Studying how language comprehension unfolds over time

Cory Shain

September 16, 2020, Computational Psycholinguistics Lab, MIT

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain's response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain's response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain's response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain's response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain’s response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Human cognition unfolds in time
  - Psycholinguistic data are generated by people with brains
  - The brain is a dynamical system that responds to its environment in time
  - Most (all?) psycholinguistic data are underlyingly time series
  - The brain’s response to a stimulus may be slow (temporally diffuse)
  - Psycholinguistic measures may capture lingering response to preceding stimuli

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements?
    (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a “buffer” in human sentence processing?
    (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading?
  - Can syntactic representations be detected in the brain?

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements?
    (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a “buffer” in human sentence processing?
    (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading?
  - Can syntactic representations be detected in the brain?

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements?
    (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a “buffer” in human sentence processing?
    (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading?
    (Grodner and Gibson 2005; Levy 2008; Levy et al. 2013)
  - Can syntactic representations be detected in the brain?

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements? (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a "buffer" in human sentence processing? (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading? (Grodner and Gibson 2005; Levy 2008; Levy et al. 2013)
  - Can syntactic representations be detected in the brain? (Henderson et al. 2015; Brennan et al. 2016)

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements?
    (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a “buffer” in human sentence processing?
    (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading?
    (Grodner and Gibson 2005; Levy 2008; Levy et al. 2013)
  - Can syntactic representations be detected in the brain?
    (Henderson et al. 2015; Brennan et al. 2016)

Time and the mind

- Lots of hypotheses directly concern time:
  - How closely time-locked are attention and eye movements?
    (Just and Carpenter 1980; Posner 1980; Reichle et al. 1998)
  - Is there a “buffer” in human sentence processing?
    (Morton 1964; Mitchell 1984; Mollica and Piantadosi 2017)
- Others depend on assumptions about timecourse
  - Are there distinct locality and predictability effects in reading?
    (Grodner and Gibson 2005; Levy 2008; Levy et al. 2013)
  - Can syntactic representations be detected in the brain?
    (Henderson et al. 2015; Brennan et al. 2016)

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG
    - Event-related potentials/fields
  - Visual world ET
    - Event-related change in fixation proportion
  - fMRI
    - Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading
  - Self-paced reading

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG
    Event-related potentials/fields
  - Visual world ET
    Event-related change in fixation proportion
  - fMRI
    Hemodynamic response function (HRF)
- Latency is often ignored for others
    Eye-tracking during reading
    Self-paced reading

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG

    Event-related potentials/fields
  - Visual world ET

    Event-related change in fixation proportion
  - fMRI

    Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading
  - Self-paced reading

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG

    Event-related potentials/fields
  - Visual world ET

    Event-related change in fixation proportion
  - fMRI

    Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading

    (Frank and Bod 2011; van Schijndel and Schuler 2015)
  - Self-paced reading

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG
    Event-related potentials/fields
  - Visual world ET
    Event-related change in fixation proportion
  - fMRI
    Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading
    (Frank and Bod 2011; van Schijndel and Schuler 2015)
  - Self-paced reading
    (Shain et al. 2016; van Schijndel and Linzen 2018)

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG
    Event-related potentials/fields
  - Visual world ET
    Event-related change in fixation proportion
  - fMRI
    Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading
    (Frank and Bod 2011; van Schijndel and Schuler 2015)
  - Self-paced reading
    (Shain et al. 2016; van Schijndel and Linzen 2018)

Time and the mind

- Latency is taken for granted for some measures
  - EEG/MEG
    Event-related potentials/fields
  - Visual world ET
    Event-related change in fixation proportion
  - fMRI
    Hemodynamic response function (HRF)
- Latency is often ignored for others
  - Eye-tracking during reading
    (Frank and Bod 2011; van Schijndel and Schuler 2015)
  - Self-paced reading
    (Shain et al. 2016; van Schijndel and Linzen 2018)

Time and the mind

- Time matters even if it's not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

Time and the mind

- Time matters even if it’s not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e-10$
    - Constituent wrap-up: $p = 2.33e-14$
  - Change spillover of one baseline variable:
    - Locality: $p =$ [illegible]
    - Constituent wrap-up: $p =$ [illegible]

Time and the mind

- Time matters even if it's not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

Time and the mind

- Time matters even if it's not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

Time and the mind

- Time matters even if it's not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

Time and the mind

- Time matters even if it’s not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

