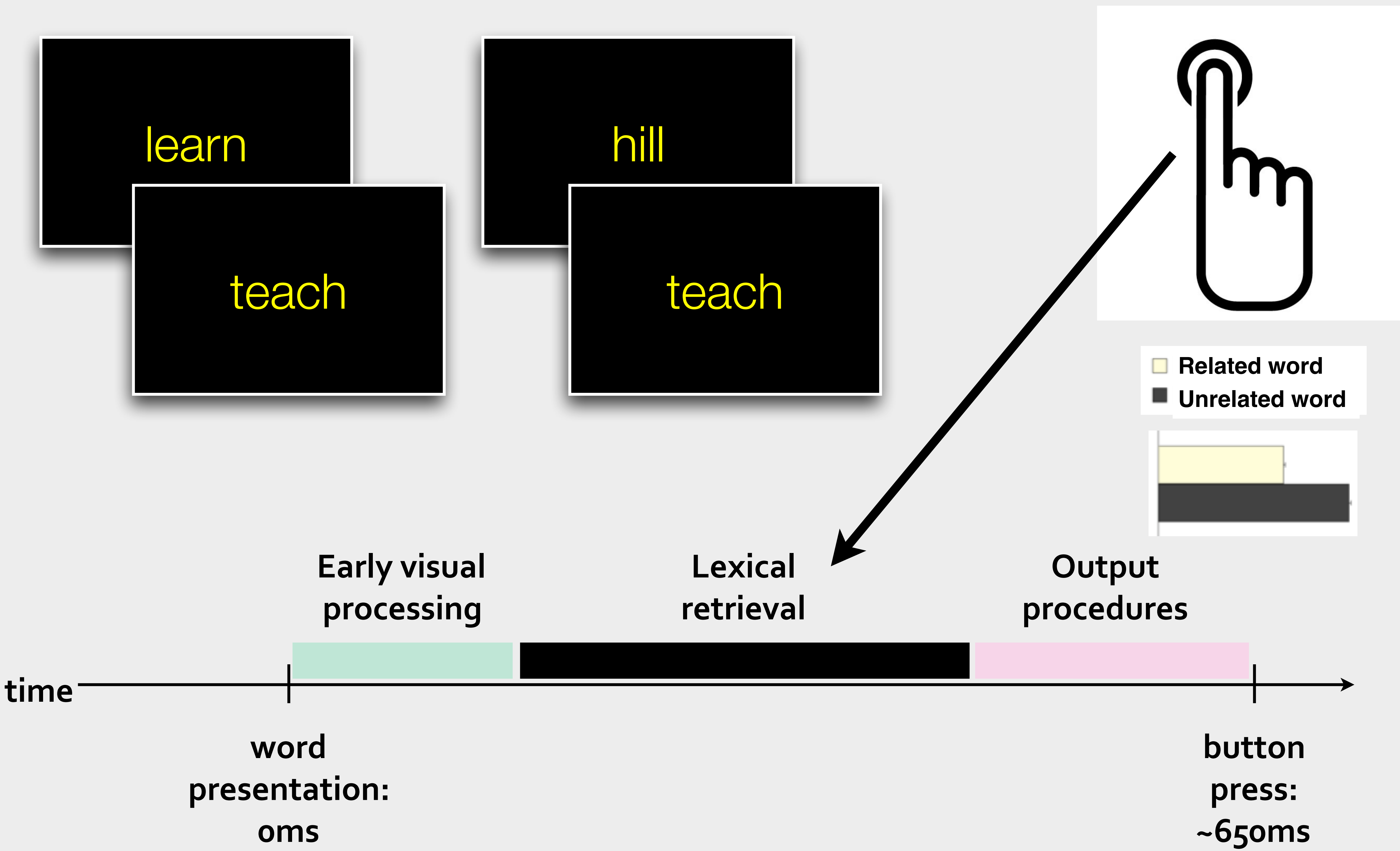




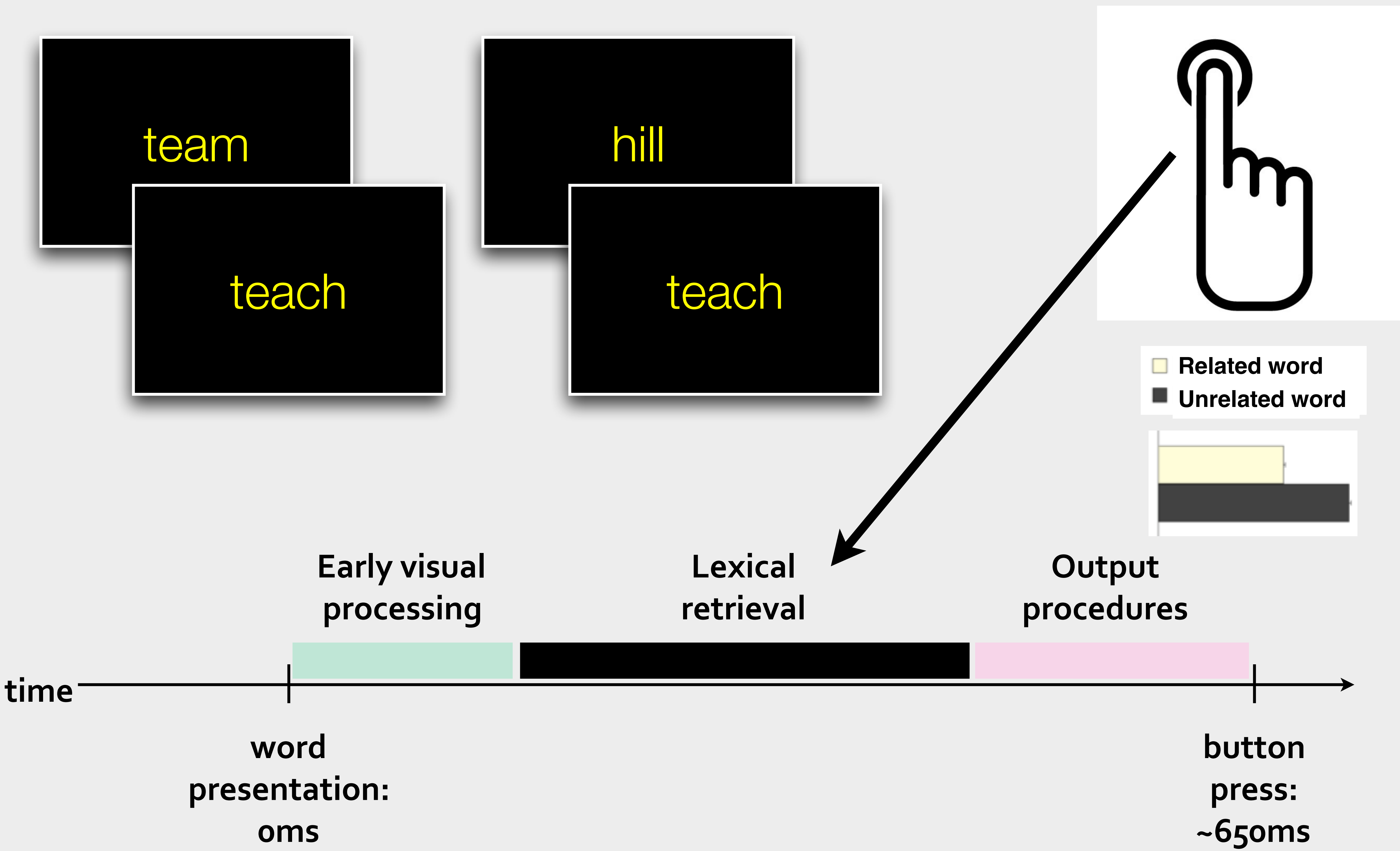
REPETITION PRIMING IN LDT



SEMANTIC PRIMING IN LDT



FORM PRIMING IN LDT



CRITICAL CASE: 'MORPHOLOGICAL' PRIMING

Repetition

teach/hill

teach

Semantic

learn/hill

teach

Form

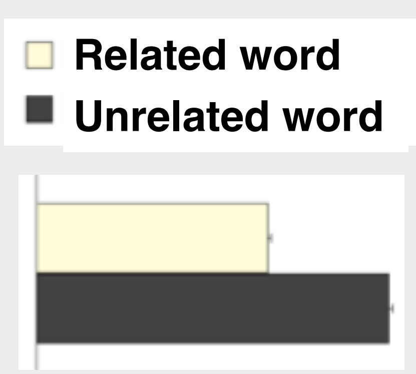
team/hill

teach

Morphological

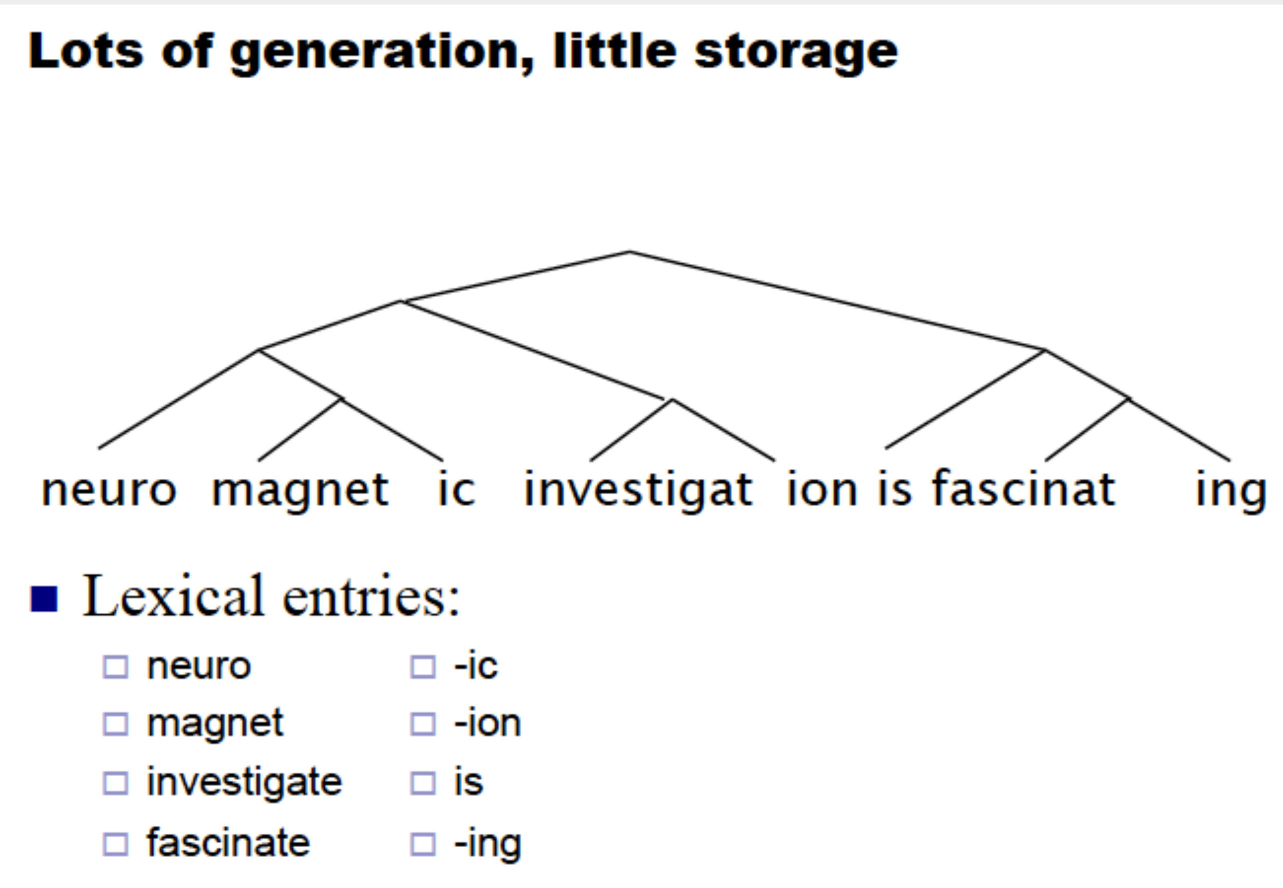
teacher/hill

teach



FULL DECOMPOSITION VS FULL STORAGE

representation vs computation in the lexicon



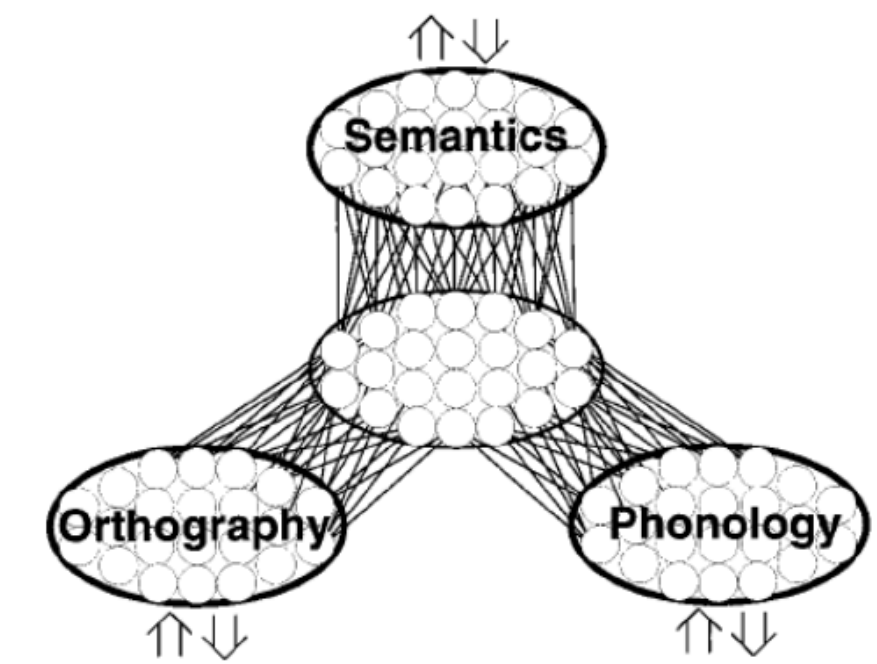


Figure 1. A connectionist framework for lexical processing. The large arrows depict inputs and outputs of the system.

