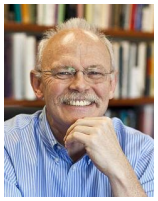
Modeling the N400 brain potential as change in a probabilistic representation of meaning

Milena Rabovsky

May 28, 2020

1

Thanks to



Jay McClelland

Steven Hansen

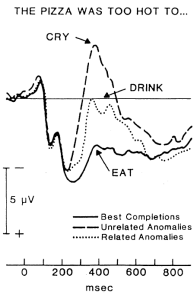
2

The N400 component of the ERP

Modulating variables
- Semantic violations,
- Contextual fit,
- Frequency,…
- ➢ > 1000 studies
- ➢ Meaning processing

➢ Functional basis?
- Lexical access (Lau et al., 2008)
- Semantic inhibition? (Debruille, 2007)
- Semantic integration? (Baggio & Hagoort, 2011)

➢ Relate to computational model



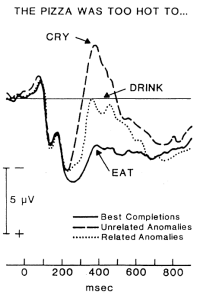
3

The N400 component of the ERP

Our account:

➢ Change of a representation of meaning that implicitly and probabilistically represents all aspects of meaning of the event described by a sentence

➢ Change in conditional probabilities of semantic features



4

Model based approaches

1) Models of brain processes
   - ➢ Neurobiologically more plausible
   - ➢ Modeling the ERP waveform (Cheyette & Plaut, 2017; Laszlo & Armstrong, 2014; Laszlo & Plaut, 2012)

2) Functional-level models of cognitive processes
   - ➢ Covariation between N400 amplitudes and model measures
   - ➢ Leaving aside physiological details (Brouwer, Crocker, Venhuizen, & Hoeks, 2017; Frank et al., 2015; Rabovsky, Hansen, & McClelland, 2018; Rabovsky & McRae, 2014)

5

Previous functional level models of N400 amplitudes

➢ Network error in a model of word meaning (Rabovsky & McRae, 2014)
   - ➢ conceptualized as implicit prediction error (McClelland, 1994)
     model generated activation = implicit prediction
     correct target activation = observation

➢ Network error in simple recurrent network model (SRN)
   → surprisal (Frank et al., 2015)

➢ Change of lexical activation in model assuming two steps
   lexical retrieval = N400
   semantic integration = P600
   (Brouwer, Crocker, Venhuizen, & Hoeks, 2017)

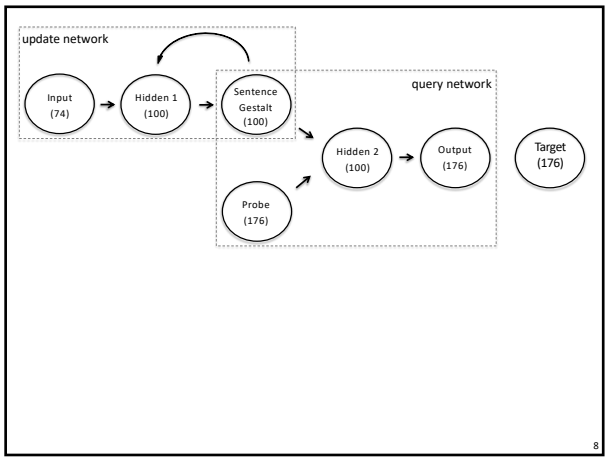
6

The Sentence Gestalt model

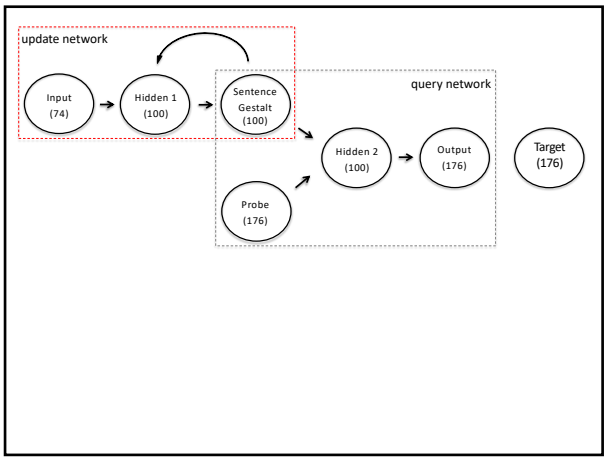
(based on McClelland, St. John, & Taraban, 1989; St. John & McClelland, 1990)

- Words as „cues to meaning“ (Rumelhart, 1979) that change the representation of sentence meaning
- Task of sentence processing:
  - Process sequences of words
  - Answer question concerning described event
- Learn representation to answer questions