Time and the mind

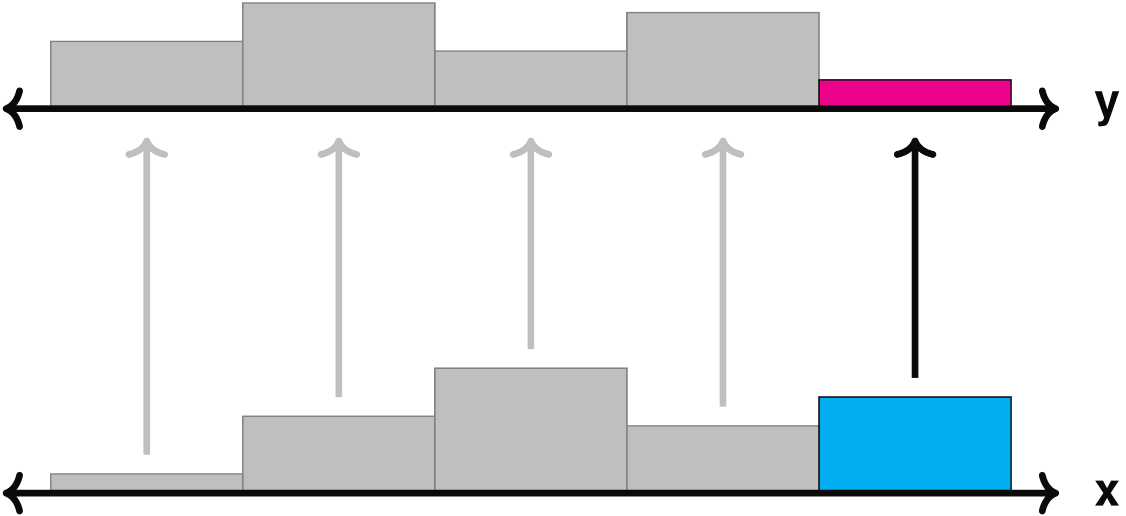
- Time matters even if it's not part of your hypothesis
  - Shain et al. 2016:
    - Locality: $p = 4.87e - 10$
    - Constituent wrap-up: $p = 2.33e - 14$
  - Change spillover of one baseline variable:
    - Locality: $p = 0.816$
    - Constituent wrap-up: $p = 0.370$

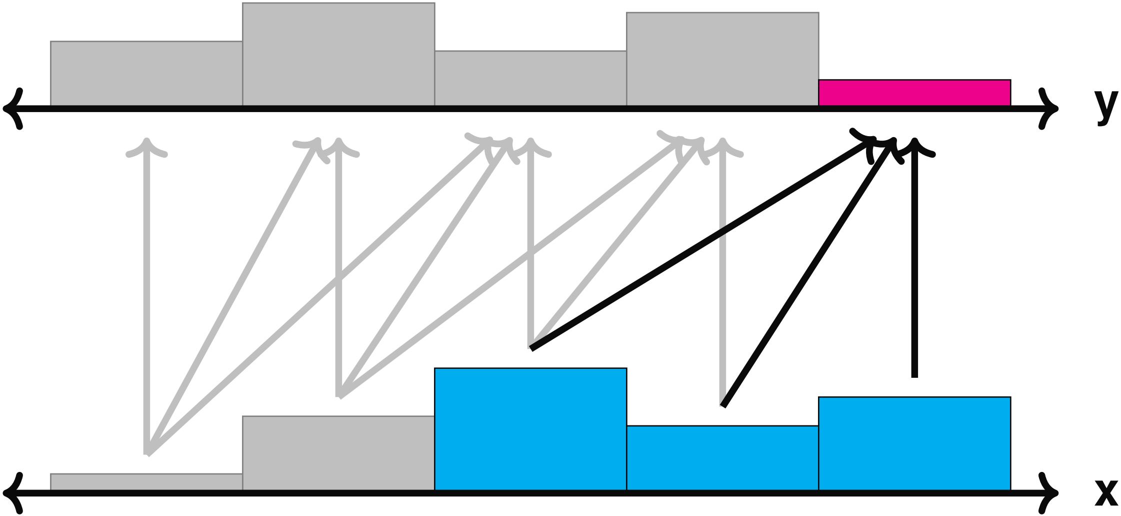
Time and the mind

- Studying cognition means dealing with temporal structure
- Wait, don't we already have spillover for this?