Gonnerman and Plaut (2000)

Words are always broken down into constituents

Word structure is an epiphenomenon (similarity to the extreme)

CRITICAL CASE: 'MORPHOLOGICAL' PRIMING

Repetition

teach/hill

teach

Semantic

learn/hill

teach

Form

team/hill

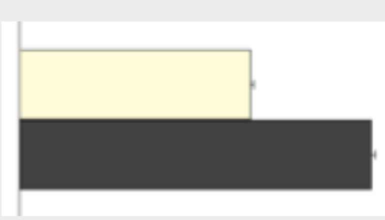
teach

Morphological

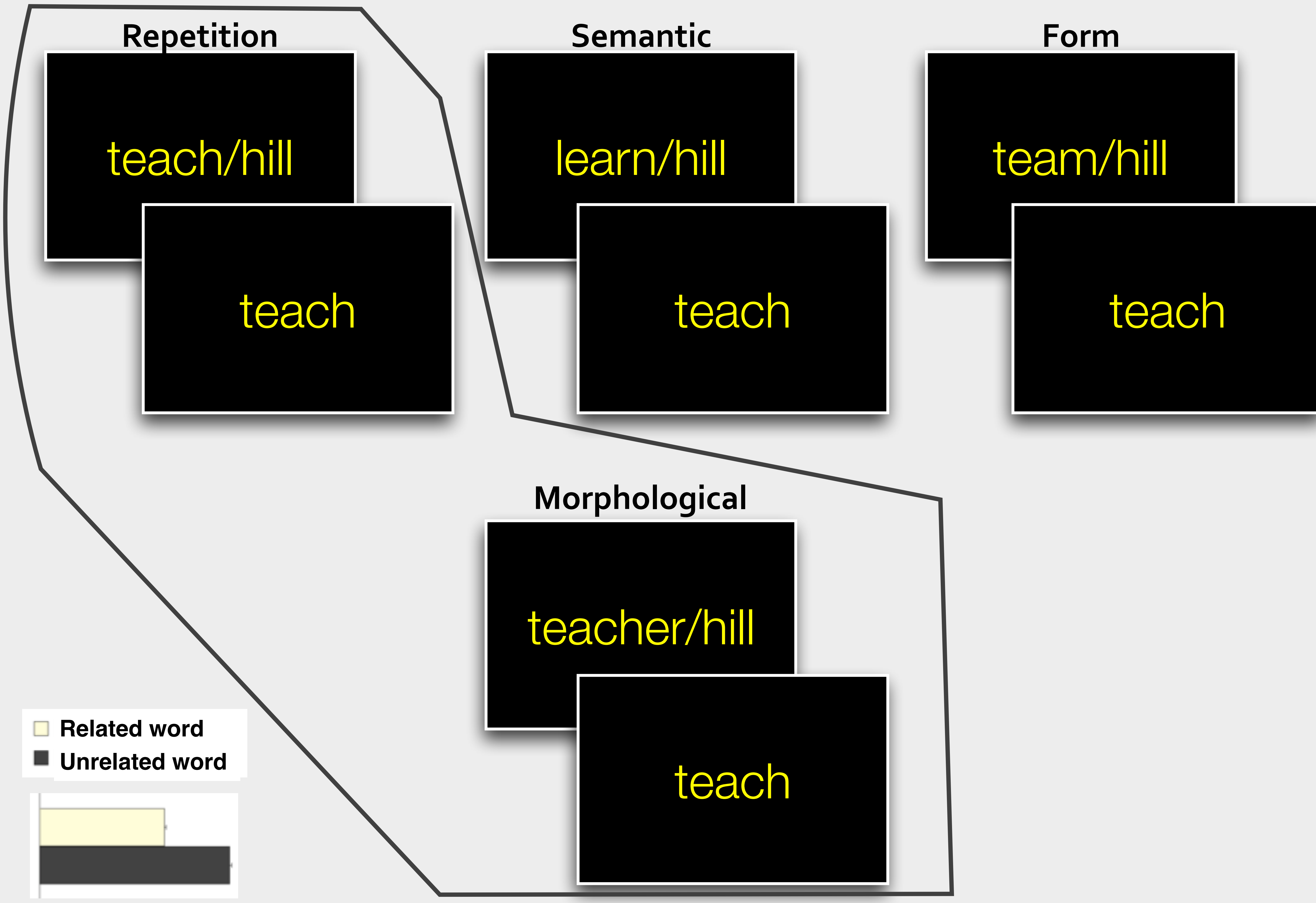
teacher/hill

teach

Related word
Unrelated word



FULL DECOMPOSITION: ‘MORPHOLOGICAL’ PRIMING = REPETITION (IDENTITY) PRIMING



FULL STORAGE:
'MORPHOLOGICAL' PRIMING = FORM + MEANING PRIMING

Repetition

teach/hill

teach

Semantic

learn/hill

teach

Form

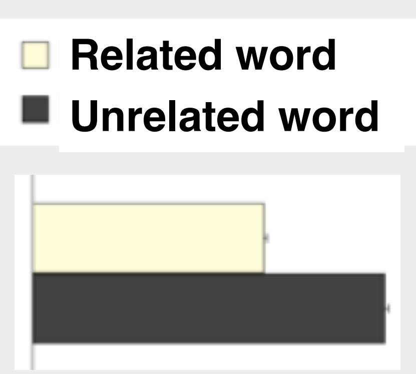
team/hill

teach

Morphological

teacher/hill

teach



HOW DO WE TEASE THESE POSSIBLE EXPLANATIONS APART?

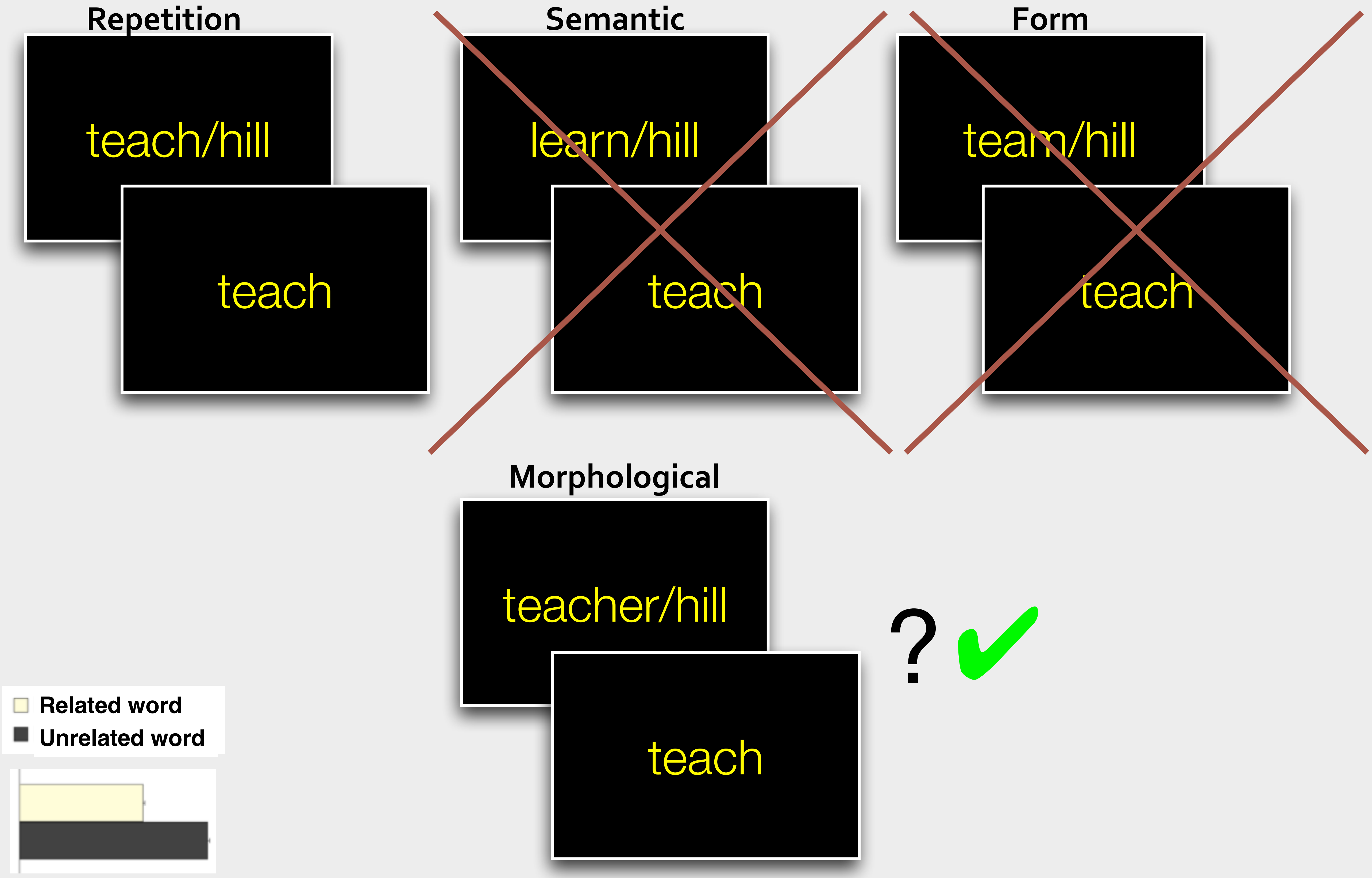
In regular paradigms, you always see priming

- repetition
- semantic
- form
- morphological

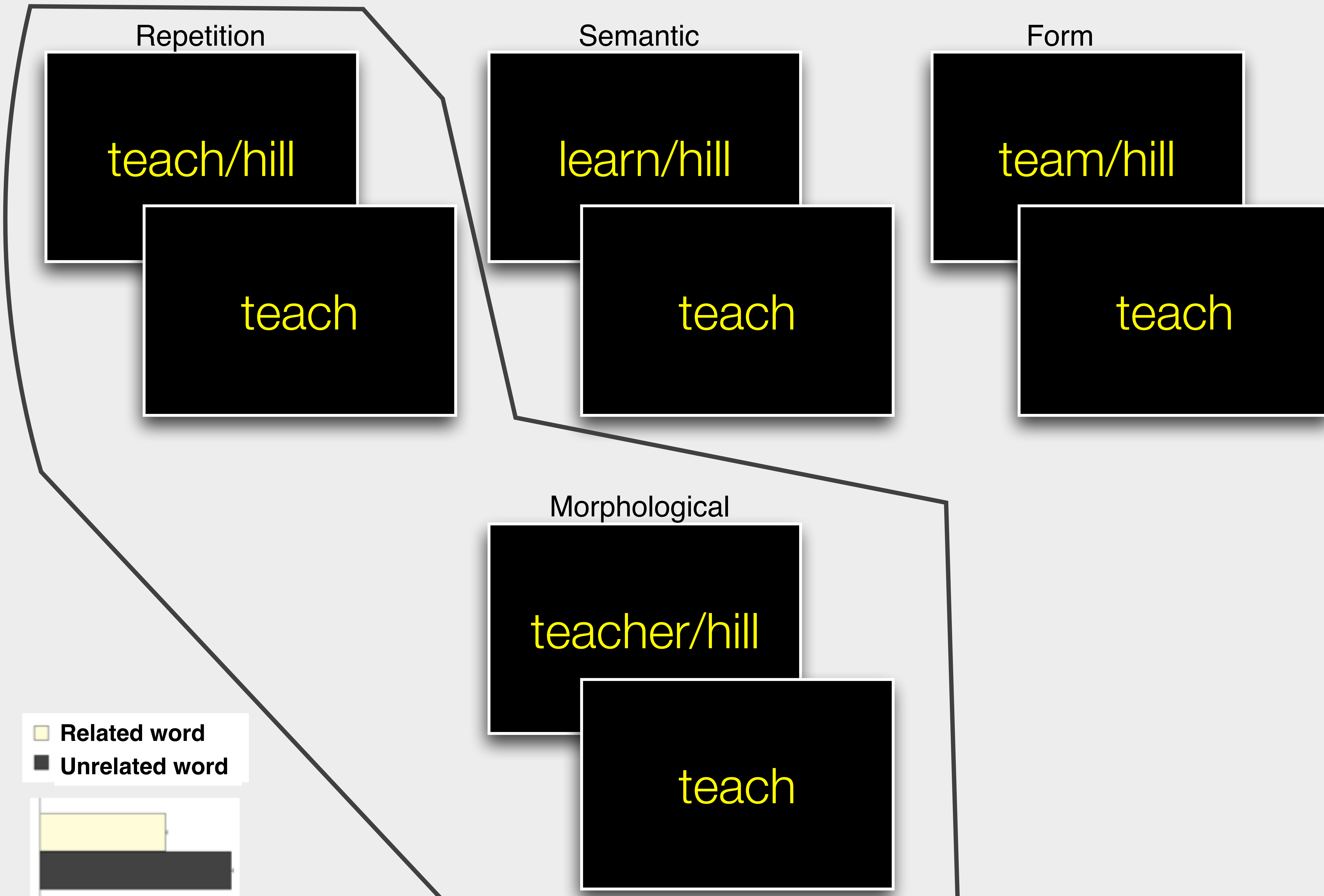
But what if you started seeing dissociations?

- Case of "Masked Priming"

MASKED PRIMING EFFECTS

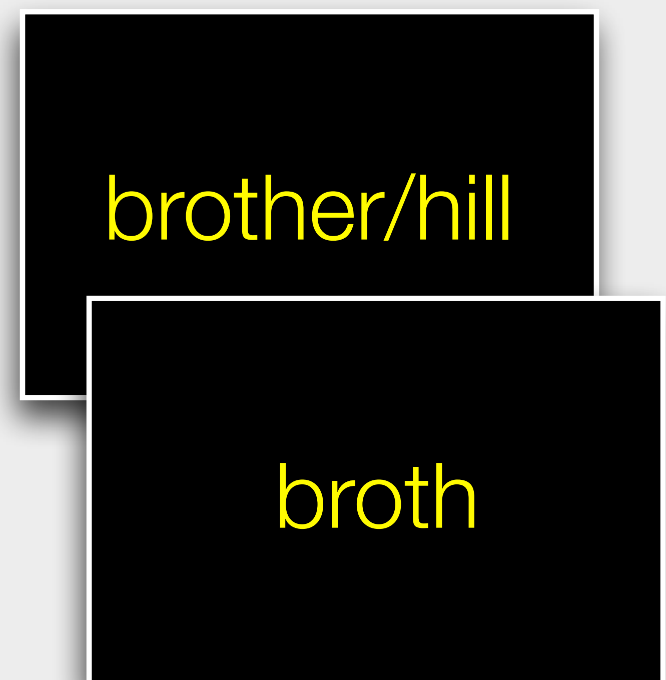


FULL DECOMPOSITION: ‘MORPHOLOGICAL’ PRIMING = REPETITION (IDENTITY) PRIMING

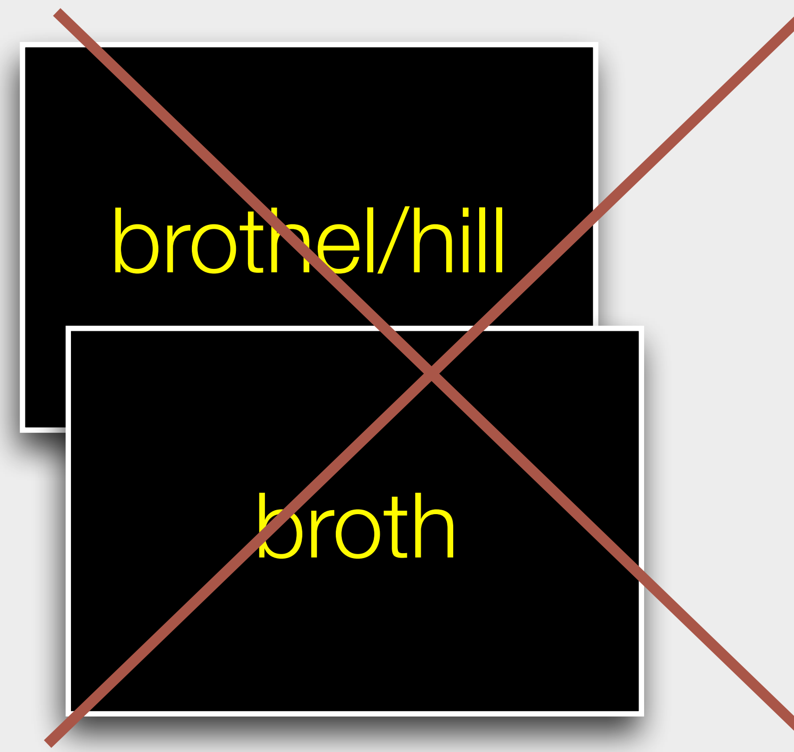


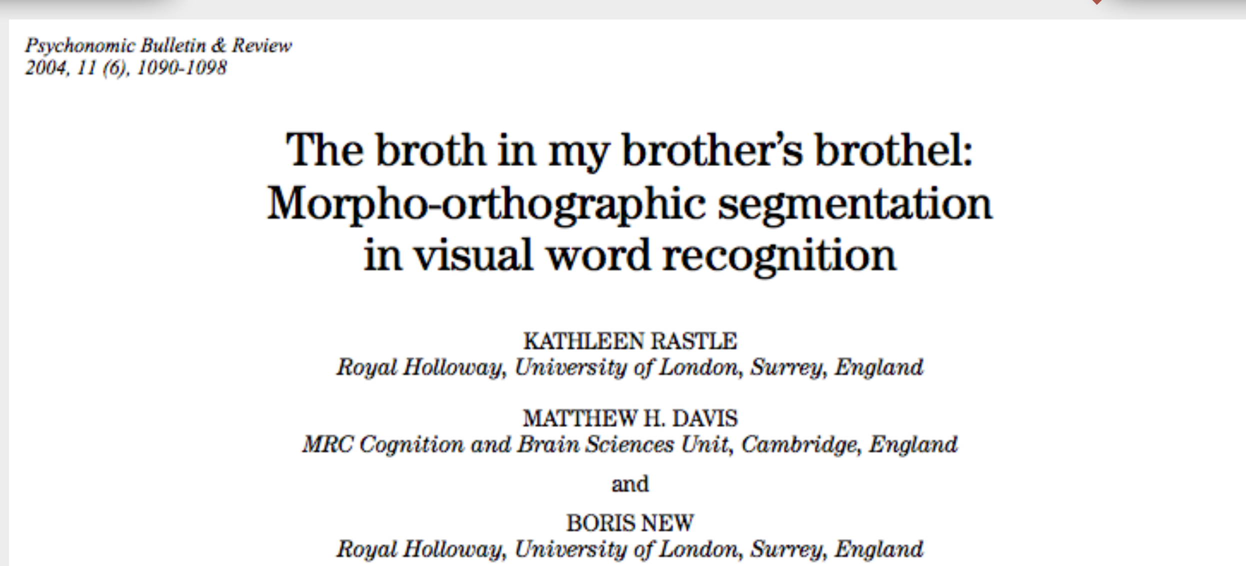
MORPHOLOGICAL PRIMING EVEN IN THE ABSENCE OF MEANING SIMILARITY

morphological decomposition available even if nonsensical, since '-er' is an affix



morphological decomposition unavailable, since '-el' is not an affix





Psychonomic Bulletin & Review
2004, 11 (6), 1090-1098

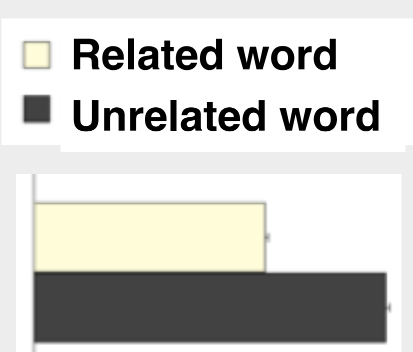
The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition

KATHLEEN RASTLE
Royal Holloway, University of London, Surrey, England

MATTHEW H. DAVIS
MRC Cognition and Brain Sciences Unit, Cambridge, England

and

BORIS NEW
Royal Holloway, University of London, Surrey, England

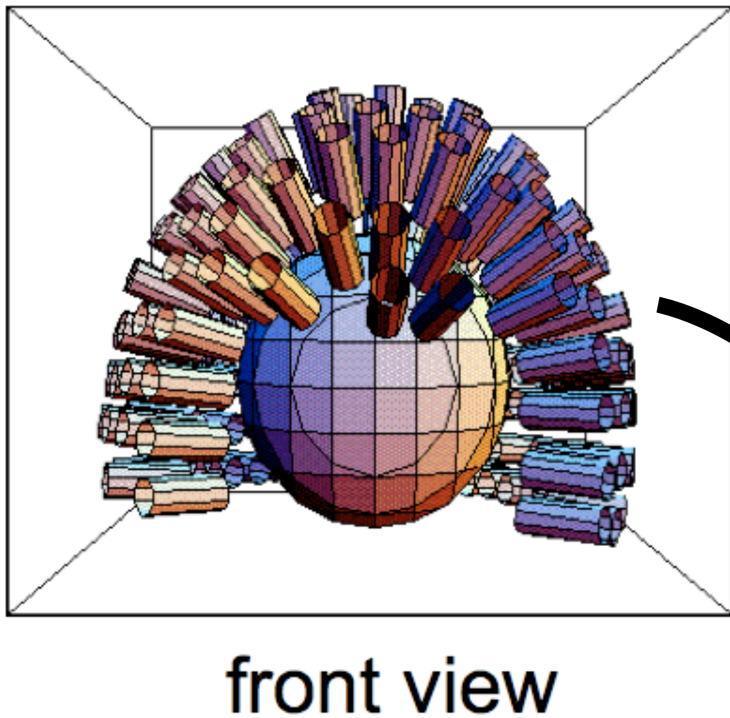


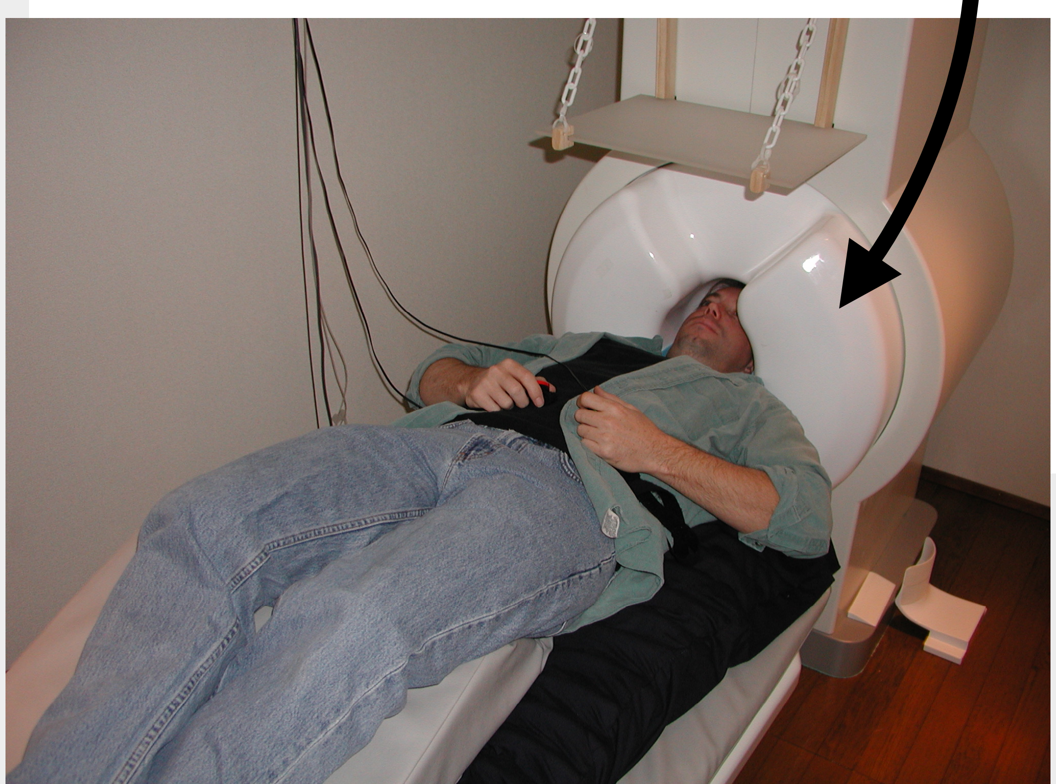
TEASING THESE EXPLANATIONS APART

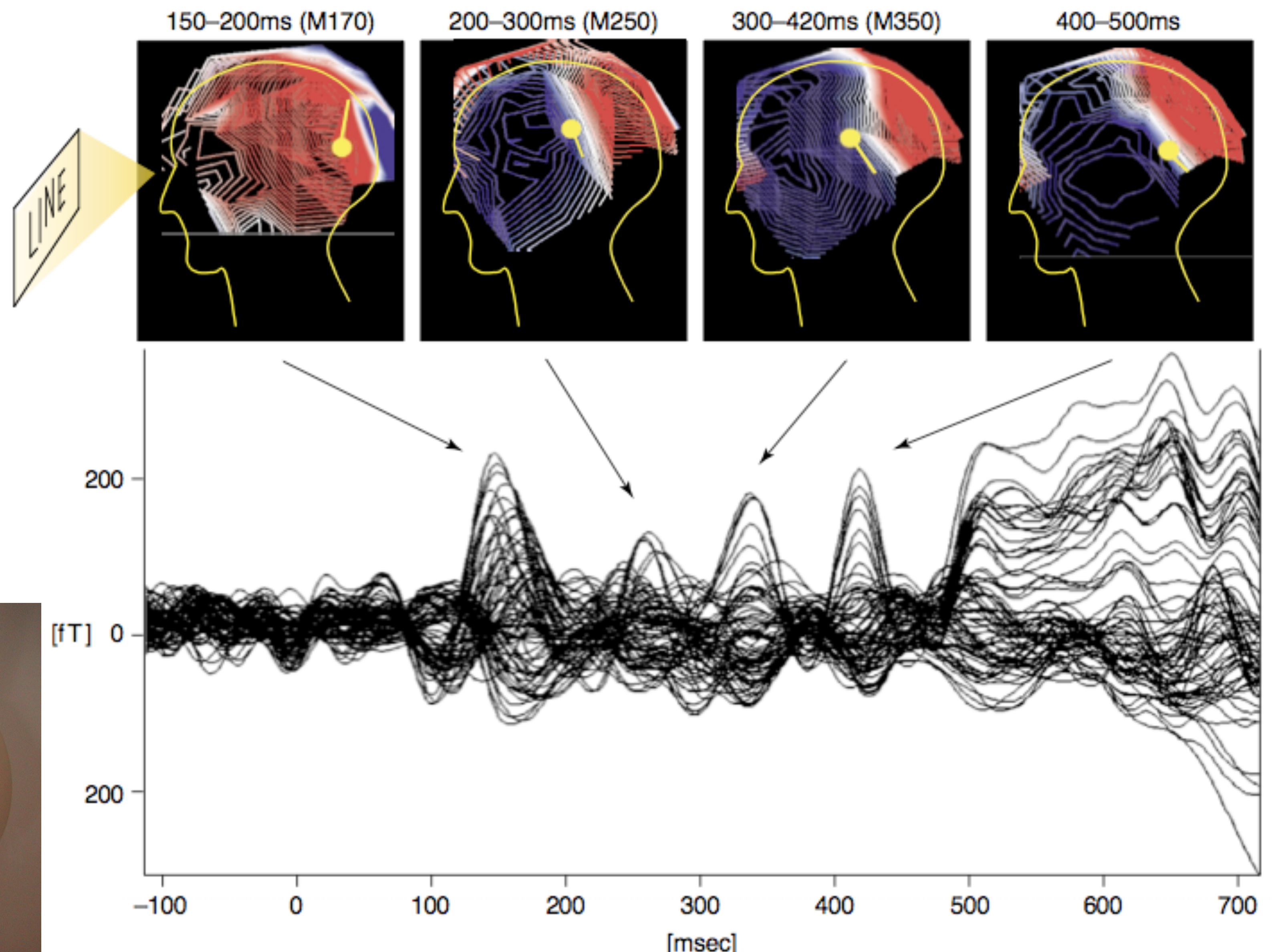
Sources of evidence for the debate:

- Very simple behavioral experiments
- Psycholinguistic experiments
- Neurolinguistic experiments

MEG



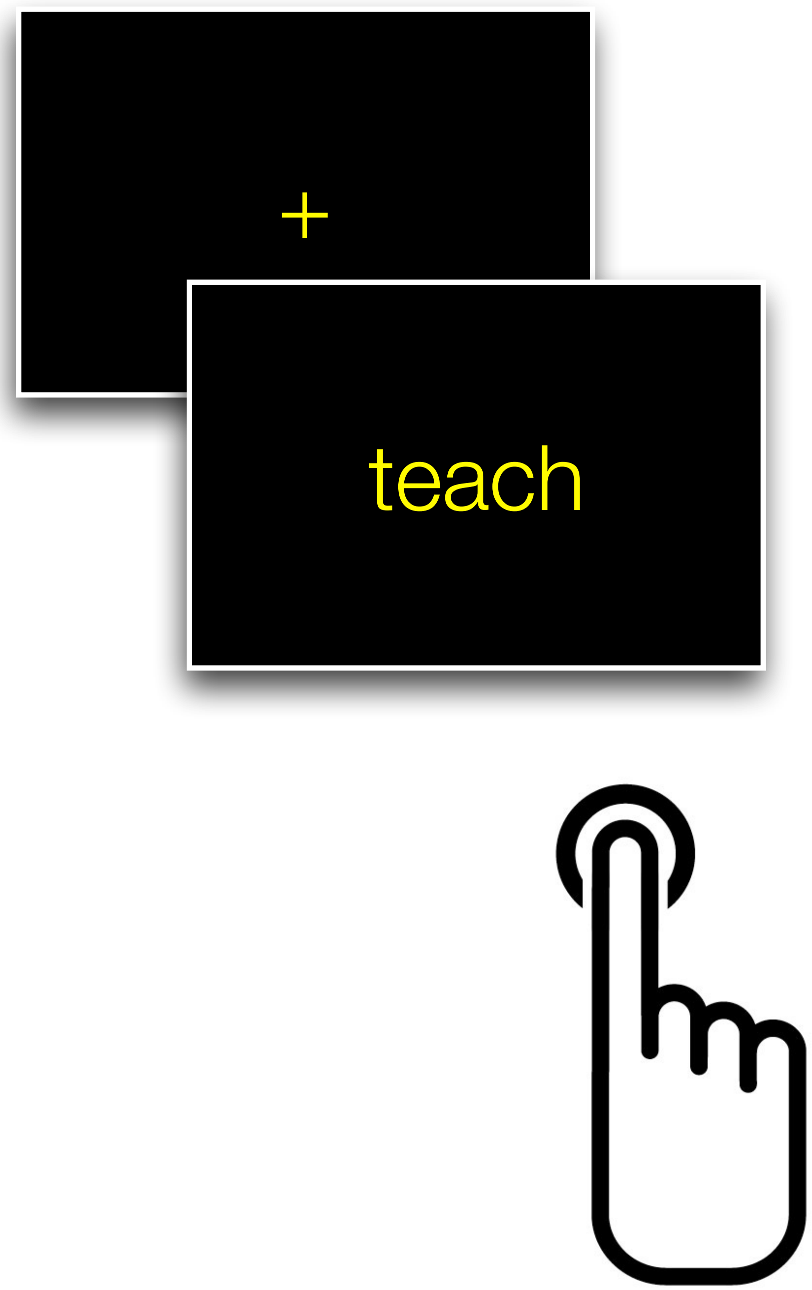


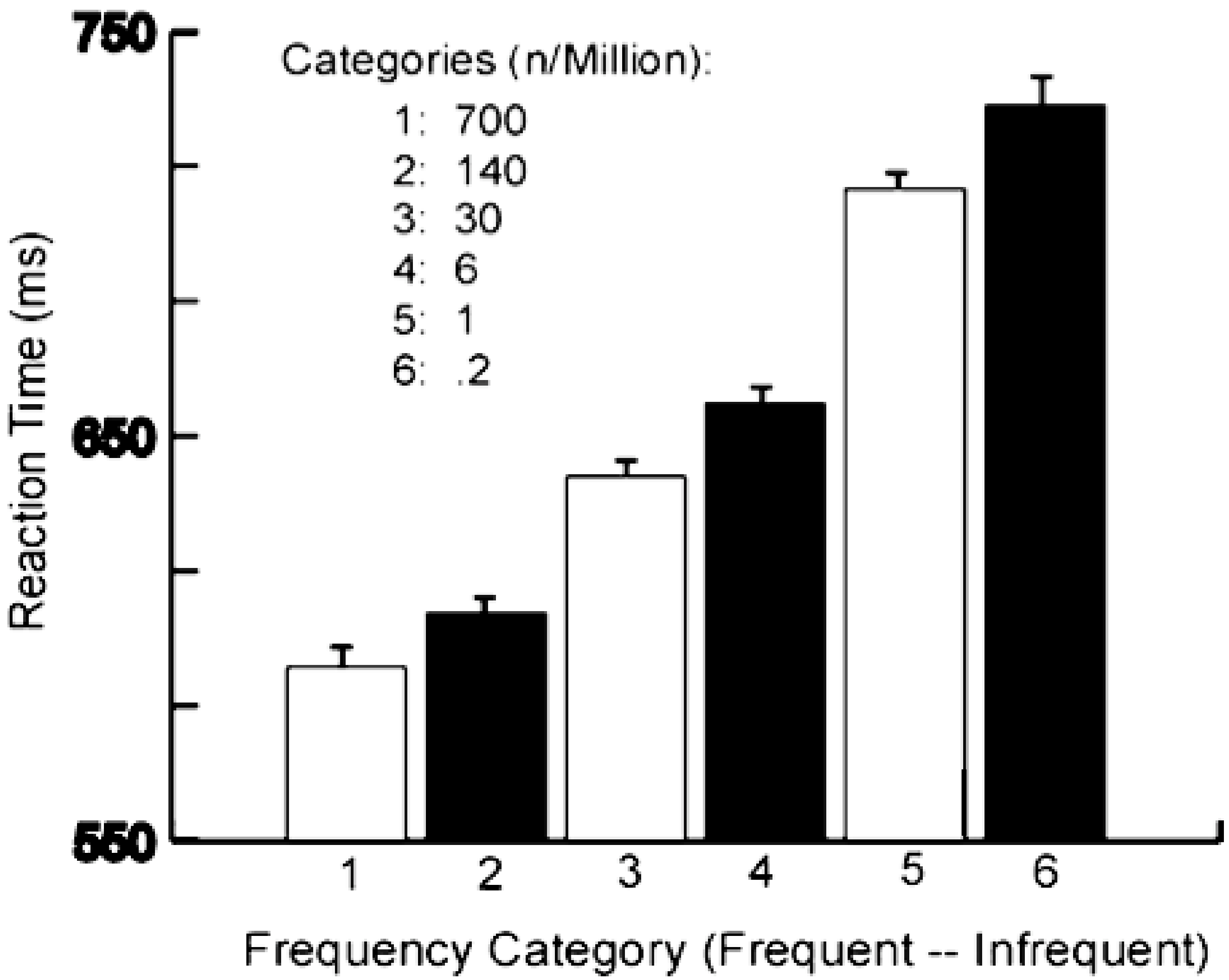


Pylkkänen & Marantz, (2003)