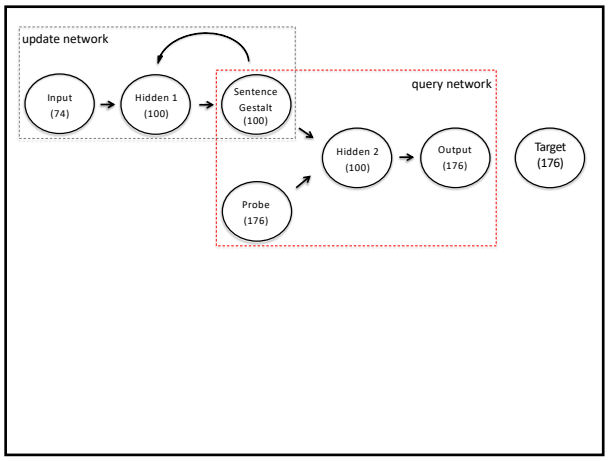
7



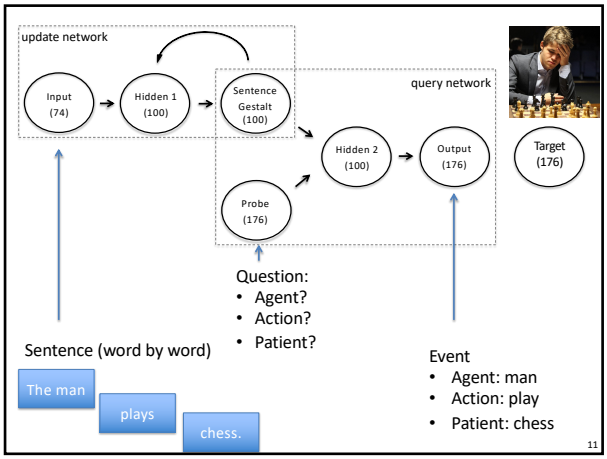
8



9



10

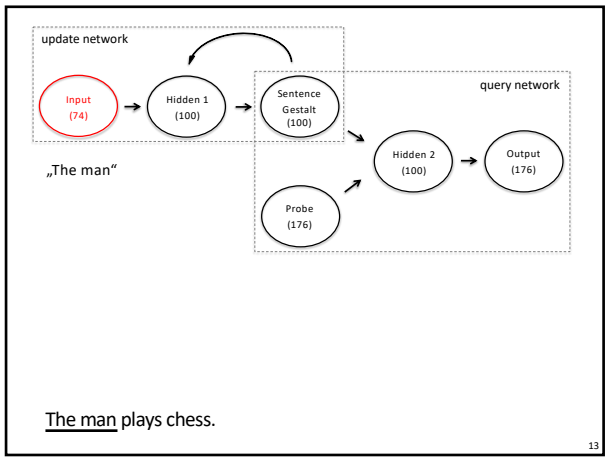


11

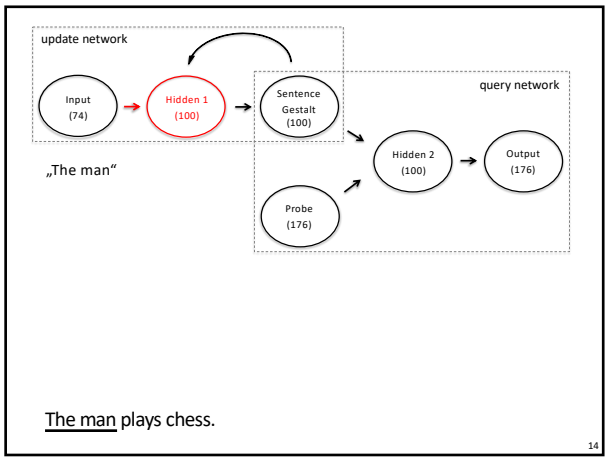
Learning

- Goal of training: Activation of each feature unit corresponds to the conditional probability of that feature in that situation (Rumelhart et al., 1995)
  - In ideally trained model, change in activation induced by each incoming word would represent change in the probabilities of semantic features induced by that word
  - also – implicitly – at the Sentence Gestalt layer
- Trained model updates with each incoming word an internal representation that probabilistically represents all aspects of meaning of the described event