Time and the mind

- Studying cognition means dealing with temporal structure
- Wait, don’t we already have spillover for this?


y
x


y
x

Time and the mind

- Spillover is a **distributed lag (DL)** or **finite impulse response (FIR)** model
- DL/FIR is poorly suited to natural language
- Key problem: non-uniform time series

- Spillover ignores variation in event duration

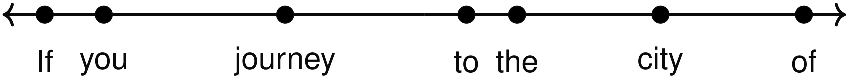
Time and the mind

- Spillover is a **distributed lag (DL)** or **finite impulse response (FIR)** model
- DL/FIR is poorly suited to natural language
- Key problem: non-uniform time series

- Spillover ignores variation in event duration

Time and the mind

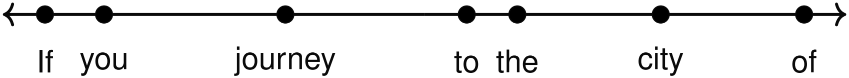
- Spillover is a **distributed lag (DL)** or **finite impulse response (FIR)** model
- DL/FIR is poorly suited to natural language
- Key problem: non-uniform time series



- Spillover ignores variation in event duration

Time and the mind

- Spillover is a **distributed lag (DL)** or **finite impulse response (FIR)** model
- DL/FIR is poorly suited to natural language
- Key problem: non-uniform time series



- Spillover ignores variation in event duration

Time and the mind

- What if we could study delayed effects over continuous time?
  - Continuous **impulse response function (IRF)** would allow convolution over time
  - Response becomes weighted sum of all preceding stimuli
  - IRF provides the weights

Time and the mind

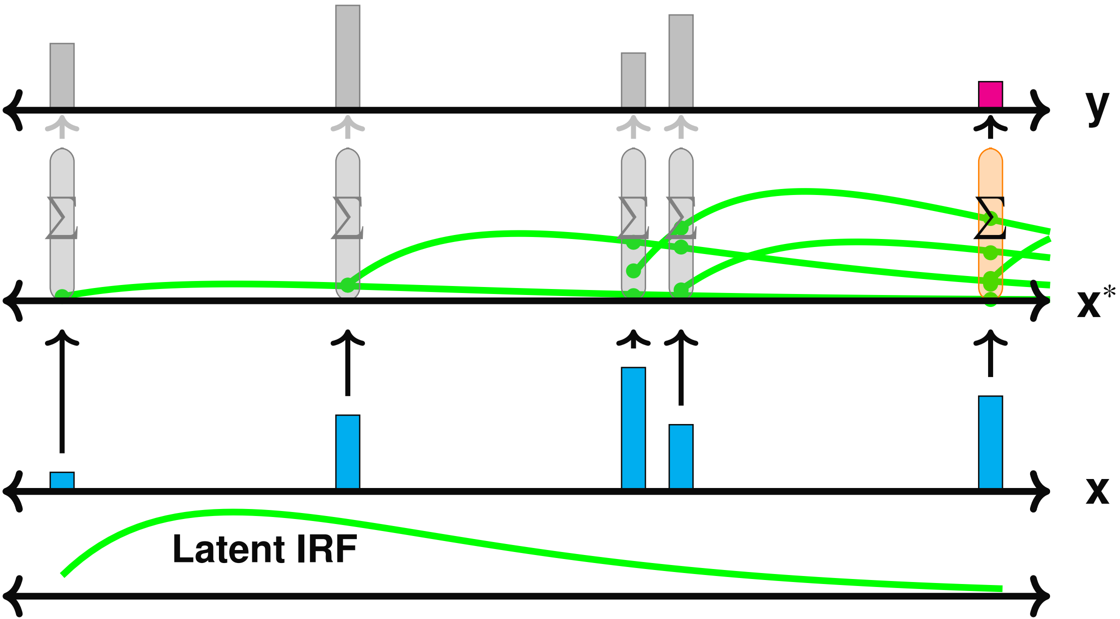
- What if we could study delayed effects over continuous time?
  - Continuous **impulse response function (IRF)** would allow convolution over time
  - Response becomes weighted sum of all preceding stimuli
  - IRF provides the weights