FREQUENCY EFFECT

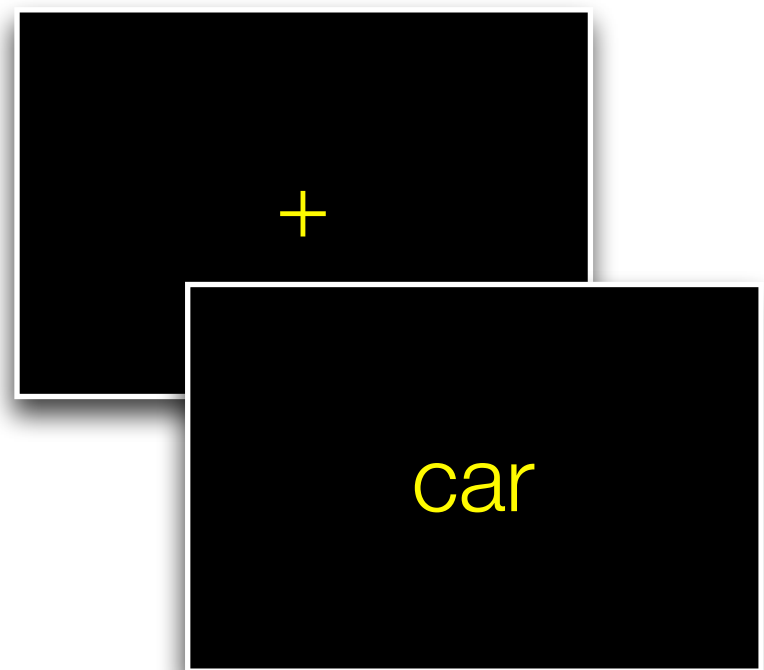
Words that are frequent in the language are responded to faster in lexical decision tasks

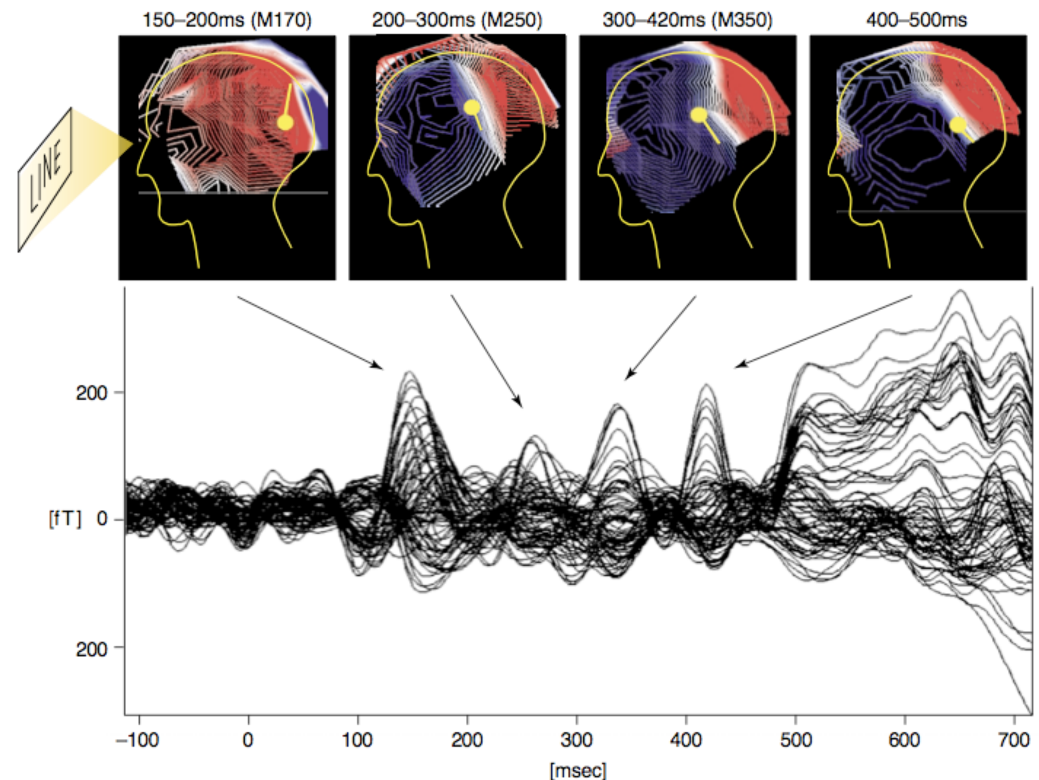


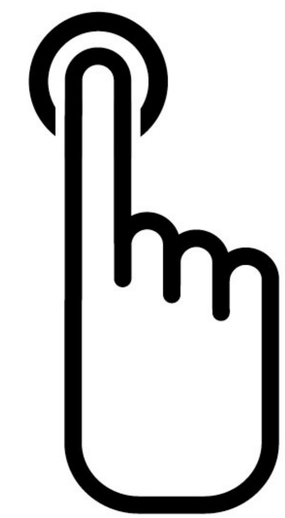


Embick et al. (2001)

MILLISECOND TEMPORAL RESOLUTION IN MONITORING BRAIN ACTIVITY USING MEG







MEG component ?
visual processing

MEG component ?
lexical activation

MEG component ?
lexical selection

MEG component ?
motor response preparation

time

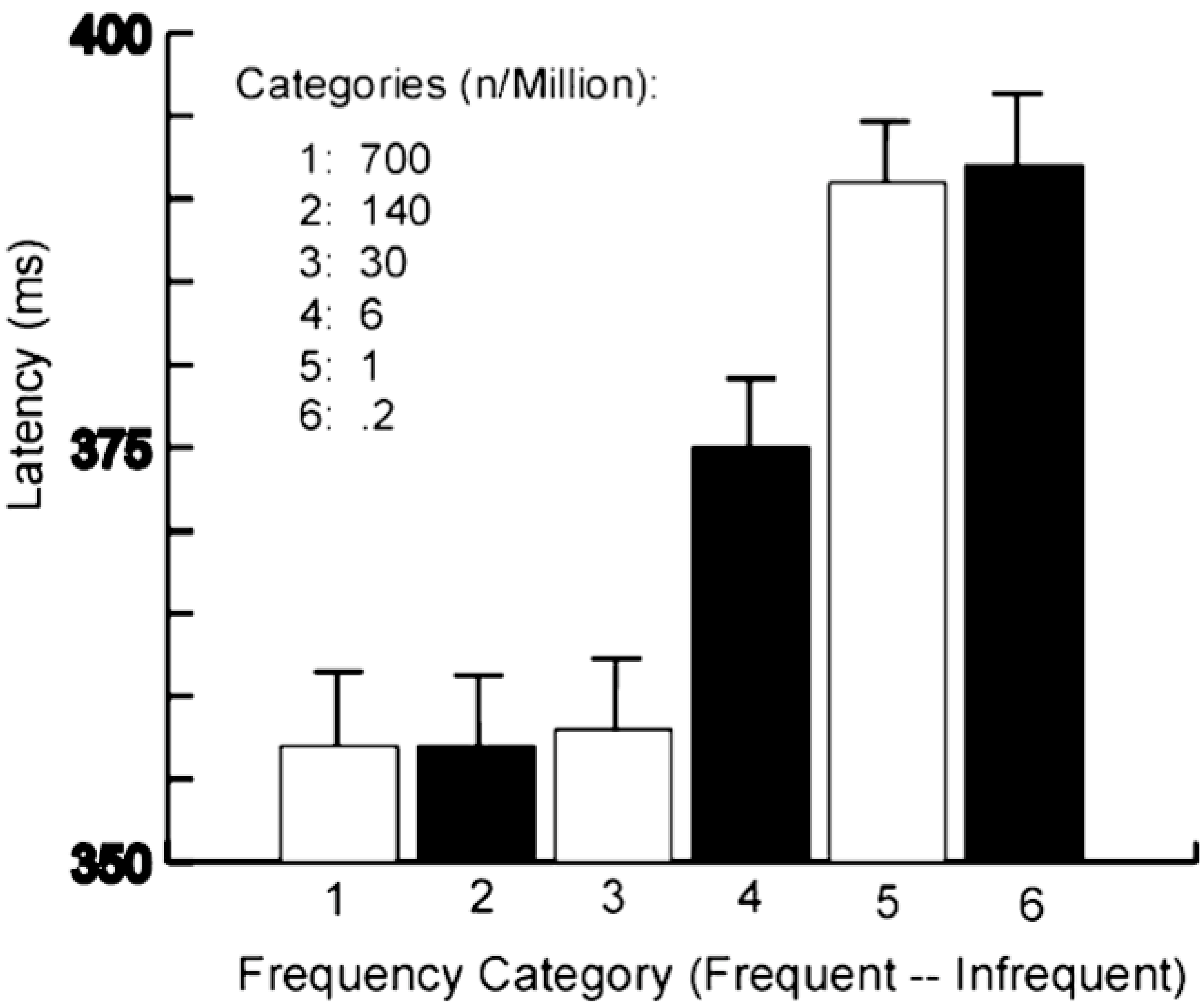
onset of presentation 0ms

button press: ~650ms

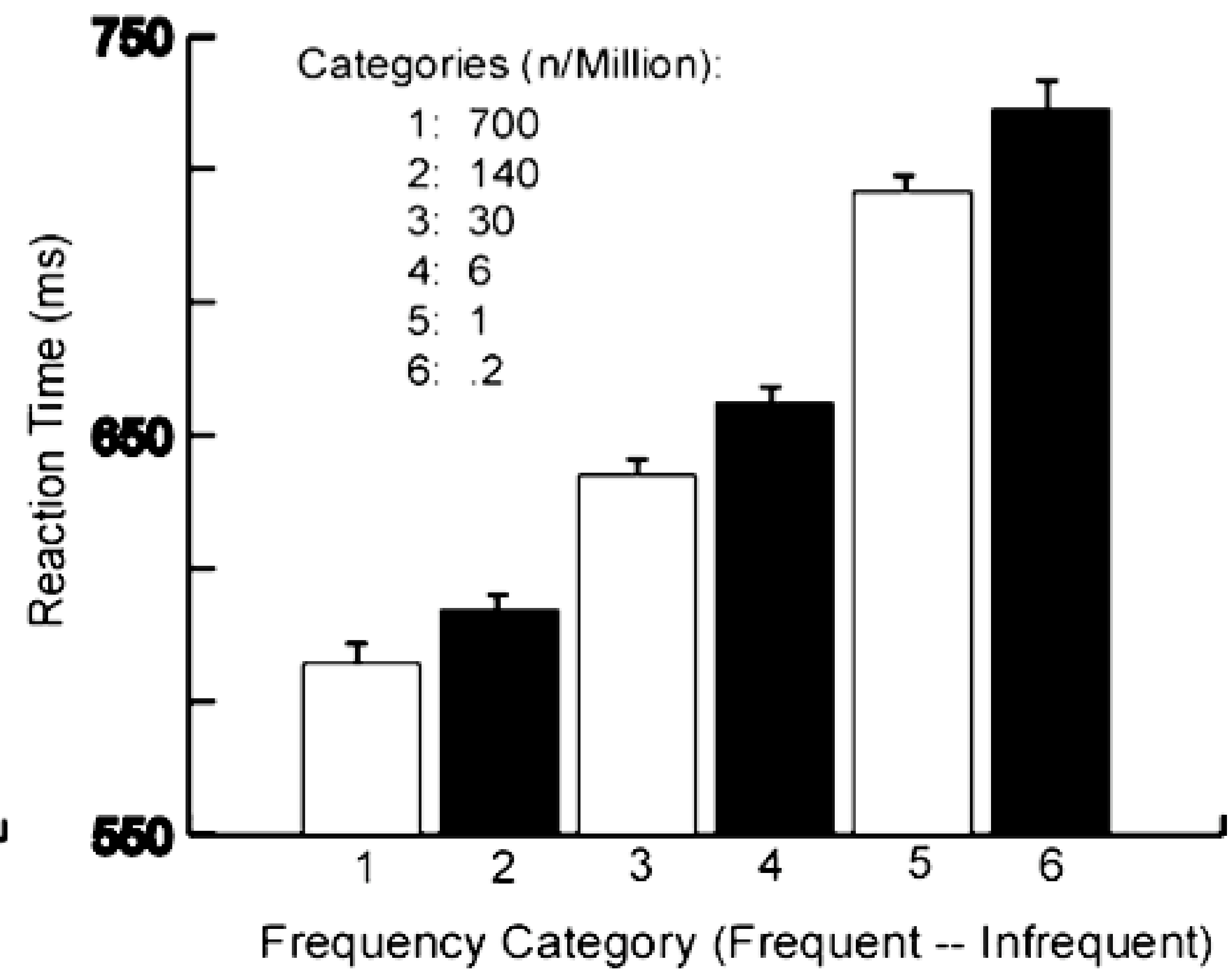
FREQUENCY EFFECT

MEG & Behavioral responses

MEG component: M350



Reaction time



Embick et al. (2001)

FIORENTINO & POEPPEL 2007

Perspective on Internal Structure	*Whole-word representation*		*Morphologically Structured Entry*		
Example	flagship	crescent	flag	ship	crescent
Log Frequency	.68	.69	*1.49*	*1.95*	.69
Length	8	8	*4*	*4*	8
Syllabicity	2	2	*1*	*1*	2

Same properties

FIORENTINO & POEPPEL 2007

Perspective on Internal Structure	*Whole-word representation*		*Morphologically Structured Entry*		
Example	flagship	crescent	flag	ship	crescent
Log Frequency	.68	.69	*1.49*	*1.95*	.69
Length	8	8	*4*	*4*	8
Syllabicity	2	2	*1*	*1*	2

Different properties

FIORENTINO & POEPPEL 2007

TABLE 1
Samples of visual lexical decision stimuli

Condition	*Mean Log. Freq.**	*Mean no. letters*	*Example*
Compound (CW)	0.451	7.82	flagship
CW 1st / 2nd *constituents*	*1.96/1.98*	*3.82/4.0*	*flag/ship*
Single word (SW)	0.459	7.78	crescent
Nonword (NW)		7.81	nishpern
W-NW Foil (WNW)		7.94	crowskep

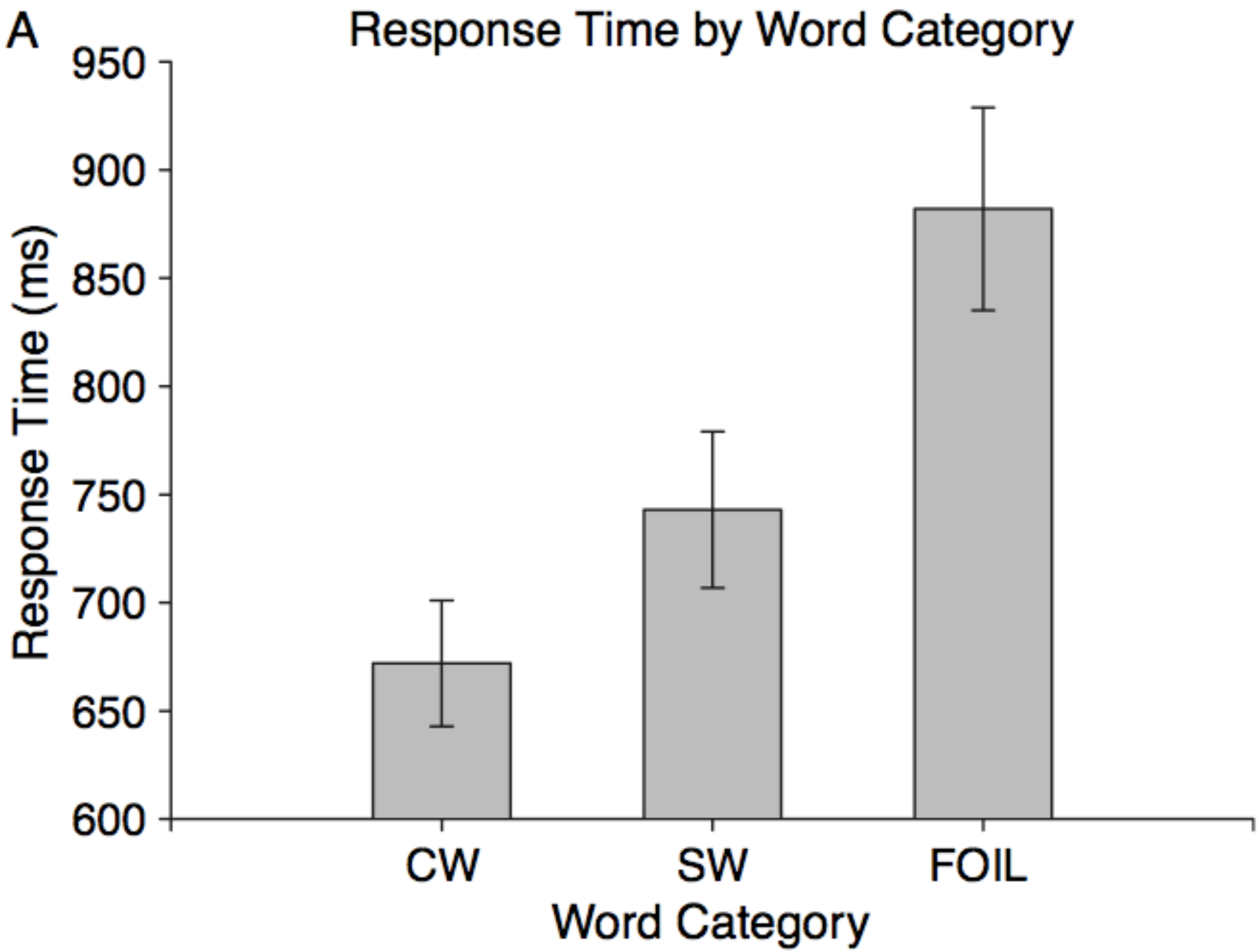
* Parts per million (ppm): CW 2.82 ppm; CW 1st / 2nd constituents 91.2/95.5 ppm; SW 2.88 ppm.