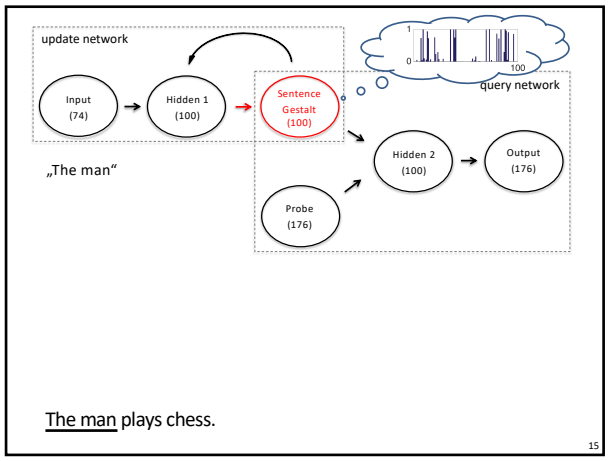
12



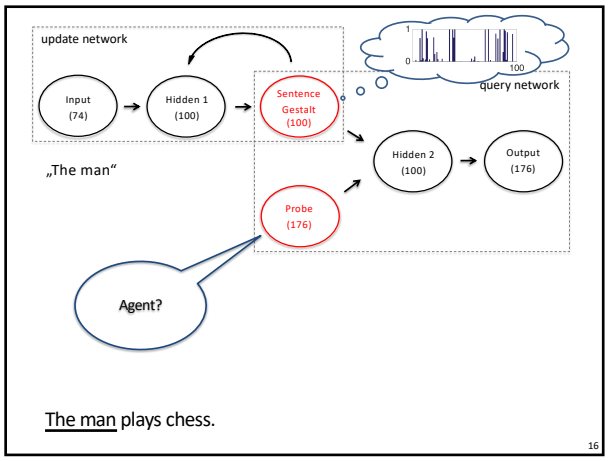
13



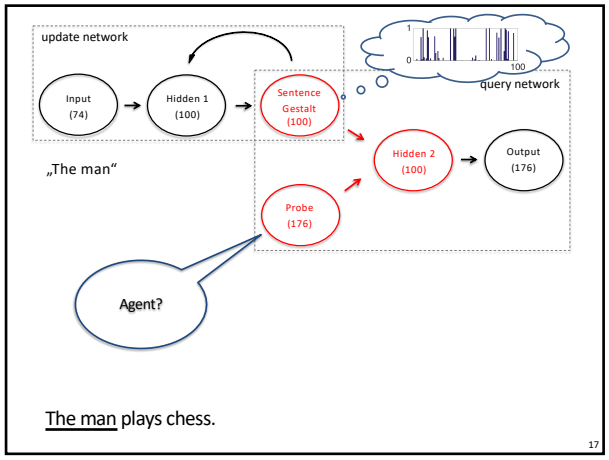
14



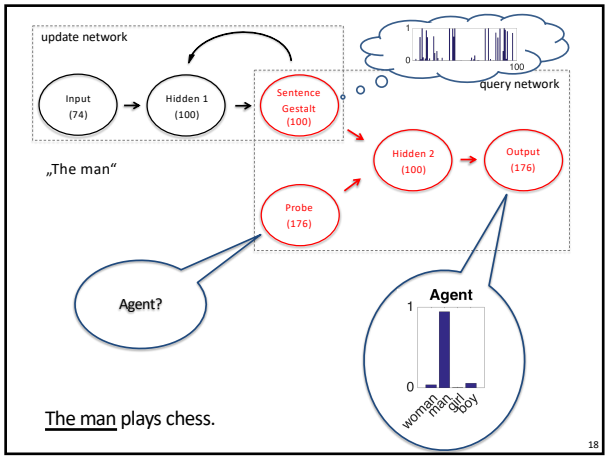
15



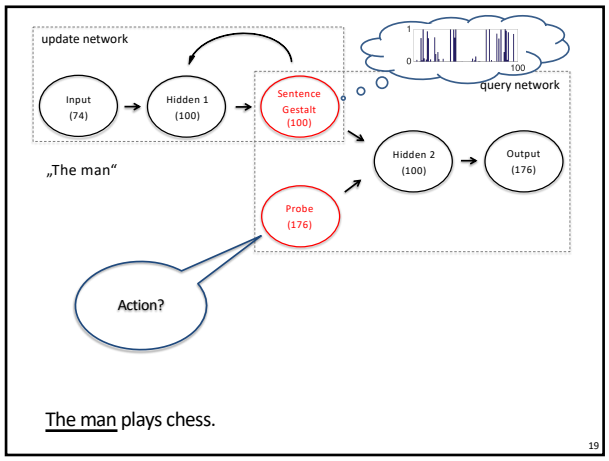
16



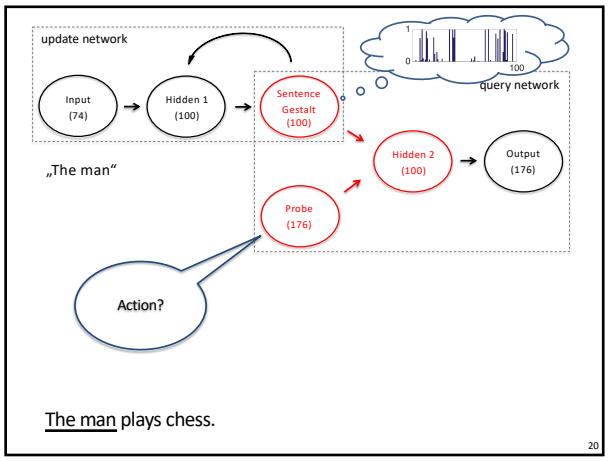
17



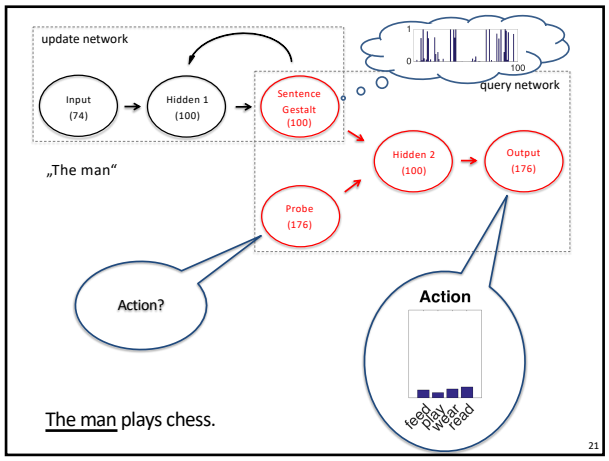
18



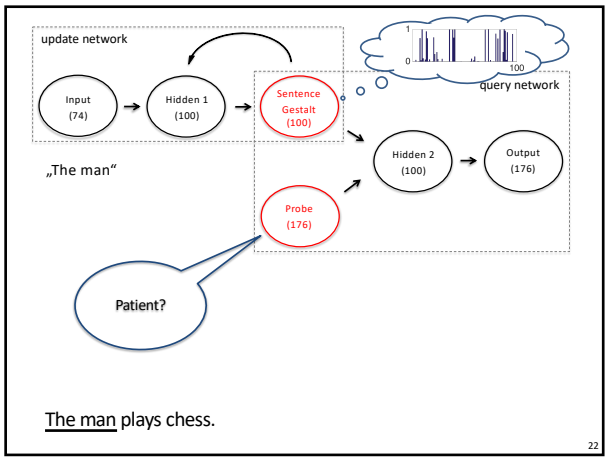
19



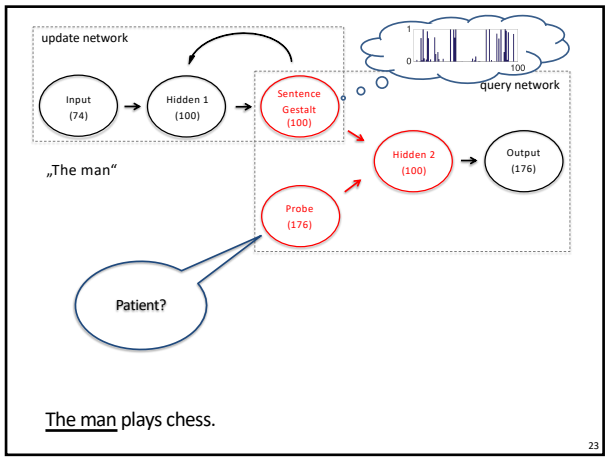
20



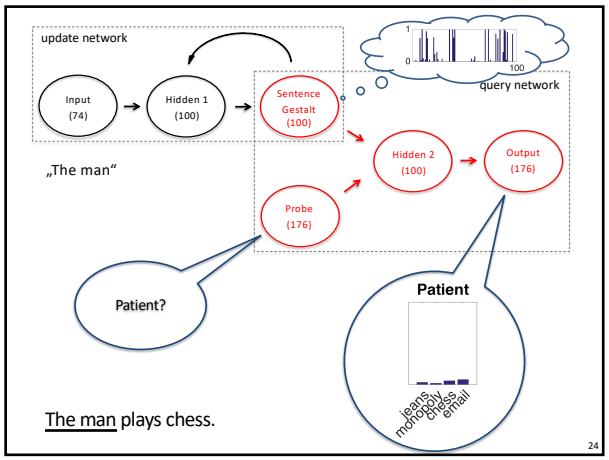
21



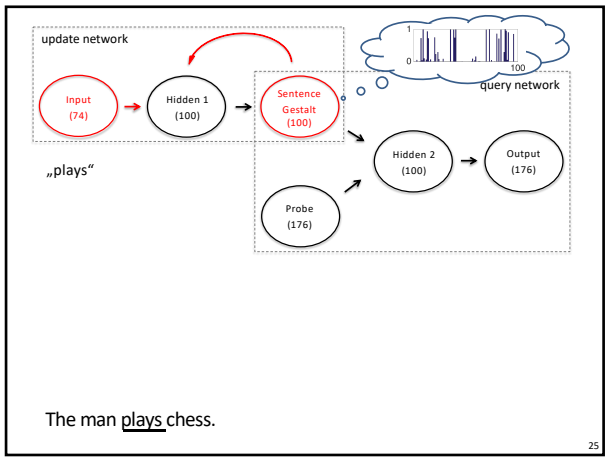
22



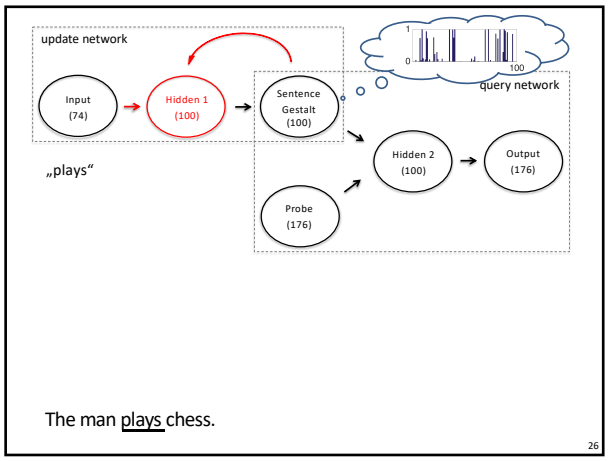
23



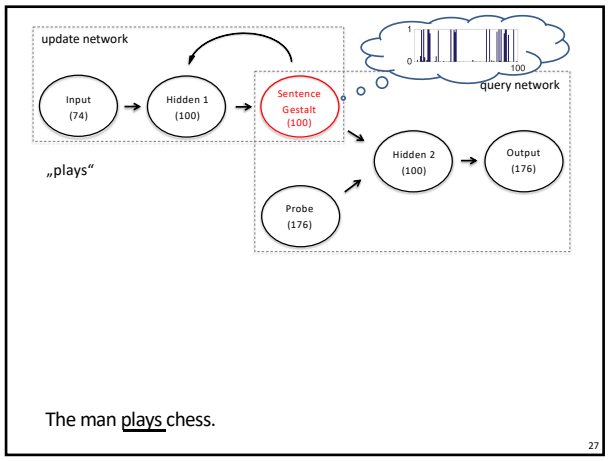
24



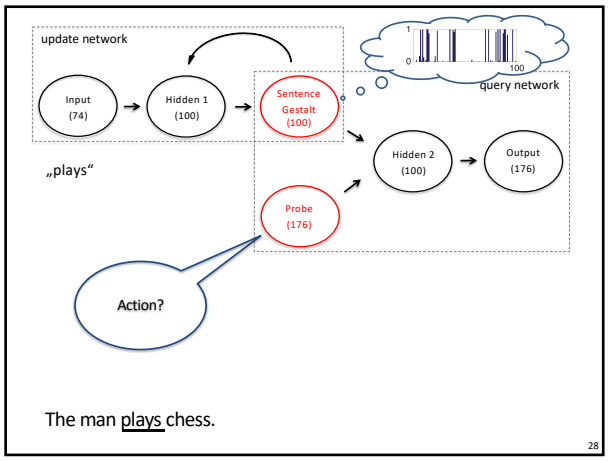
25



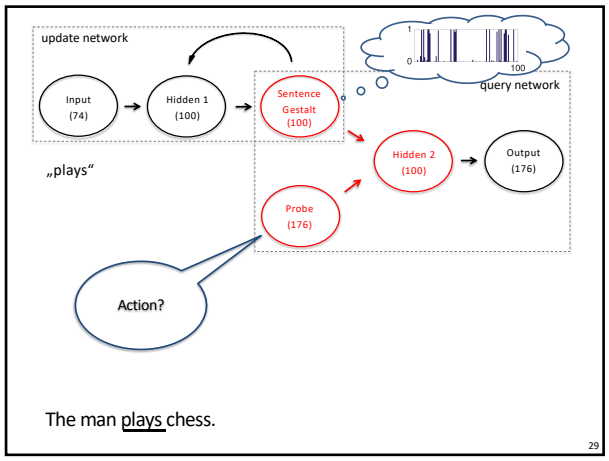
26



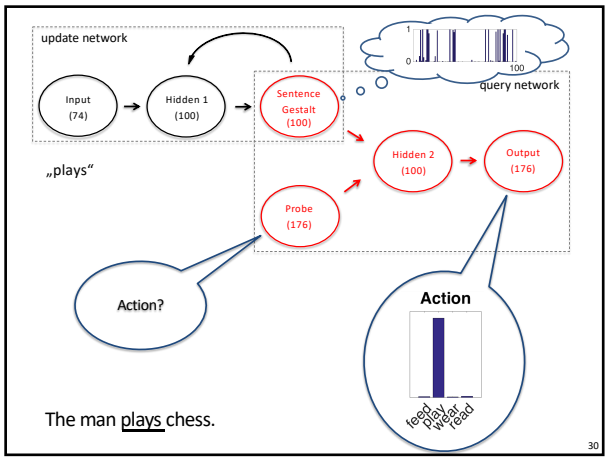
27