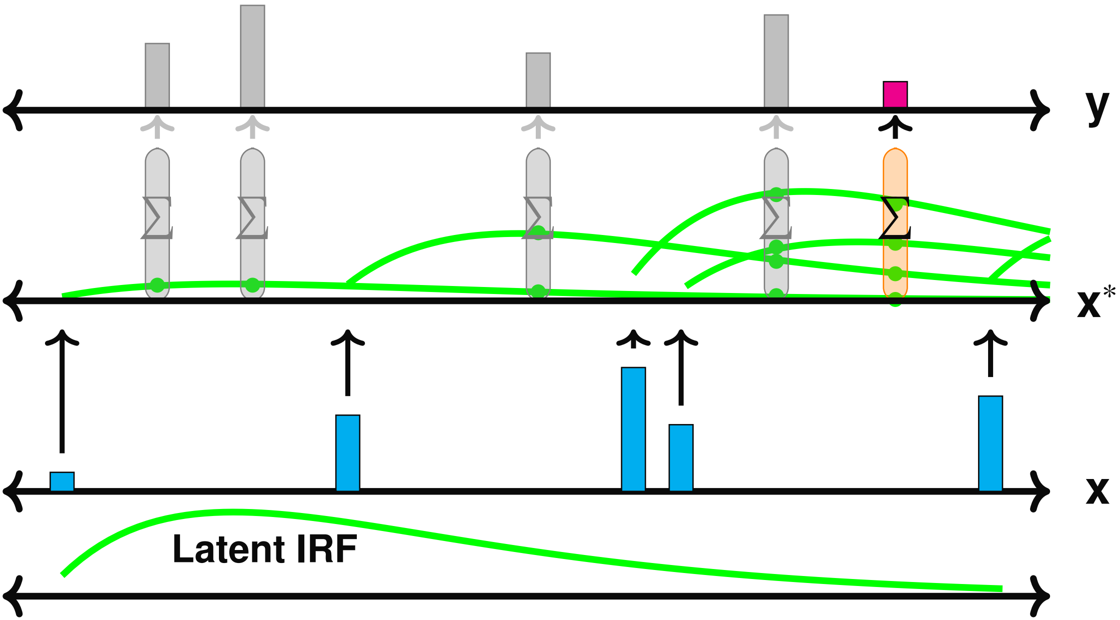
Time and the mind

- What if we could study delayed effects over continuous time?
  - Continuous **impulse response function (IRF)** would allow convolution over time
  - Response becomes weighted sum of all preceding stimuli
  - IRF provides the weights

Time and the mind

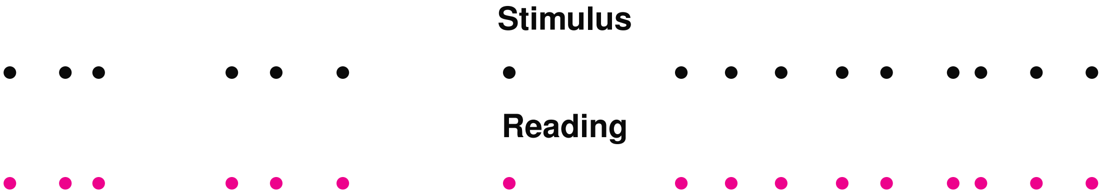
- What if we could study delayed effects over continuous time?
  - Continuous **impulse response function (IRF)** would allow convolution over time
  - Response becomes weighted sum of all preceding stimuli
  - IRF provides the weights