FIORENTINO & POEPPEL 2007

TABLE 2
Response times (mean, *SE*) and accuracy (Pct.) for compounds, single words, word-nonword foils, and other nonwords

Word category	*Mean RT (ms)*	*SE*	*Accuracy (%)*
Compound	672	29	97%
Single word	743	36	90%
Word-nonword foil	882	47	99%
Other nonwords	793	36	99%

FIORENTINO & POEPPEL 2007



FIORENTINO & POEPPEL 2007

TABLE 4
Mean peak latency and SE, for compounds, single words and word-nonword foils

Word category	*Mean latency (ms)*	*SE*
Compound	333	13
Single word	361	16
Word-nonword foils	340	12

FIORENTINO & POEPPEL 2007

TABLE 4
Mean peak latency and SE, for compounds, single words and word-nonword foils

Word category	*Mean latency (ms)*	*SE*
Compound	333	13
Single word	361	16
Word-nonword foils	340	12

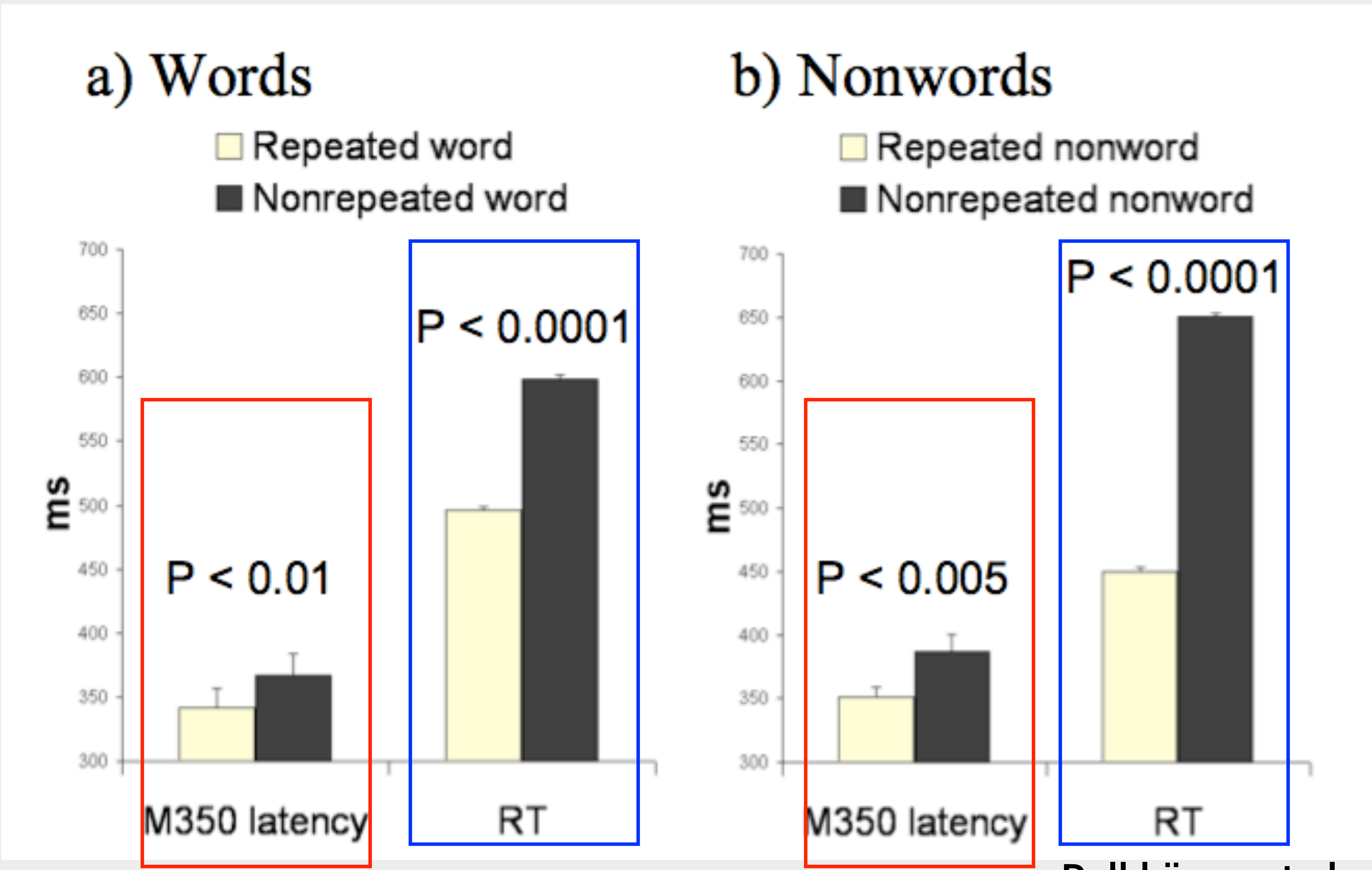
FIORENTINO & POEPPEL 2007

TABLE 4
Mean peak latency and SE, for compounds, single words and word-nonword foils

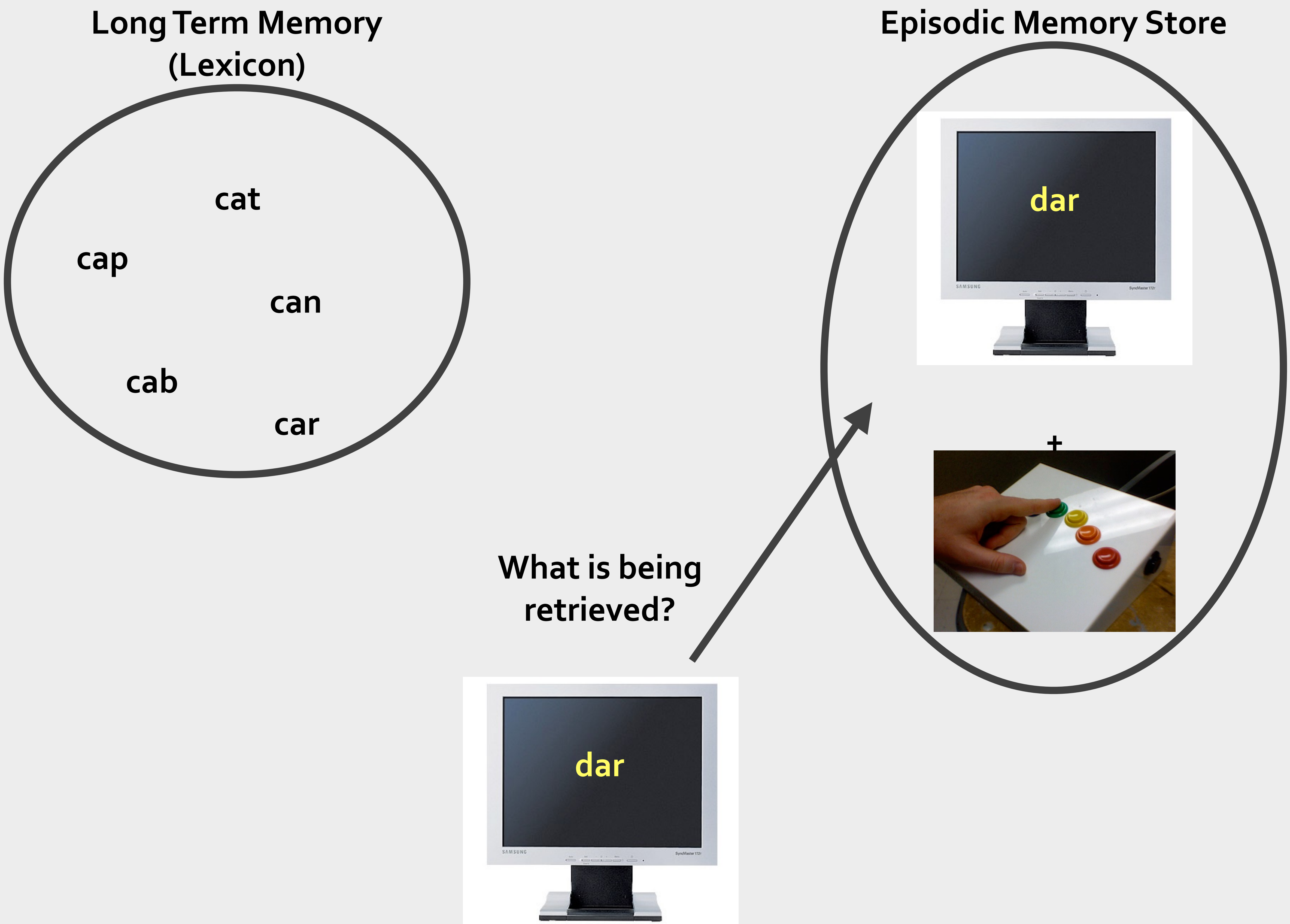
Word category	*Mean latency (ms)*	*SE*
Compound	333	13
Single word	361	16
Word-nonword foils	340	12

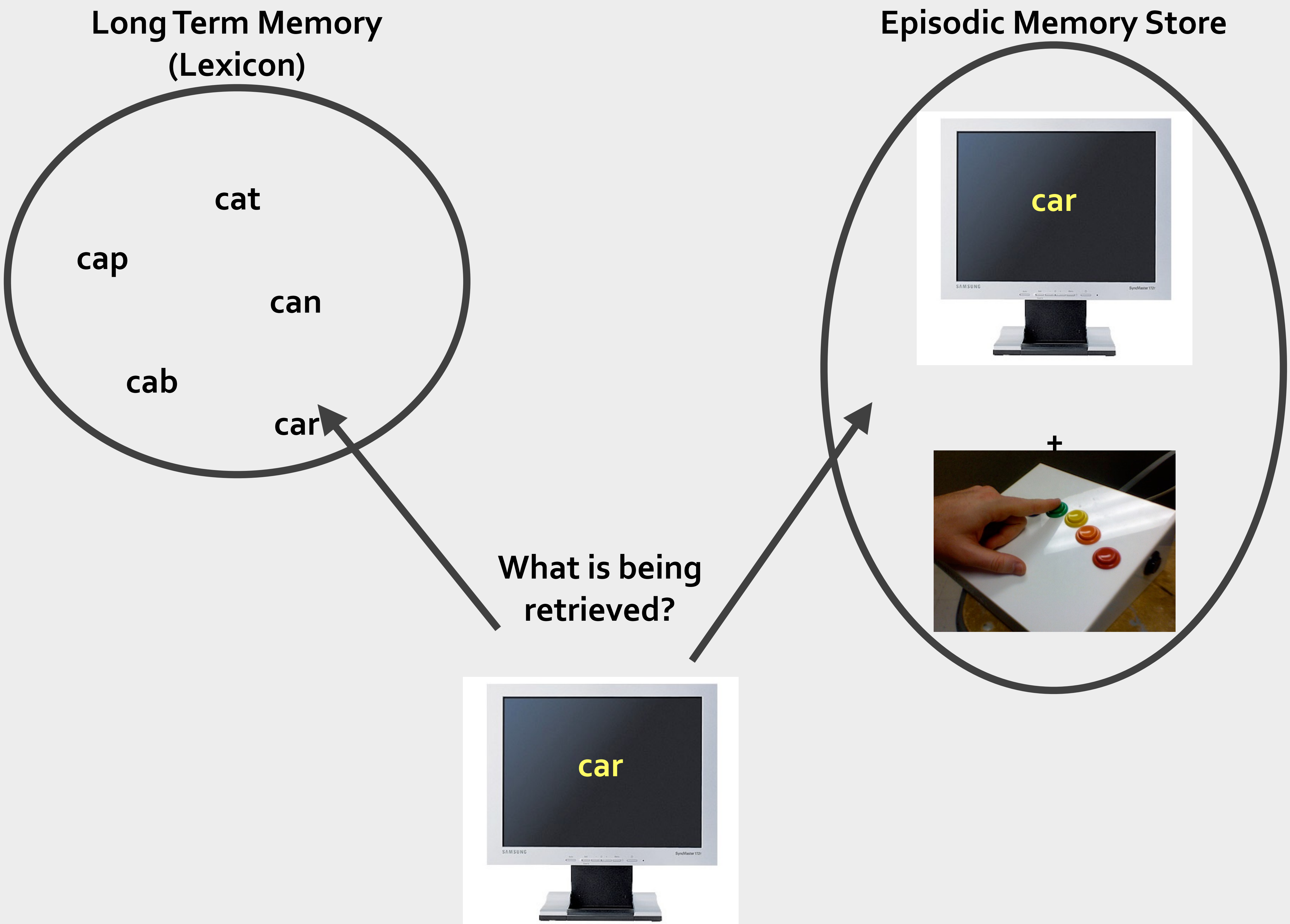
M350 reflects lexical frequency (Embick et al. 2001)
M350 reflects frequency of constituents!
Evidence for early decomposition!

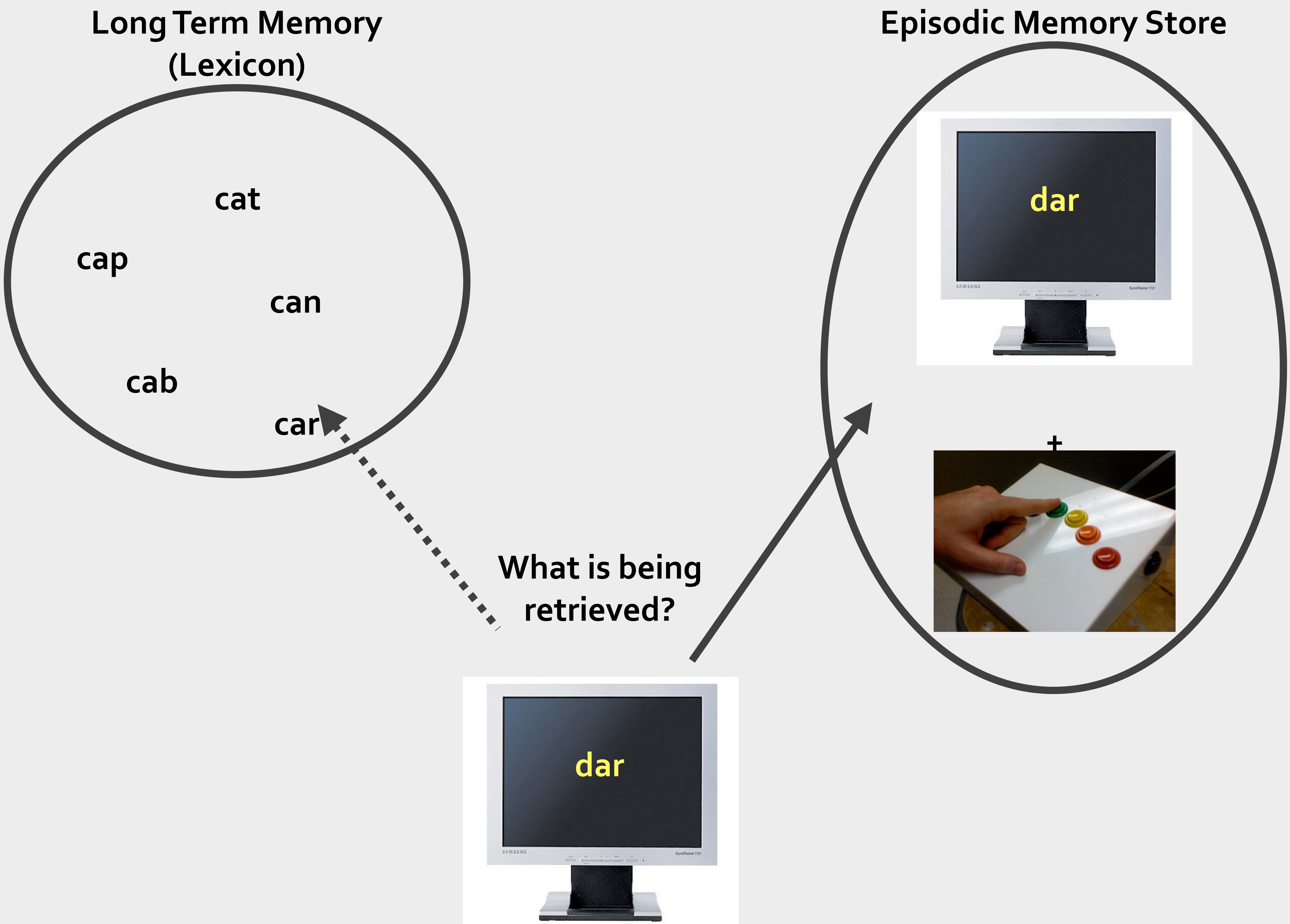
HOWEVER… M350 DOES NOT DISTINGUISH BETWEEN REPETITION OF WORDS AND NONWORDS!