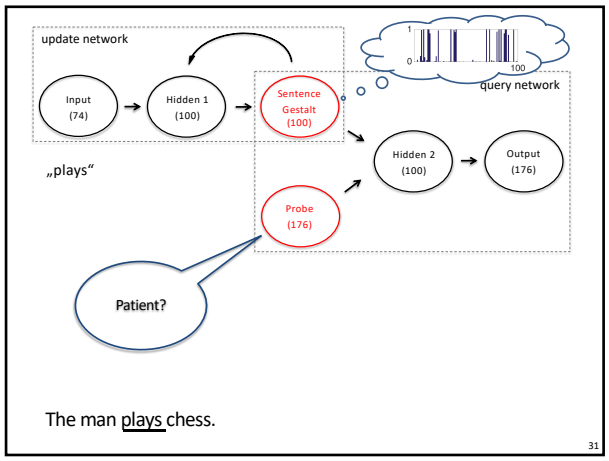
28



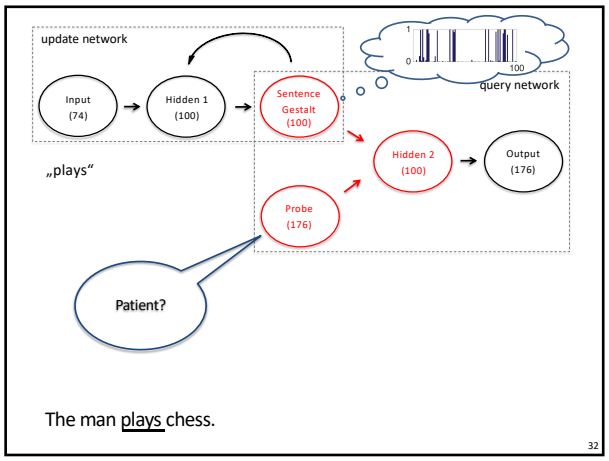
29



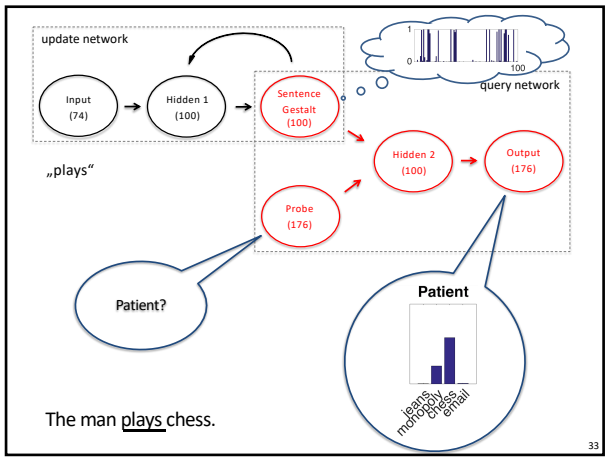
30



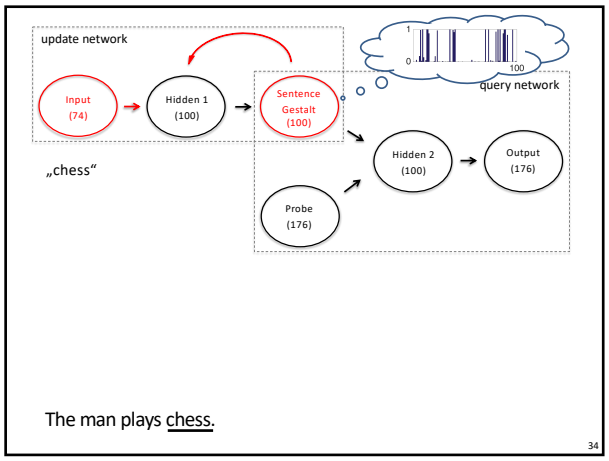
31



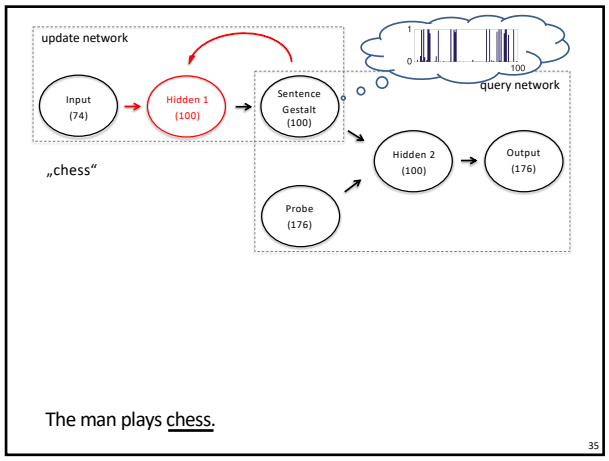
32



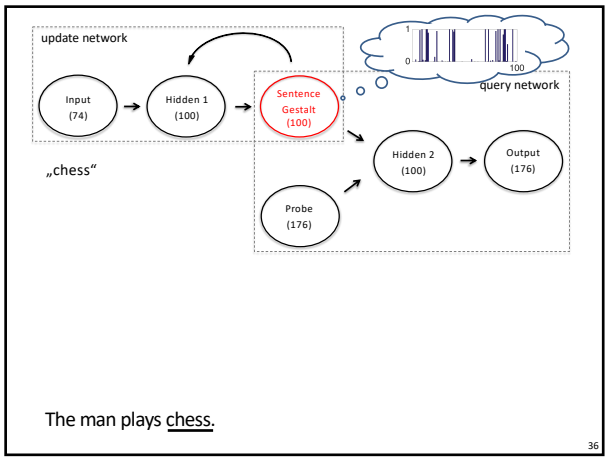
33



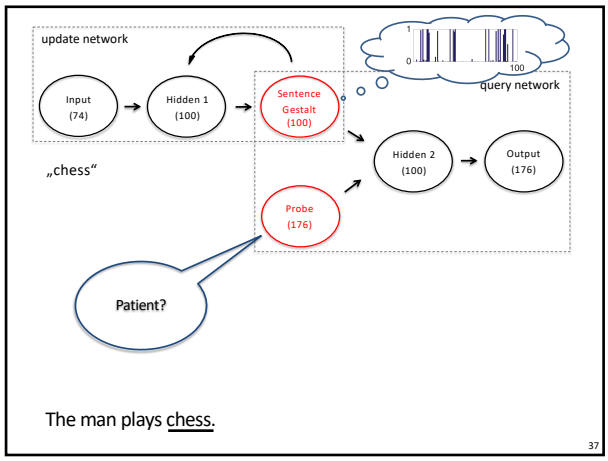
34



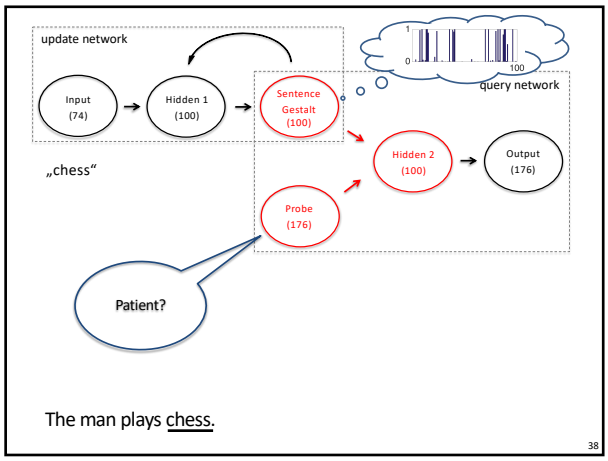
35



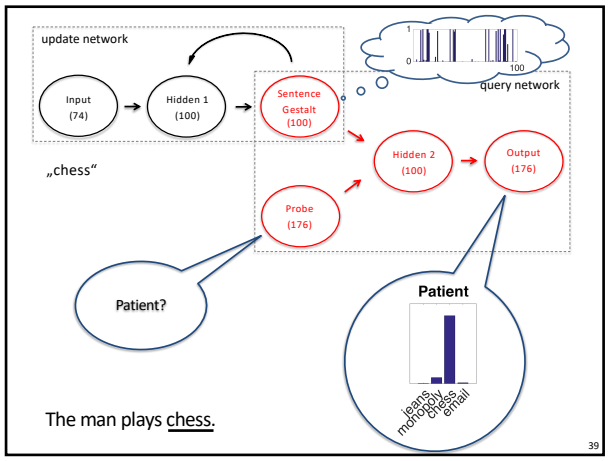
36



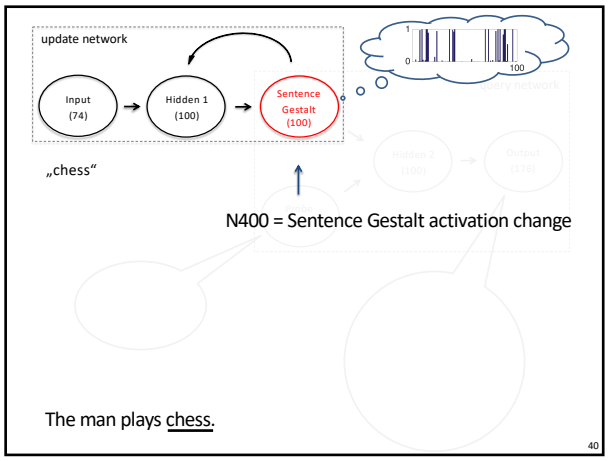
37



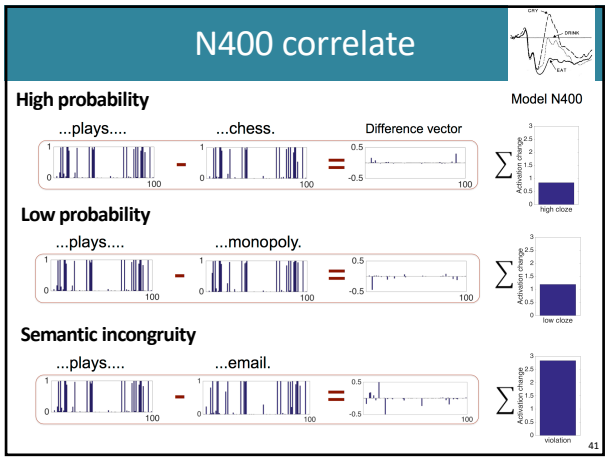
38



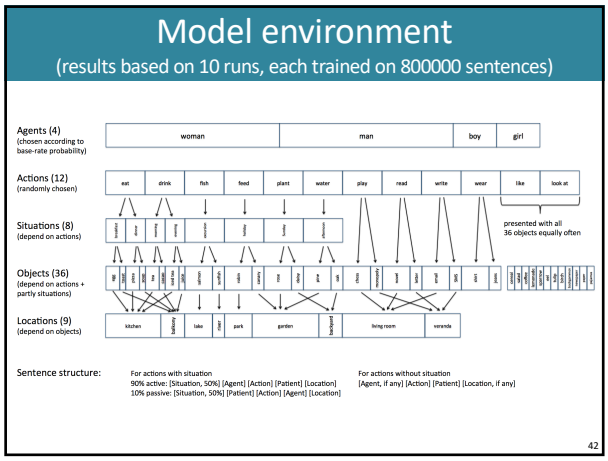
39



40



41



42

Experimental Manipulation
Semantic congruency
Cloze probability
Position in sentence
Categorical relation of incongruent completion
Repetition
Associative priming