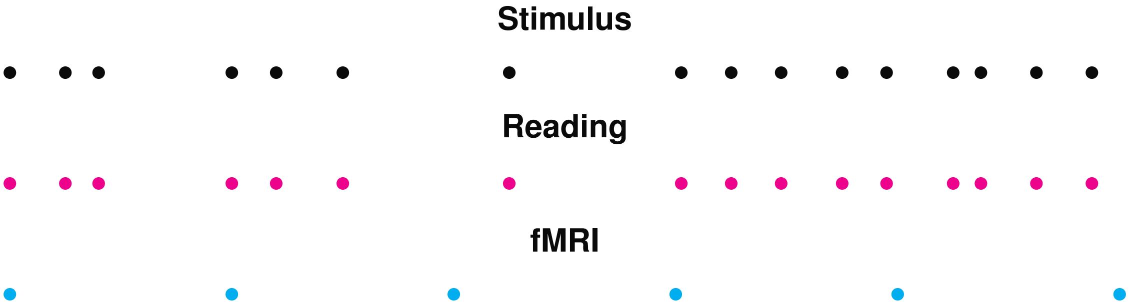

y
x*
x
Latent IRF

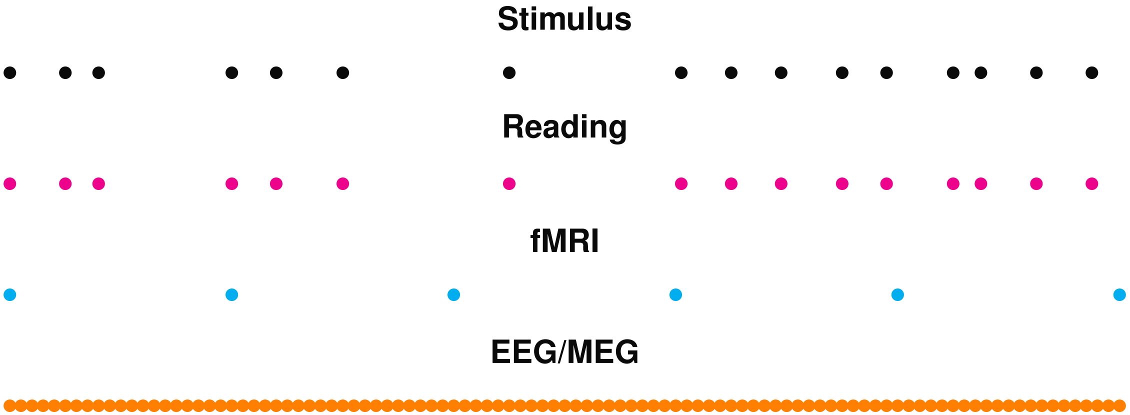


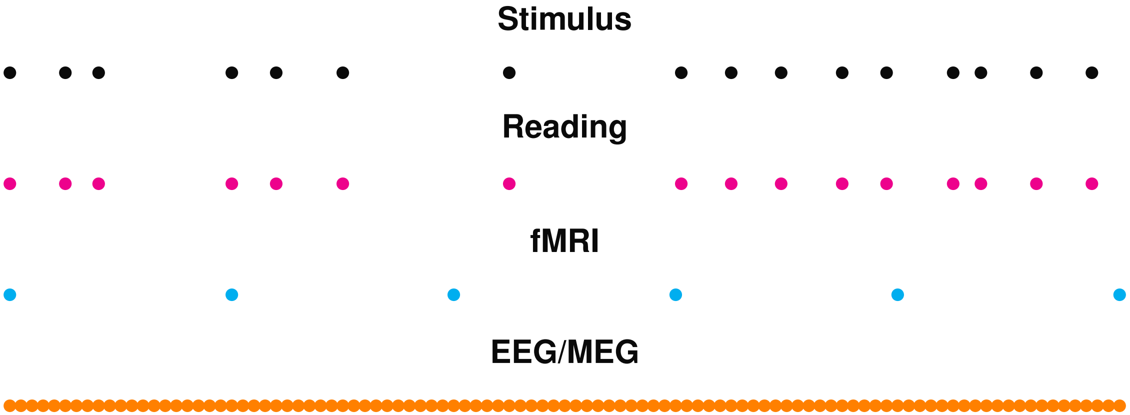
y
x*
x
Latent IRF

Stimulus


Stimulus
Reading


Stimulus
Reading
fMRI


Stimulus
Reading
fMRI
EEG/MEG



Each of these domains is typically analyzed with different methods.
CDR provides a unified approach.

CDR at a high level

- Design stimuli
- Collect responses
- Specify IRF kernel (exponential, Gaussian, gamma, etc.)
- Estimate IRF parameters from data
- Compare goodness of fit on held-out evaluation set

CDR at a high level

- Design stimuli
- Collect responses
- Specify IRF kernel (exponential, Gaussian, gamma, etc.)
- Estimate IRF parameters from data
- Compare goodness of fit on held-out evaluation set

CDR at a high level

- Design stimuli
- Collect responses
- Specify IRF kernel (exponential, Gaussian, gamma, etc.)
- Estimate IRF parameters from data
- Compare goodness of fit on held-out evaluation set

CDR at a high level

- Design stimuli
- Collect responses
- Specify IRF kernel (exponential, Gaussian, gamma, etc.)
- Estimate IRF parameters from data
- Compare goodness of fit on held-out evaluation set

CDR at a high level

- Design stimuli
- Collect responses
- Specify IRF kernel (exponential, Gaussian, gamma, etc.)
- Estimate IRF parameters from data
- Compare goodness of fit on held-out evaluation set

Does it work?

Yes.

Synthetic Experiments

Synthetic Experiments

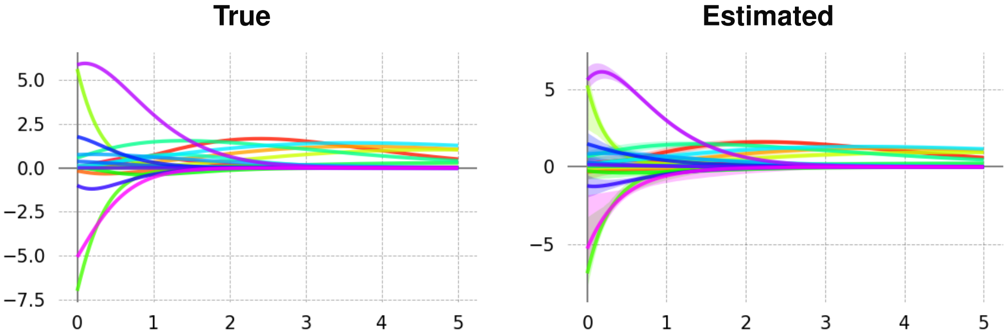
- Generate data from a temporally diffuse process
- Fit CDR to that data
- Check that CDR recovers the underlying process

Synthetic Experiments

- Generate data from a temporally diffuse process
- Fit CDR to that data
- Check that CDR recovers the underlying process

Synthetic Experiments

- Generate data from a temporally diffuse process
- Fit CDR to that data
- Check that CDR recovers the underlying process