Pylkkänen et al., 2001

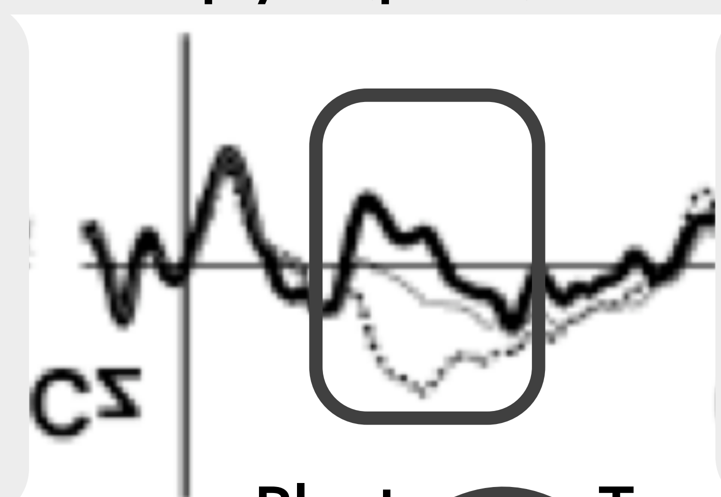

Long Term Memory
(Lexicon)
cat
cap
can
cab
car
Episodic Memory Store
dar
+
What is being
retrieved?
dar


Long Term Memory
(Lexicon)
cat
cap
can
cab
car
Episodic Memory Store
car
+
What is being
retrieved?
car


Long Term Memory
(Lexicon)
cat
cap
can
cab
car
Episodic Memory Store
dar
+
What is being
retrieved?
dar

N400 EVIDENCE THAT NONWORDS ACCESS THE LEXICON

Nonwords that "resemble" real words:
wolm (worm)
plynt (plant)



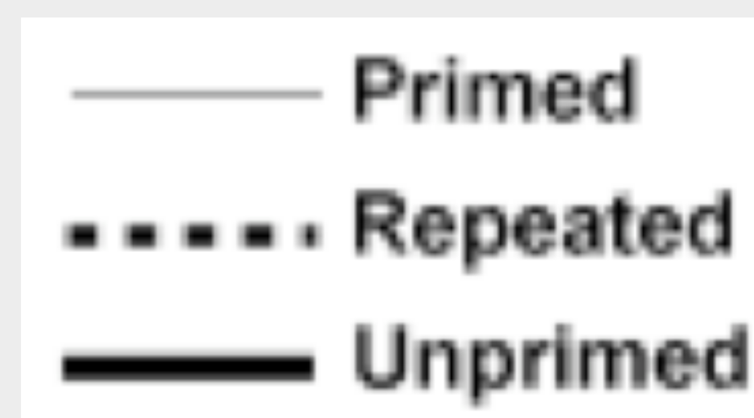
Nonwords DO access the lexicon

Plant **Tree**

Plynt **Tlee**

Wolm **Wolm**

Stairs **Putterfly**

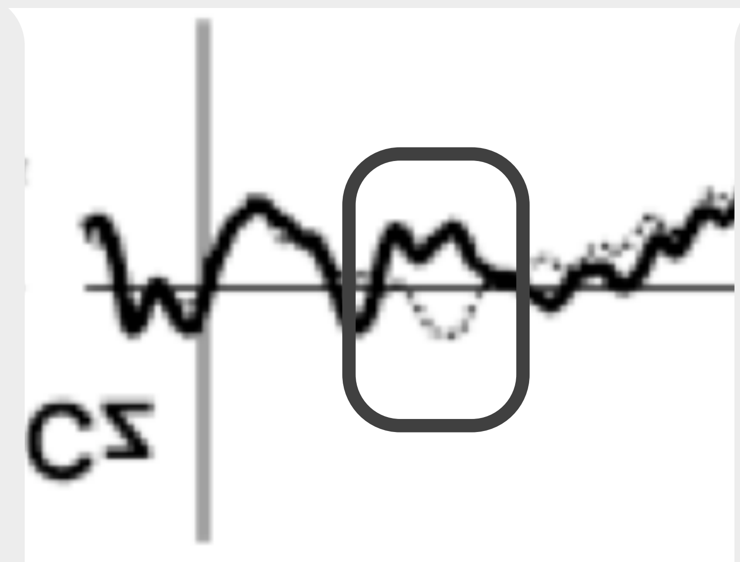


Deacon et al (2004)

BUT THIS IS HARDER TO EXPLAIN…

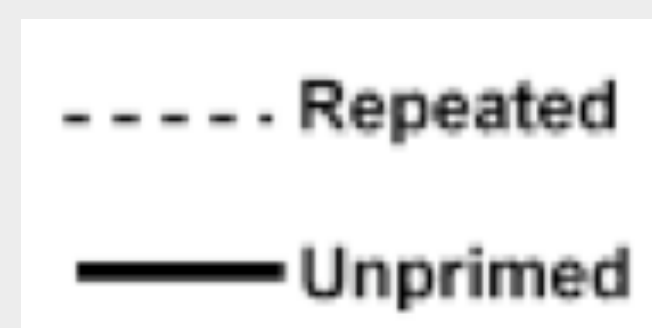
Nonwords that do not "resemble" real words:
loppir
quapt

Supposedly, no lexical access, but still N400 repetition effect



Loppir - Loppir

Tunnel - Quapt



Deacon et al (2004)

WHAT IS GOING ON?

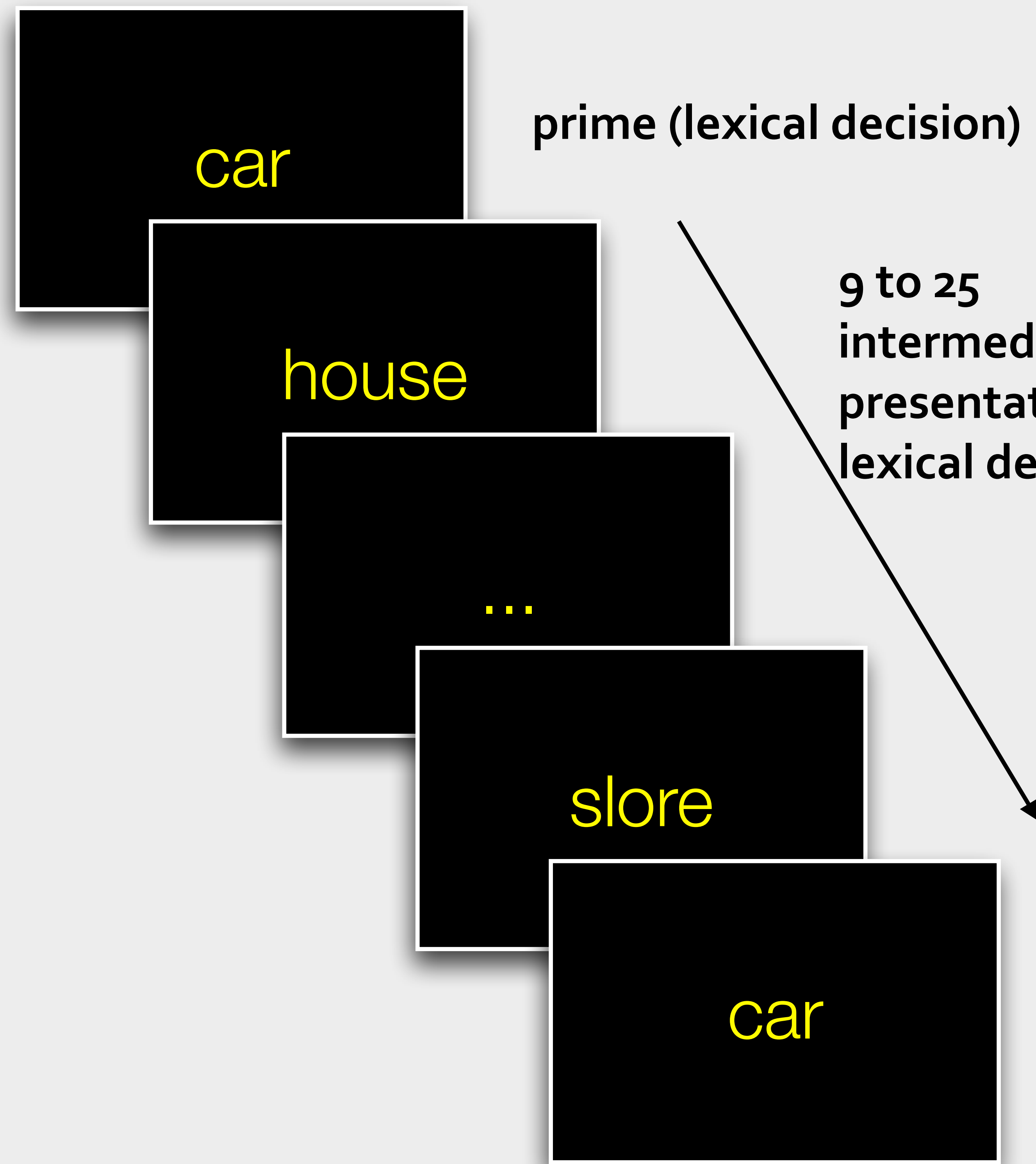
RT, N400 and M350:

- Do not distinguish between repetition of words and repetition of nonwords.

Hypothesis: these measures do NOT reflect lexical access alone

- Earlier response (either P2, N/M170) reflects lexical access processes
- N400/M450 reflects top down predictability of lexico-semantic processes (eg. Kutas 1993, Dambacher et al 2006).

ALMEIDA & POEPPEL, 2013



prime (lexical decision)

9 to 25
intermediate
presentations with
lexical decisions

target (lexical decision)

Words:
high frequency noun list

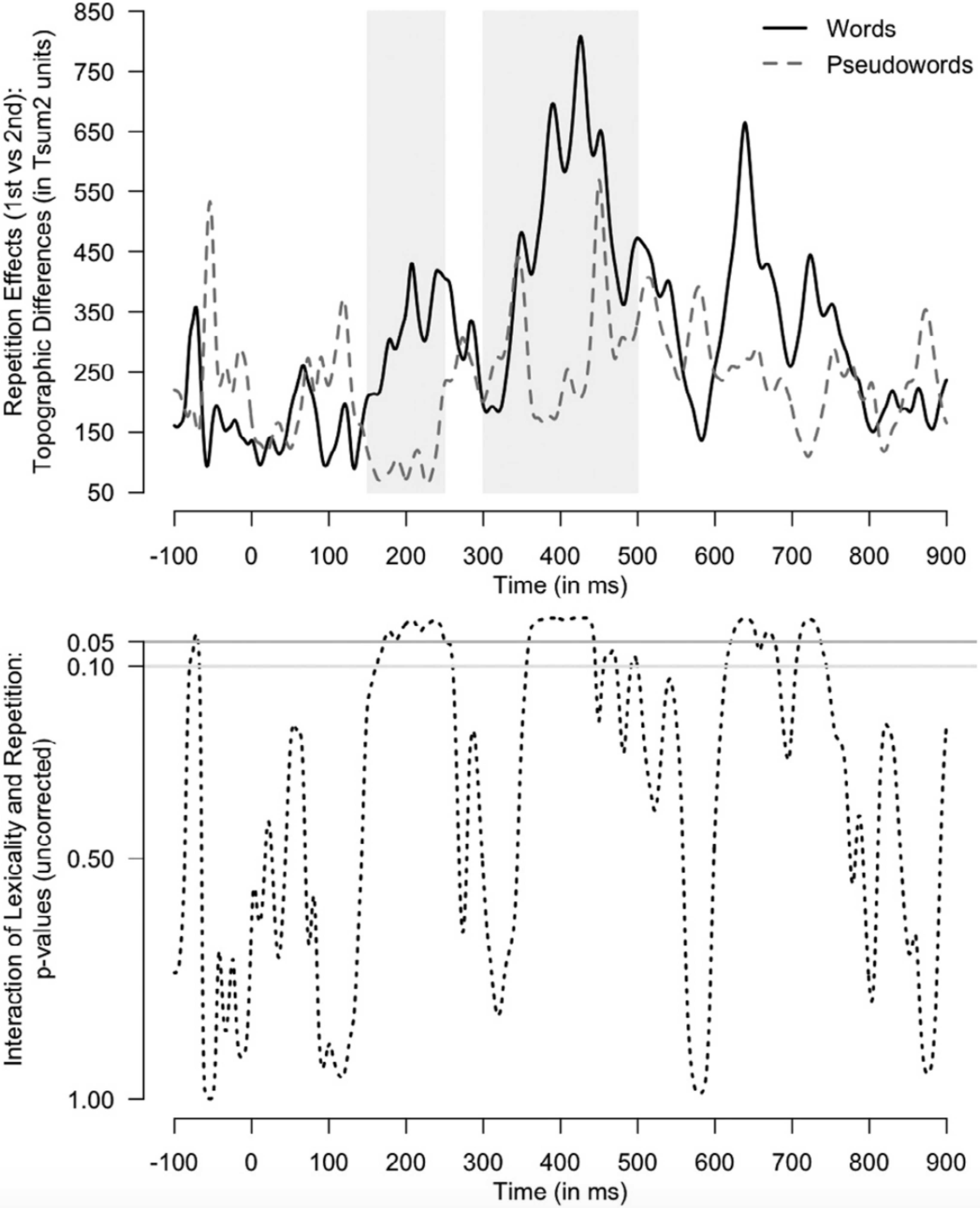
Pseudowords:
phonotactically legal, pronounceable strings

BEHAVIORAL & ERP RESULTS

Table 2
Behavioral results from the lexical decision task. Reaction times are in milliseconds, and standard errors are in parentheses. MOP = magnitude of priming (in ms).

	1st Presentation	2nd Presentation	MOP	Lexicality influence in MOP
Word	675 (33.6)	637 (32.5)	38	−9
Pseudoword	731 (44.1)	684 (32.1)	47	

Almeida & Poeppel, 2013



Almeida & Poeppel, 2013

BEHAVIORAL & ERP RESULTS

Table 2
Behavioral results from the lexical decision task. Reaction times are in milliseconds, and standard errors are in parentheses. MOP = magnitude of priming (in ms).