Semantic priming
Lexical frequency
Constraint (unexpected endings)
Reversal anomaly
Syntactic violation
Priming during chance performance
Development
Semantic congruency x rep.

43

43

Experimental Manipulation	N400 (empirical)
Semantic congruency	incongruent > congruent
Cloze probability	low > high
Position in sentence	early > late
Categorical relation of incongruent completion	incongr. unrel > incongr. rel
Repetition	first pres. > repetition
Associative priming	unrelated > related
Semantic priming	related > related
Lexical frequency	high < low
Constraint (unexpected endings)	no effect
Reversal anomaly	Congruent = reversal < incongr.
Syntactic violation	no effect
Priming during chance performance	unrelated > related
Development	Very young < young > old
Semantic congruency x rep.	interaction

44

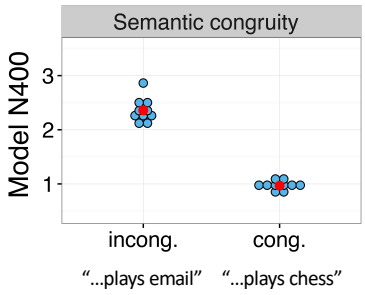
44

Experimental Manipulation	N400 (empirical)	
Semantic congruency	incongruent > congruent	?
Cloze probability	low > high	?
Position in sentence	early > late	?
Categorical relation of incongruent completion	incongr. unrel > incongr. rel	?
Repetition	first pres. > repetition	?
Associative priming	unrelated > related	?
Semantic priming	related > related	?
Lexical frequency	high < low	?
Constraint (unexpected endings)	no effect	?
Reversal anomaly	Congruent = reversal < incongr.	?
Syntactic violation	no effect	?
Priming during chance performance	unrelated > related	?
Development	Very young < young > old	?
Semantic congruency x rep.	interaction	?