True
5.0
2.5
0.0
−2.5
−5.0
−7.5
0
1
2
3
4
5
Estimated
5
0
−5
0
1
2
3
4
5

Synthetic Experiments

- Manipulations
  - Noise
  - Multicollinearity
  - Misspecified IRF kernel

Synthetic Experiments

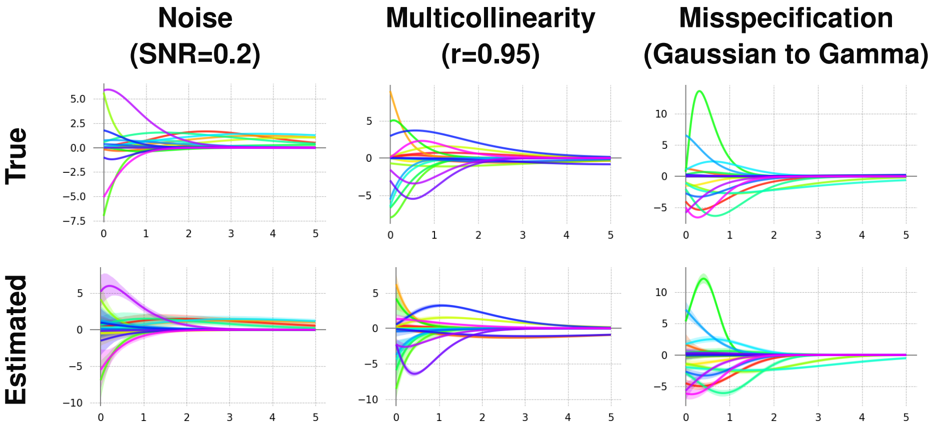
- Manipulations
  - Noise
  - Multicollinearity
  - Misspecified IRF kernel

Synthetic Experiments

- Manipulations
  - Noise
  - Multicollinearity
  - Misspecified IRF kernel

Synthetic Experiments

- Manipulations
  - Noise
  - Multicollinearity
  - Misspecified IRF kernel


Noise
(SNR=0.2)
Multicollinearity
(r=0.95)
Misspecification
(Gaussian to Gamma)
True
Estimated

Naturalistic Reading Experiments

- Are CDR estimates
  - Plausible?
  - Consistent?
  - Externally valid?

Naturalistic Reading Experiments

- Are CDR estimates
  - Plausible?
  - Consistent?
  - Externally valid?

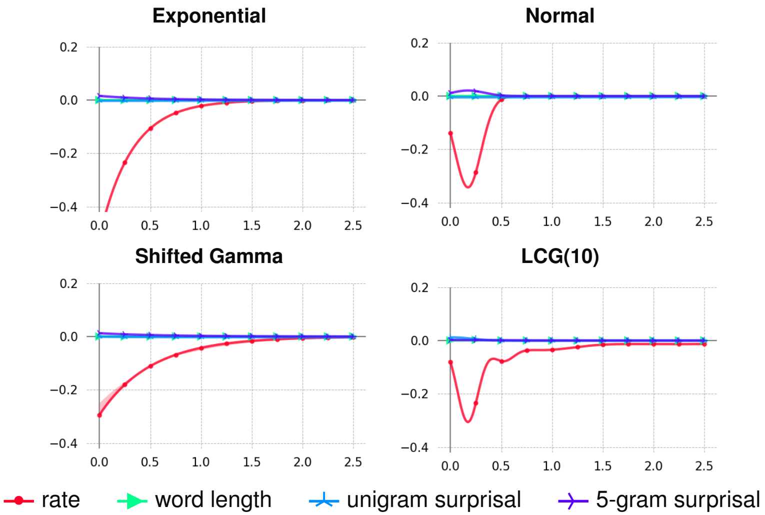
Naturalistic Reading Experiments

- Are CDR estimates
  - Plausible?
  - Consistent?
  - Externally valid?

Naturalistic Reading Experiments

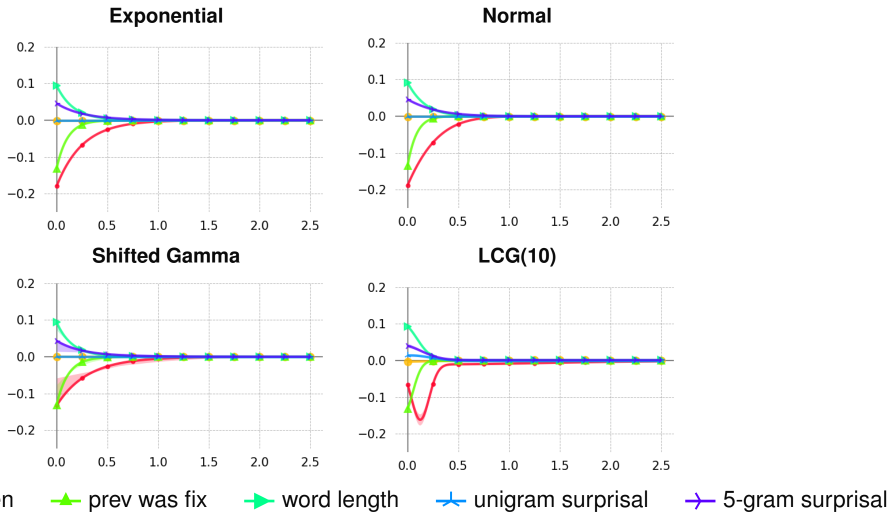
- Are CDR estimates
  - Plausible?
  - Consistent?
  - Externally valid?