	1st Presentation	2nd Presentation	MOP	Lexicality influence in MOP
Word	675 (33.6)	637 (32.5)	38	−9
Pseudoword	731 (44.1)	684 (32.1)	47	

	150–250 ms Window	300–500 ms Window
Interaction of lexicality and order of presentation		
$Tsum2_{words}–Tsum2_{pseudowords}$	**337 (0.0001)**	**257 (0.0001)**
Repetition effect for words		
Tsum2	**406 (0.0043)**	**617 (0.0018)**
Topographic dissimilarity	**0.017 (0.0032)**	**0.07 (0.0005)**
Repetition effect for nonwords		
Tsum2	69 (0.9875)	**360 (0.0257)**
Topographic dissimilarity	0.0047 (0.9895)	0.02 (0.2944)

Almeida & Poeppel, 2013

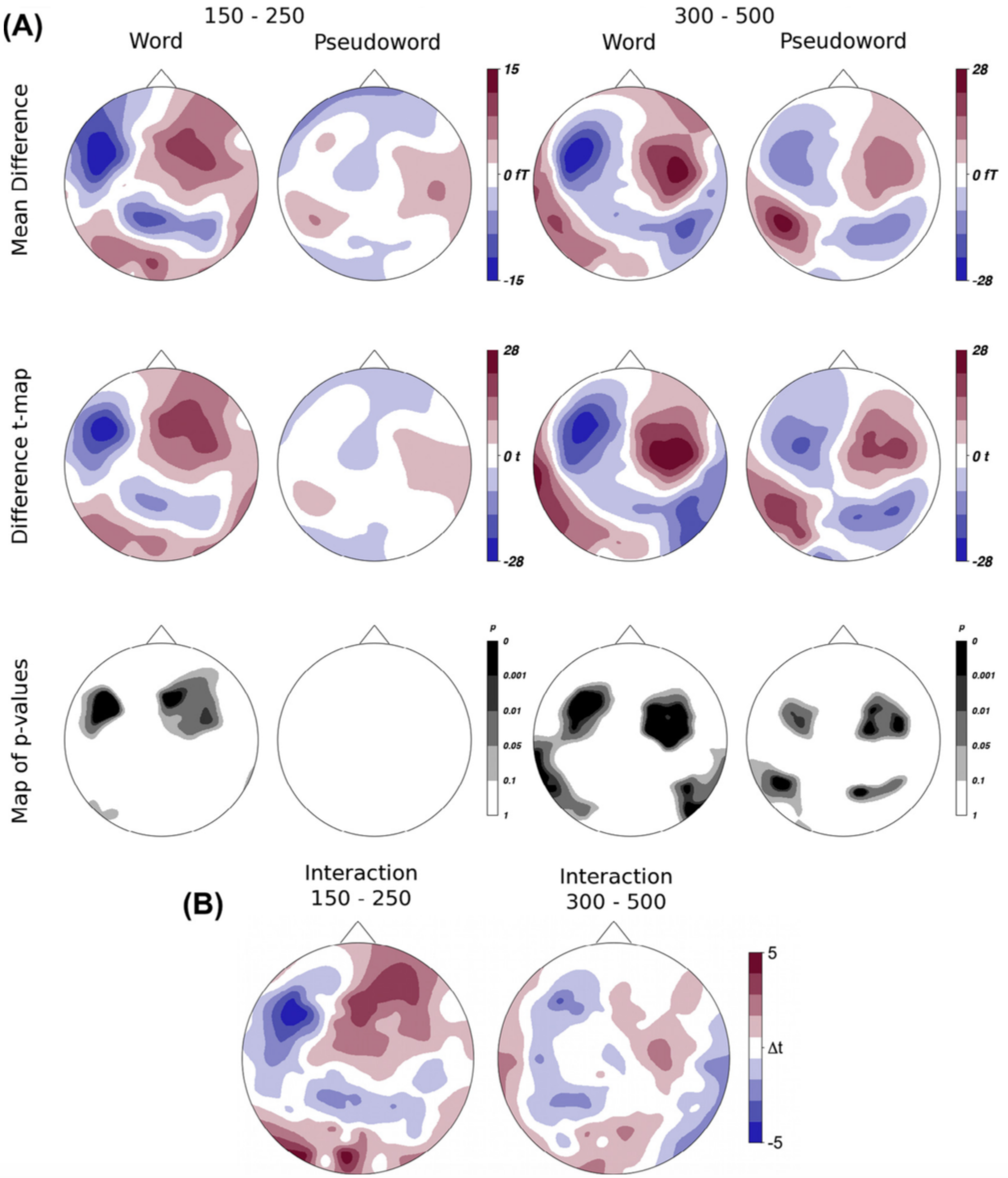
BEHAVIORAL & ERP RESULTS

Table 2
Behavioral results from the lexical decision task. Reaction times are in milliseconds, and standard errors are in parentheses. MOP = magnitude of priming (in ms).

	1st Presentation	2nd Presentation	MOP	Lexicality influence in MOP
Word	675 (33.6)	637 (32.5)	38	−9
Pseudoword	731 (44.1)	684 (32.1)	47	

	150–250 ms Window	300–500 ms Window
Interaction of lexicality and order of presentation		
$Tsum2_{words} – Tsum2_{pseudowords}$	**337 (0.0001)**	**257 (0.0001)**
Repetition effect for words		
Tsum2	**406 (0.0043)**	**617 (0.0018)**
Topographic dissimilarity	**0.017 (0.0032)**	**0.07 (0.0005)**
Repetition effect for nonwords		
Tsum2	69 (0.9875)	**360 (0.0257)**
Topographic dissimilarity	0.0047 (0.9895)	0.02 (0.2944)

Almeida & Poeppel, 2013



Almeida & Poeppel, 2013

CONCLUSIONS

Early bilateral anterior response (P2m)

- Tracks solely the repetition of lexical items
- Independent of predictability (see also Misra et al 2003)

Converging evidence that we access the orthographic-phonological form of words by 200 ms post-stimulus onset.

CONCLUSIONS

N400/M350

- Tracks repetition of lexical items AND pseudowords
- But the underlying circuits are *not* identical
  - **Words:** Changes in the configuration of the underlying neural sources
  - **Pseudowords:** Evidence for simple strength (not configuration) difference in the same underlying neural sources

WHAT'S IN A WORD?

Major controversy: Storage vs computation in the lexicon

- Everyone agrees that there is storage
- Disagreement is about the amount of *computation/parsing* in the lexicon
- Two poles of the debate
  - **Full decomposition (but neuro data needs more work)**
  - Full storage

VISUAL WORD RECOGNITION AND LANGUAGE

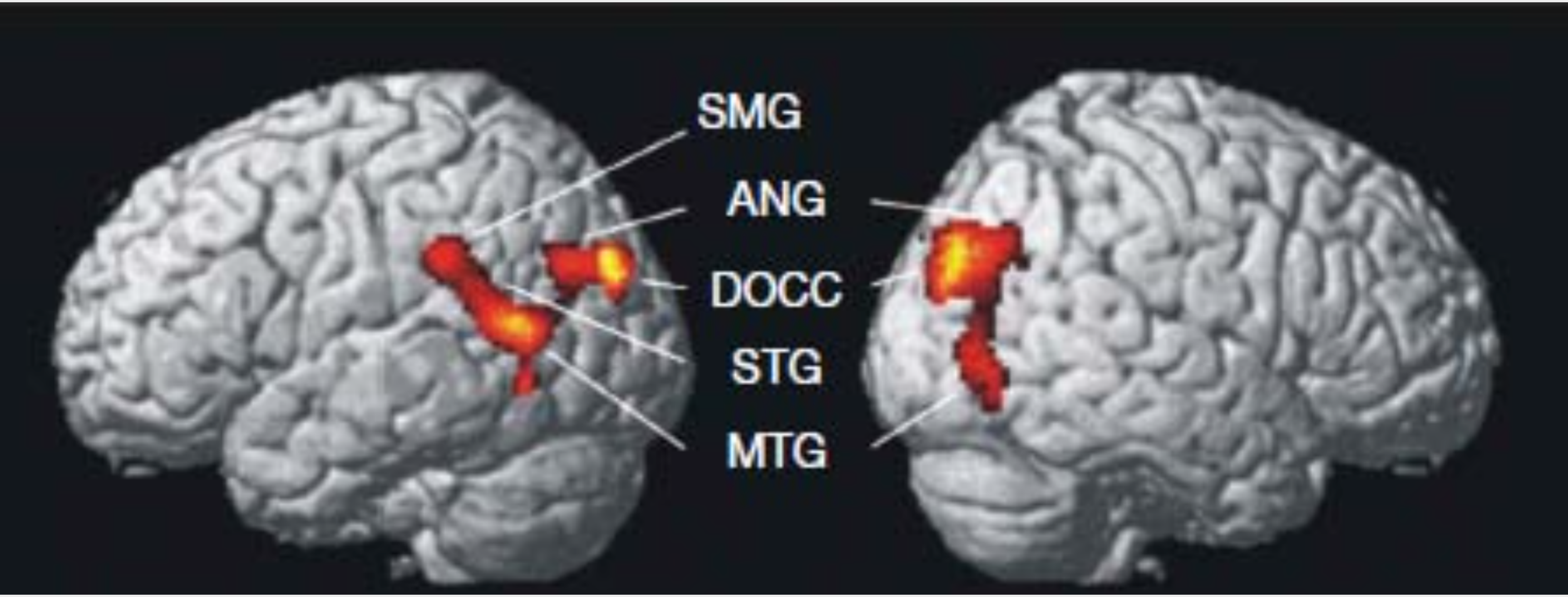
VISUAL WORD RECOGNITION AND LANGUAGE

What does visual word recognition actually tell us about language?

- There ARE important differences between visual word recognition and auditory word recognition
- Written language bias
  - Reading/Writing is historically weird for humans
- We are studying the interaction between language and a sensory system!

VISUAL WORD RECOGNITION AND THE VISUAL SYSTEM

READING CHANGES THE BRAIN



Carreiras et al. (2009)

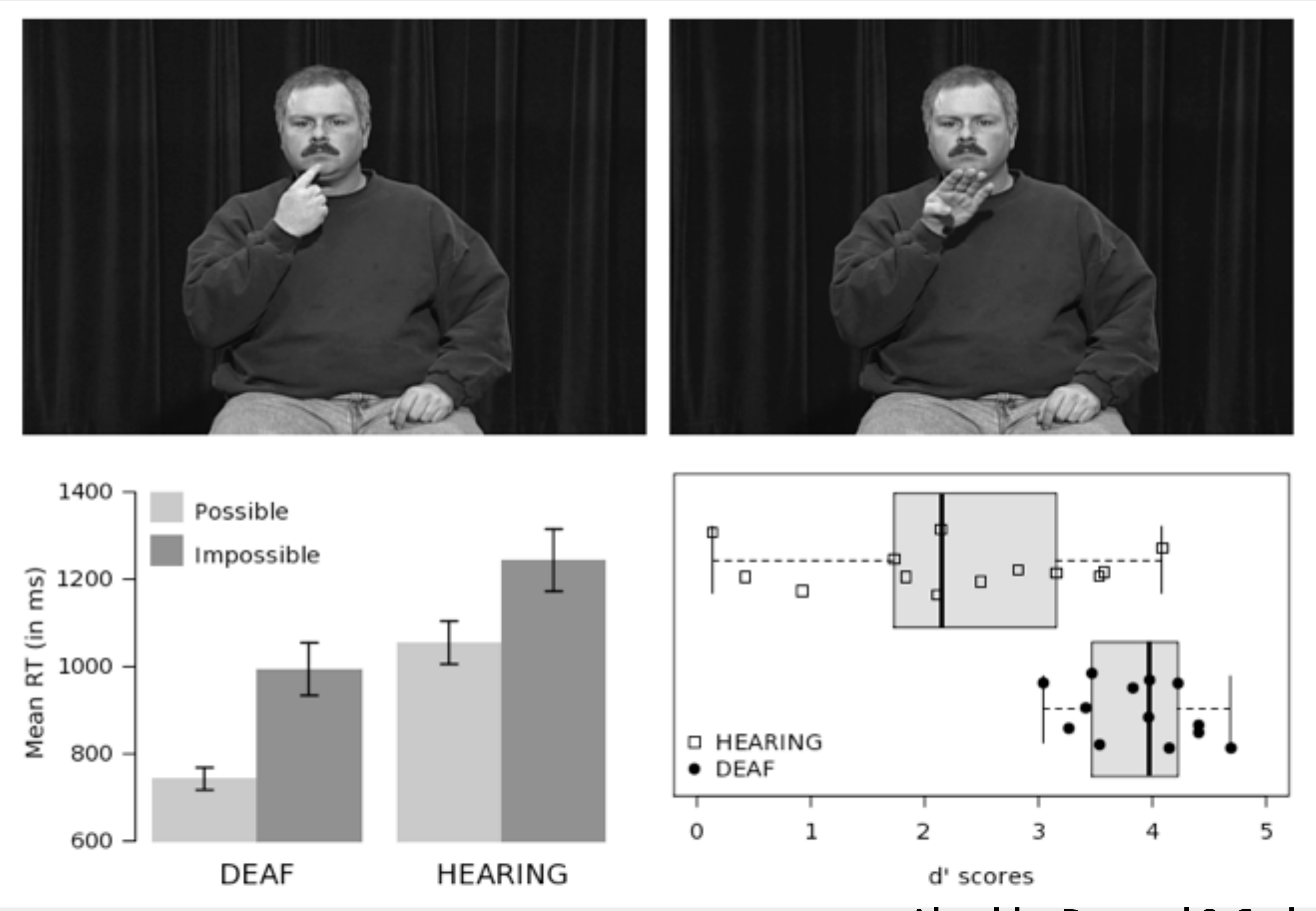
VISUAL SYSTEMS CAN BE FINE-TUNED TO PROCESS HIGHLY SPECIFIC STIMULI

possible gesture
to miss (someone)

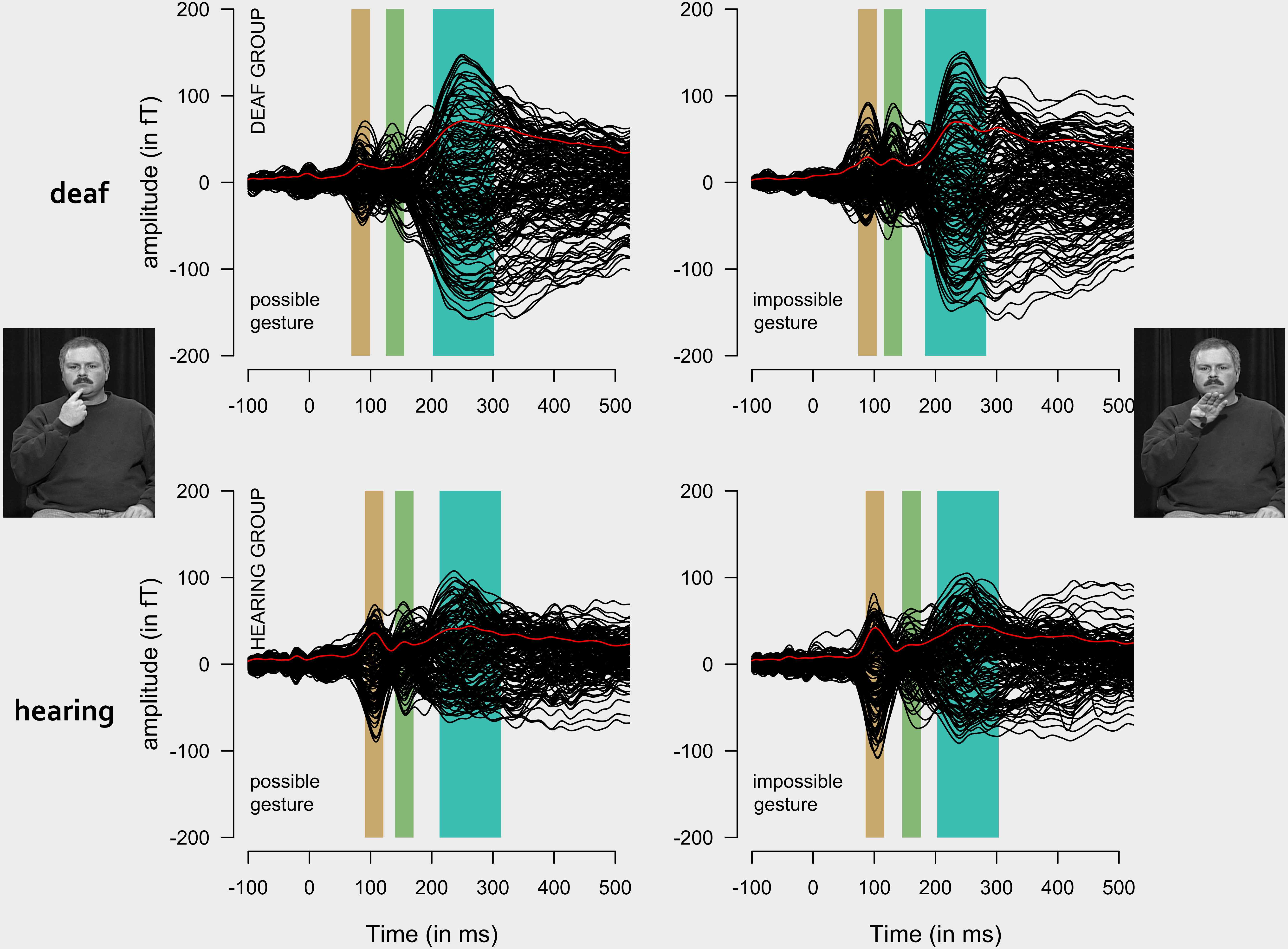
impossible gesture
nonword

Almeida, Poeppel & Corina (2015)