45

45

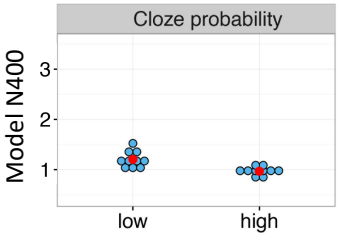
Semantic congruity: incongruent > congruent



46

46

Cloze probability: Lower > higher

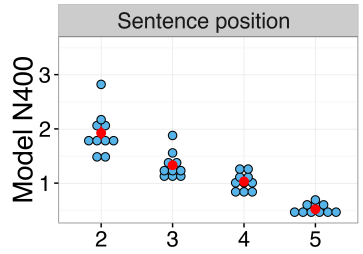


47

47

Position in sentence: Early > late

"At breakfast, the boy eats eggs in the kitchen."



48

48

N400 as word surprisal?

- Correlation between N400 and word suprisal measured by a simple recurrent network (Frank et al., 2015)

However

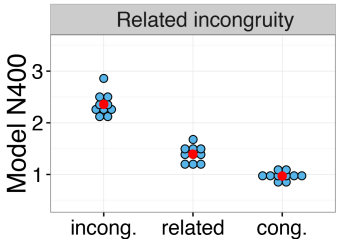
"They wanted to make the hotel look more like a tropical resort. So along the driveway they planted rows of…"

palms < pines < tulips

(Federmeier & Kutas, 1999)

49

Semantically related incongruities



50

N400 as the effort of semantic integration?

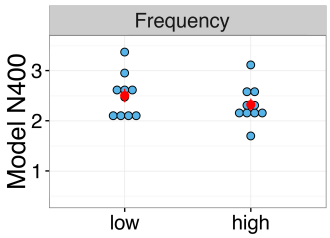
- But: No sentence context needed

➢ N400 effects for single words and words pairs

➢ Use SG model to simulate N400 effects outside of a sentence context

51

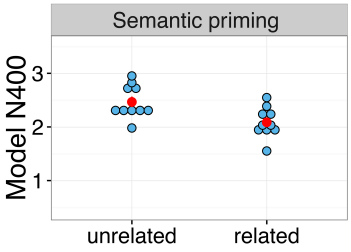
Lexical frequency: Low > high



→ reflects the encoding of base rate probabilities

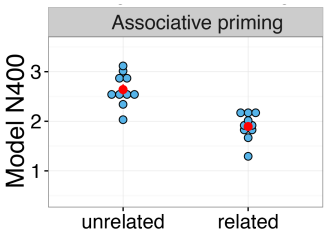
52

Semantic priming: Related (cat – dog) < unrelated (butter – dog)



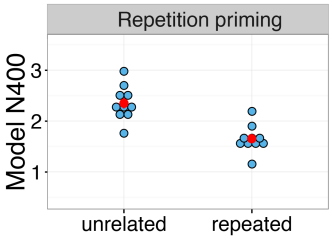
53

Associative priming: Related (sleep – bed) < unrelated (sing – bed)



54

Repetition priming: Repeated (dog – dog) < unrelated (butter – dog)



55

Word and sentence meaning

- Model accounts for both, N400 effects at the word and sentence level
  - No assumption of a separate semantic system for word meanings, separate from overall meaning
  - All stimuli produce change in activation state in the same semantic system
- Also important: Specificity
  - Variables that do not influence the N400 should not influence the model's N400 correlate

56

Reversal anomalies

N400 data:

"Every morning at breakfast, the eggs would eat…" =<

"Every morning at breakfast, the boys would eat…" <

"Every morning at breakfast, the boys would plant…"

(Kuperberg et al., 2003)

→ N400: word meaning, not sentence meaning?

(Brouwer, Fitz, & Hoeks, 2012, Brouwer et al., 2017)

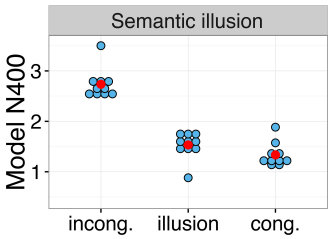
Simulation:

- "At breakfast, the egg eats…"
- "At breakfast, the boy eats…"
- "At breakfast, the boy plants…"

10% passive sentences during training

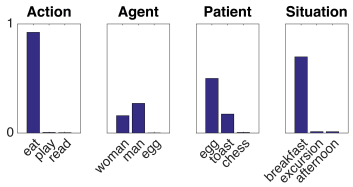
57

Reversal anomaly
incong > rev. anom. >= cong.



58

"At breakfast, the egg eats…"



→ "Semantic illusion" (Kim & Osterhout, 2005)

→ Language processing can be just "good enough"

(Ferreira et al., 2002; Sanford & Stuart, 2002)

P600 increase

→ Re-analysis of the sentence?

→ N400 in reversal anomalies is consistent with N400 as update of representation of sentence meaning

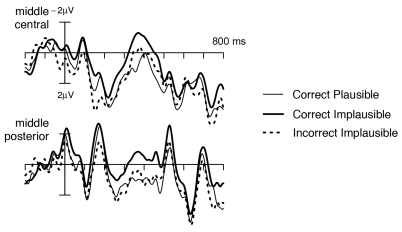
59

Does the N400 reflect the update of a "good enough" representation of sentence meaning?

Rabovsky, Matsuki, & McRae (in preparation):

"The patient was treated by the doctor…" (plausible)

"The clinician was treated by the client…" (implausible)



60

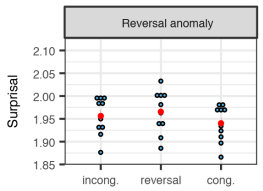
Does the N400 reflect the update of a "good enough" representation of sentence meaning?

- Influence of "good enough" interpretation on N400 amplitudes is in line with our model
- Model linking the N400 to lexical access does not predict this relationship (Brouwer et al., 2017)

61

Word surprisal is large in reversal anomalies

- Correlation between N400 and word suprisal measured by a simple recurrent network (Frank et al., 2015)
- However, word surprisal is large in reversal anomalies ("Every morning at breakfast, the eggs would eat…")

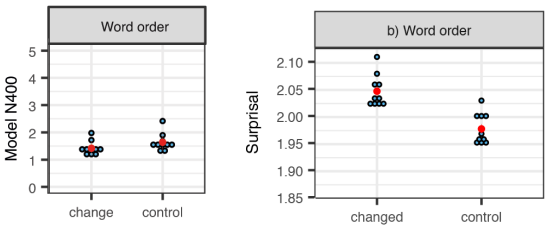


→ unlike N400

62

Syntactic violations
Changes in word order: No effect

"The girl was very satisfied with the ironed neatly linen."