Naturalistic Reading Experiments: Plausible and Consistent



Natural Stories (SPR; Futrell et al. 2018)

Naturalistic Reading Experiments: Plausible and Consistent



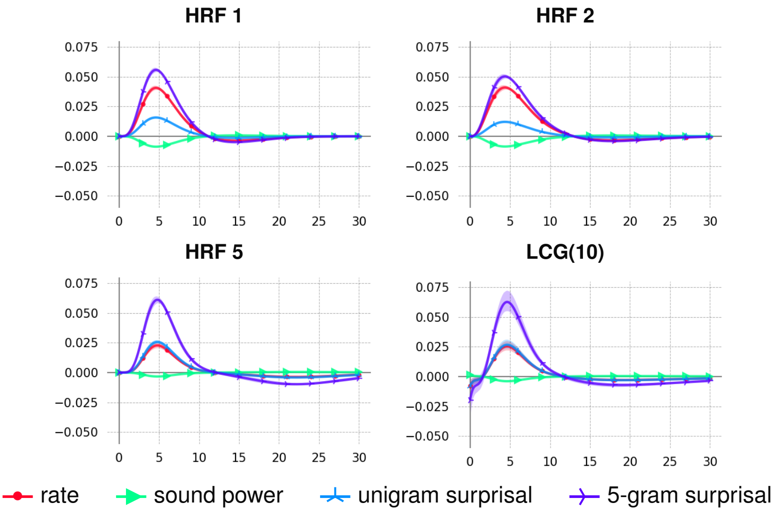
Dundee (ET, go-past; Kennedy et al. 2003)

Naturalistic Reading Experiments: Externally Valid

Baseline	Permutation test *p*
LME	**1.0e-4***
LME-S	**1.0e-4***
GAM	**1.0e-4***
GAM-S	**1.0e-4***

CDR generalizes better

Naturalistic fMRI Experiments: Plausible and Consistent



Natural Stories (fMRI; Shain et al. 2019)

Naturalistic fMRI Experiments: Externally Valid

Baseline	Permutation test p
Canonical HRF	**4.0e-4***
Interpolated	**1.0e-4***
Averaged	**1.0e-4***
Lanczos	**1.0e-4***

CDR generalizes better

What can CDR tell us about the mind?

Frequency vs. Predictability (Shain 2019)

- Are frequency and predictability effects driven by different mechanisms?
  - **Yes**

    (Seidenberg and McClelland 1989; Coltheart et al. 2001; Harm and Seidenberg 2004)
  - **No**

    (Norris 2006; Levy 2008; Rasmussen and Schuler 2018)
- CDR analysis of Dundee, Natural Stories, and UCL does **not** support a dissociation

Frequency vs. Predictability (Shain 2019)

- Are frequency and predictability effects driven by different mechanisms?
  - **Yes**
    (Seidenberg and McClelland 1989; Coltheart et al. 2001; Harm and Seidenberg 2004)
  - **No**
    (Norris 2006; Levy 2008; Rasmussen and Schuler 2018)
- CDR analysis of Dundee, Natural Stories, and UCL does **not** support a dissociation

Frequency vs. Predictability (Shain 2019)

- Are frequency and predictability effects driven by different mechanisms?
  - **Yes**
    (Seidenberg and McClelland 1989; Coltheart et al. 2001; Harm and Seidenberg 2004)
  - **No**
    (Norris 2006; Levy 2008; Rasmussen and Schuler 2018)
- CDR analysis of Dundee, Natural Stories, and UCL does **not** support a dissociation

Frequency vs. Predictability (Shain 2019)

- Are frequency and predictability effects driven by different mechanisms?
  - **Yes**
    (Seidenberg and McClelland 1989; Coltheart et al. 2001; Harm and Seidenberg 2004)
  - **No**
    (Norris 2006; Levy 2008; Rasmussen and Schuler 2018)
- CDR analysis of Dundee, Natural Stories, and UCL does **not** support a dissociation

Domain-Generality and Syntax-Sensitivity of Prediction Effects (Shain et al. 2019)