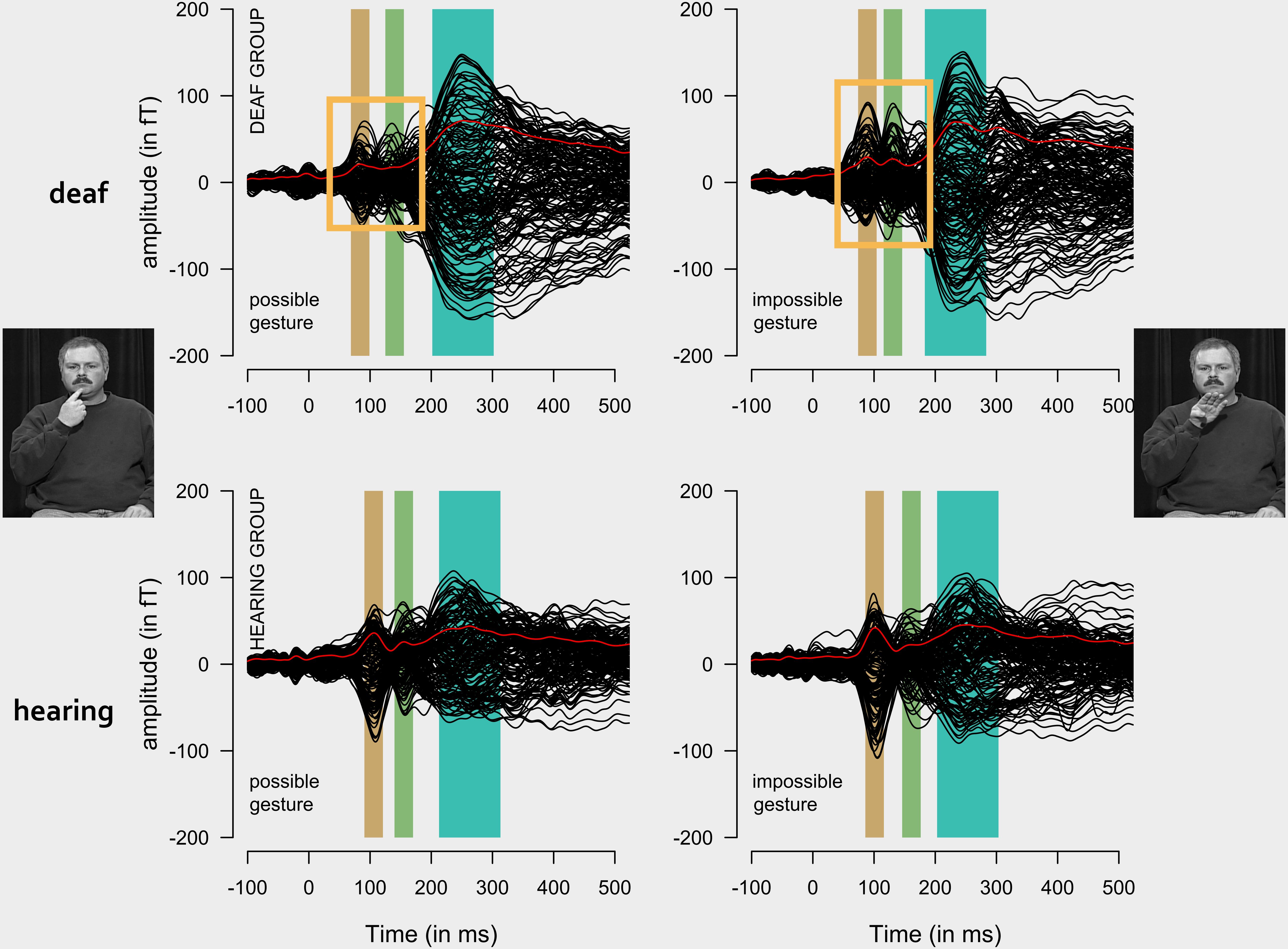
VISUAL SYSTEMS CAN BE FINE-TUNED TO PROCESS HIGHLY SPECIFIC STIMULI



Almeida, Poeppel & Corina (2015)



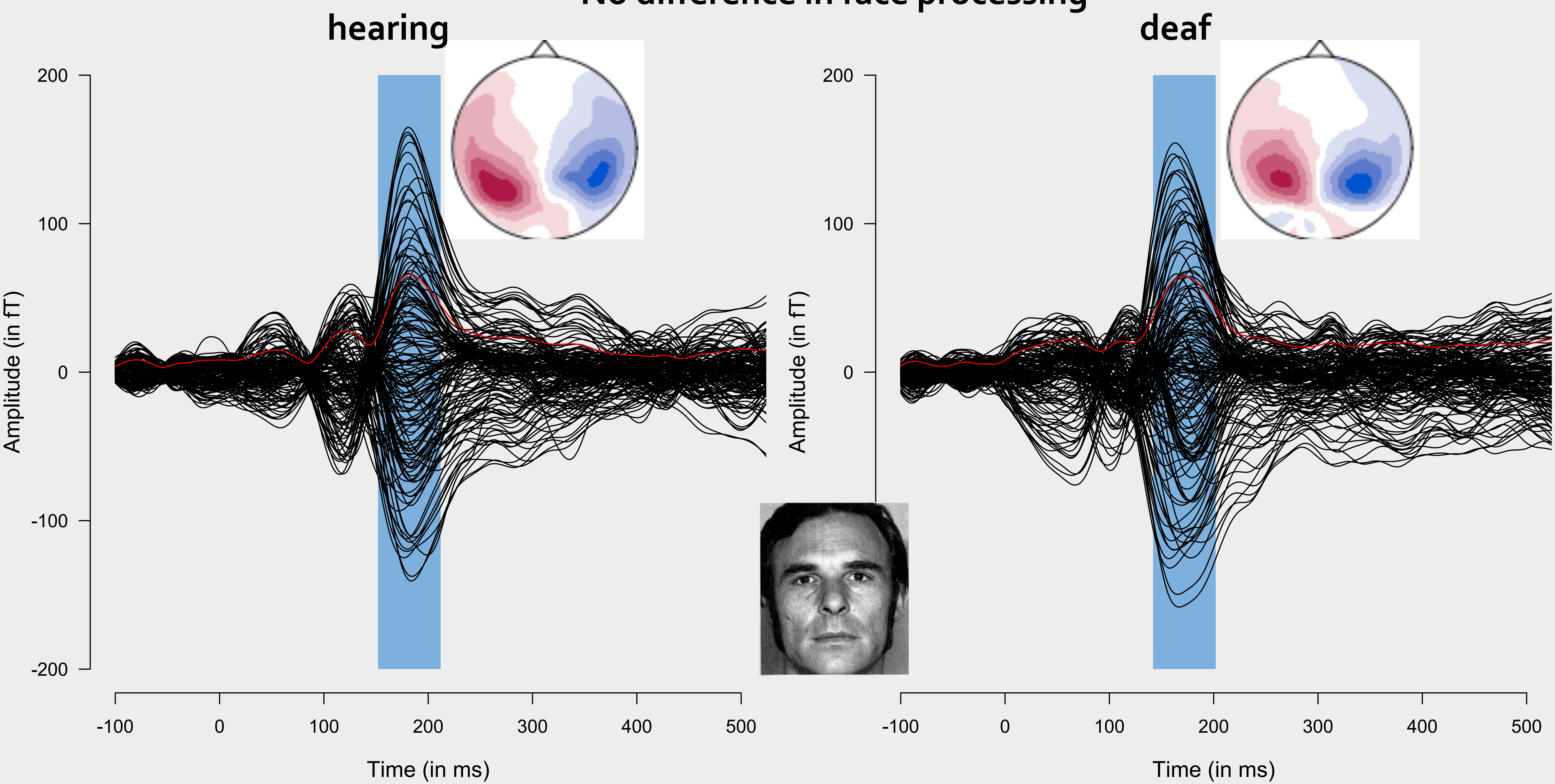
deaf
hearing
DEAF GROUP
HEARING GROUP
amplitude (in fT)
possible gesture
impossible gesture
Time (in ms)
200
100
0
-100
-200
-100
0
100
200
300
400
500



deaf
hearing
amplitude (in fT)
DEAF GROUP
HEARING GROUP
possible
gesture
impossible
gesture
200
100
0
-100
-200
-100
0
100
200
300
400
500
Time (in ms)

VISUAL SYSTEMS CAN BE FINE-TUNED TO PROCESS HIGHLY SPECIFIC STIMULI

No difference in face processing



Almeida, Poeppel & Corina (2015)

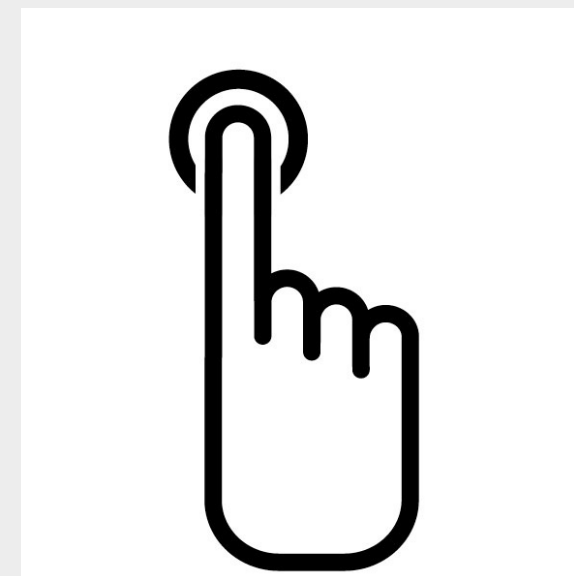
CONCLUSIONS

Early (100 and 130 ms) brain responses distinguish between anatomically possible vs anatomically impossible gestures

- but only in the deaf participants
- at 100 ms, there is a difference in strength of the sources
- at 130 ms, there is already a difference in configuration of the sources
- Probably feedforward model is different (because of the early latency of the responses)

CONCLUSIONS

CONCLUSIONS



Reading Experience
or Extensive exposure
to ASL lexicon

different
forward
models

Early visual processing

(Some) top-down predictions

Output procedures

time

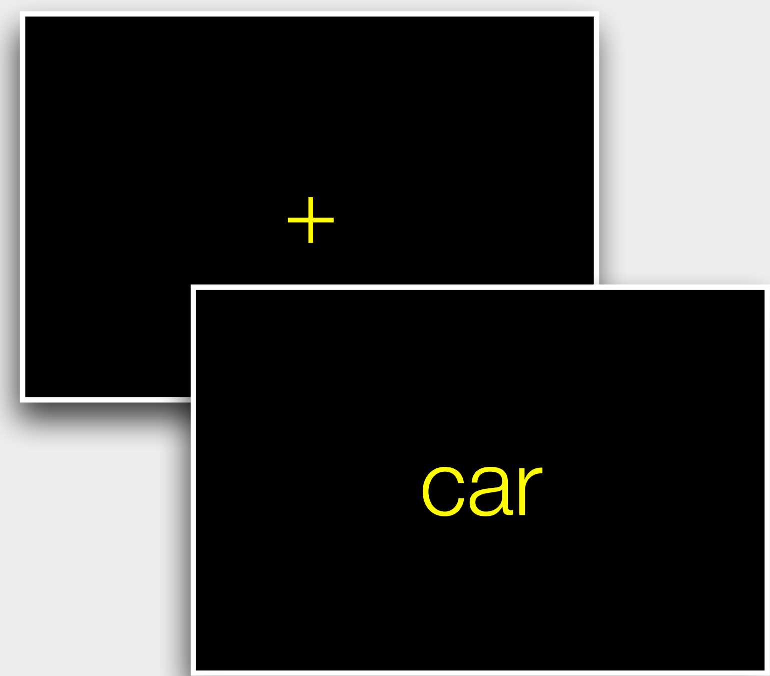
word presentation: 0 ms

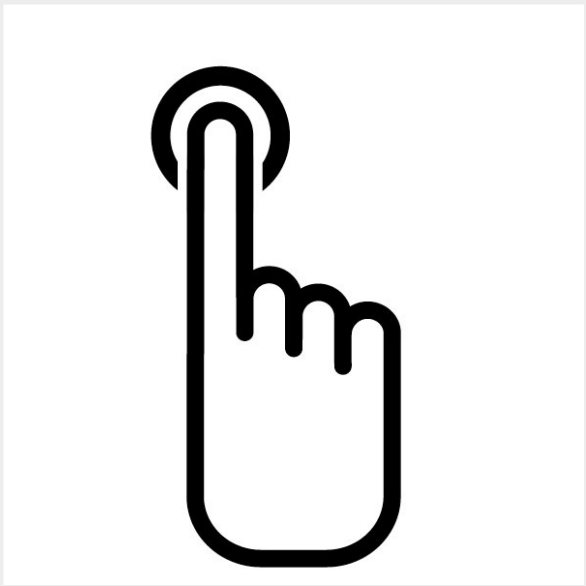
~170ms

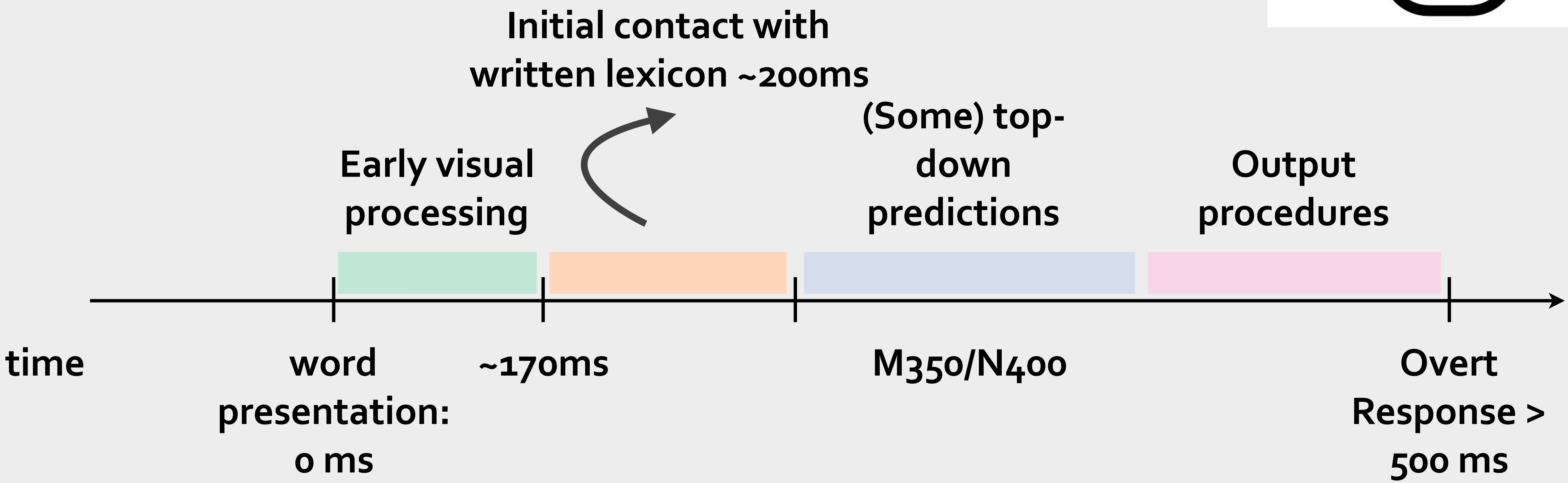
M350/N400

Overt Response > 500 ms

CONCLUSIONS







CONCLUSIONS

Beware written language bias

- Reading is not the same as hearing
- Visual word recognition: interaction between the visual and the language system
  - reading and sign language!

Reading *can* be informative about language!

- Just be careful when generalizing to language as a whole!

COLLABORATORS

David Poeppel (Max Planck/NYU) & David Corina (UCDavis)

Ali Idrissi (Qatar University)

Kevin Schluter (University of Minnesota)

Matt Tucker (Oakland University, now at Amazon)

Souaad Al Helou (Ruhr University Bochum)

Kefei “Fisher” Wu & Rosy Tahan (NYUAD ’16 and ’17)

THANK YOU!
ANY QUESTIONS?