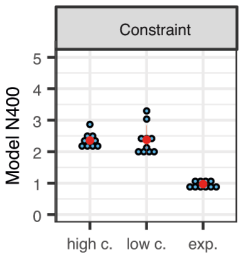


→ unlike N400

63

Constraint: No effect

"The man likes the email."
"The man eats the email."



→ Amount of unexpected semantic information, not violation signal

64

Development

N400 data:

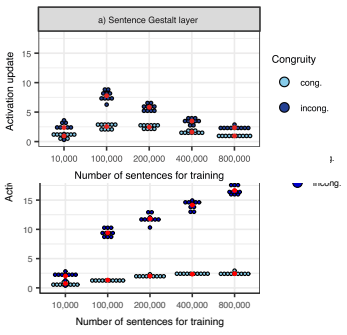
- Increase with comprehension skills in babies (Friedrich et al., 2009)
- Later: decrease with age from childhood through adulthood (Atchley et al., 2006; Kutas & Iragui, 1998)

Simulation:

- Influences of semantic congruity at different points in training

65

Development: Very young < young > old



More efficient connections → small changes at SG sufficient to produce big changes in output activation (~ decreased activation with increased practice)

N400 does not directly reflect change in explicit estimates of feature probabilities but the change of an internal representation that implicitly represents these probabilities such that they can be made explicit when queried.

66

N400 effects during chance performance

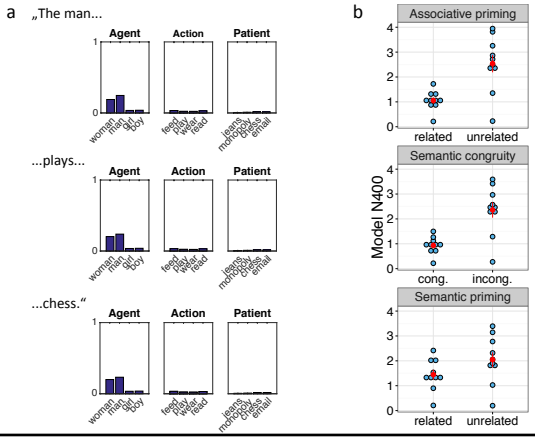
N400 data:

- Learners of new language showed N400 effects of semantic relatedness while performance in lexical decision task was still near chance (McLaughlin et a., 2004)

Simulation:

- Interrupt training after 10000 sentences.

67



68

Semantic congruity X repetition

Experiment (Besson et al., 1992):

- Congruent ("...plays chess") and incongruent ("...plays email") sentences
- All sentences presented twice (in two blocks)

N400 data:

- incongruent > congruent sentence completions
- 1st presentation > (delayed) repetition
- Incongruent (1st – repeated) > congruent (1st – repeated)

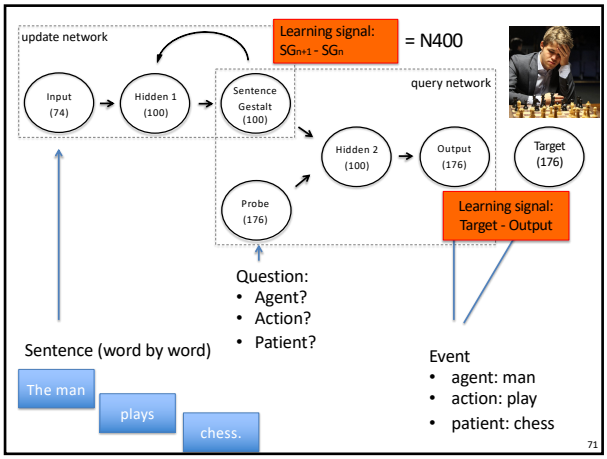
69

69

Semantic congruity X repetition

- Repetition effects as consequences of connection weight adaptations (McClelland & Rumelhart, 1985)
  → Learning operative during first presentation

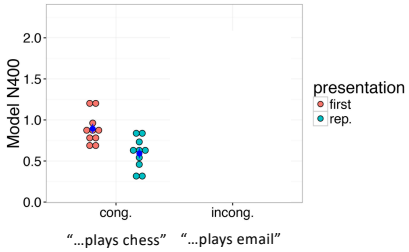
70



71

71

Simulation results



In general: Larger N400 should trigger stronger adaptation

72

72

N400 and adaptation?

- Larger N400-like negativity to single words during study predict enhanced implicit memory (stem completion in absence of explicit memory) during test (Schott et al., 2002)
- Words presented as incongruent sentence completions during study later elicit smaller N400 when presented in isolation (Meyer et al., 2007)

73

Experimental Manipulation	N400 (empirical)	
Semantic congruency	incongruent > congruent	✓
Cloze probability	low > high	✓
Position in sentence	early > late	✓
Categorical relation of incongruent completion	incongr. unrel > incongr. rel	✓
Repetition	first pres. > repetition	✓
Associative priming	unrelated > related	✓
Semantic priming	related > related	✓
Lexical frequency	high < low	✓
Constraint (unexpected endings)	no effect	✓
Reversal anomaly	Congruent = reversal < incongr.	✓
Syntactic violation	no effect	✓
Priming during chance performance	unrelated > related	✓
Development	Very young < young > old	✓
Semantic congruency x rep.	interaction	✓

74

74

Conclusion

- N400 reflects stimulus-driven change in an implicit and probabilistic representation of meaning
- Discrepancy between probabilistically anticipated and encountered features
- Corresponds to learning signal driving adaptation in semantic memory

Rabovsky, Hansen, & McClelland, 2018, *Nature Human Behaviour*

76

76

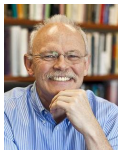
Outlook

Large-scale training of the SG model based on large-scale semantic role corpus (Sayeed et al., 2018; new improved version by Asad Sayeed and Yuval Marton)

77

77

Thank you very much!



Jay McClelland

78