- Does linguistic prediction recruit domain-general mechanisms?
  (Kaan and Swaab 2002; Novick et al. 2005; Federmeier et al. 2010; Gambi et al. 2018)
- Is linguistic prediction sensitive to syntactic structure?
  (Frank and Bod 2011; Frank and Christiansen 2018)
- CDR analysis of fMRI supports **domain-specific**, **syntax-sensitive** word prediction

Domain-Generality and Syntax-Sensitivity of Prediction Effects (Shain et al. 2019)

- Does linguistic prediction recruit domain-general mechanisms?
  (Kaan and Swaab 2002; Novick et al. 2005; Federmeier et al. 2010; Gambi et al. 2018)
- Is linguistic prediction sensitive to syntactic structure?
  (Frank and Bod 2011; Frank and Christiansen 2018)
- CDR analysis of fMRI supports **domain-specific**, **syntax-sensitive** word prediction

Domain-Generality and Syntax-Sensitivity of Prediction Effects (Shain et al. 2019)

- Does linguistic prediction recruit domain-general mechanisms?
  (Kaan and Swaab 2002; Novick et al. 2005; Federmeier et al. 2010; Gambi et al. 2018)
- Is linguistic prediction sensitive to syntactic structure?
  (Frank and Bod 2011; Frank and Christiansen 2018)
- CDR analysis of fMRI supports **domain-specific**, **syntax-sensitive** word prediction

A Step Further: CDRNN

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

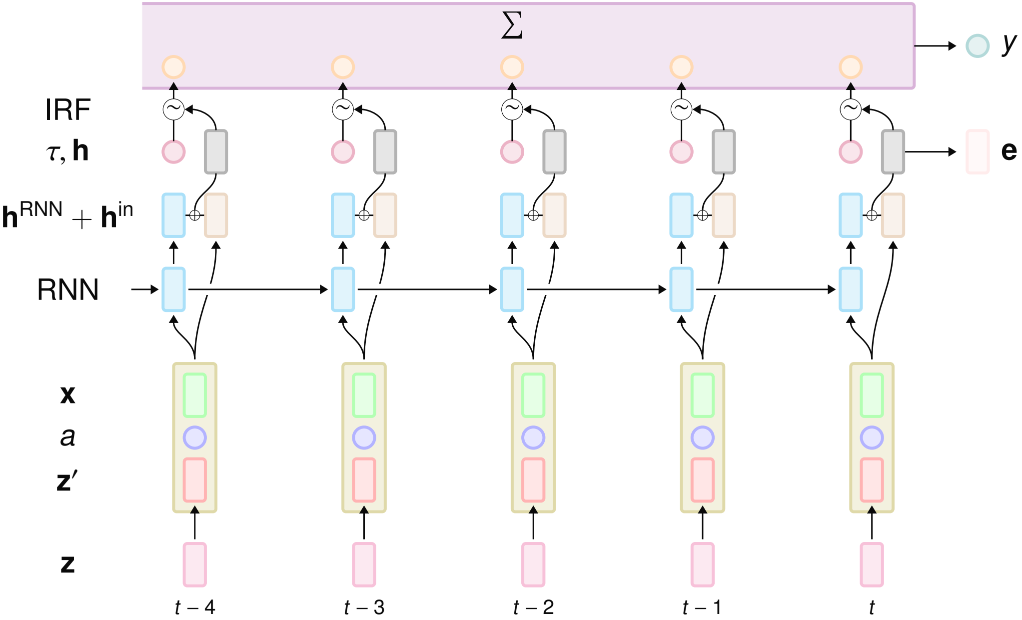
- CDR directly models influence of past on present
- Still makes lots of strict (i.e. implausible) assumptions:
  - Fixed (stationary) and context-independent IRF
  - Linear effects
  - Additive effects
  - Constant variance (homoskedasticity)
- We can relax these with a neural net and still have an interpretable IRF...

A Step Further: CDRNN

- CDRNN: parameterize IRF using deep neural transform of time series
- IRF = multi-dimensional manifold over predictors + time

A Step Further: CDRNN

- CDRNN: parameterize IRF using deep neural transform of time series
- IRF = multi-dimensional manifold over predictors + time



Σ
y
IRF
$\tau, \mathbf{h}$
$\mathbf{e}$
$\mathbf{h}^{\text{RNN}} + \mathbf{h}^{\text{in}}$
RNN
$\mathbf{x}$
a
$\mathbf{z}'$
$\mathbf{z}$
$t-4$
$t-3$
$t-2$
$t-1$
t

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

A Step Further: CDRNN

- Captures:
  - Influence of time and context on IRF
  - Changing error distribution over time
  - Non-linear interactions
- IRF can be queried *post hoc* at
  - Any combination of predictor values
  - Any timepoint
  - Any initial state
- Test hypotheses using out-of-sample goodness-of-fit
- Detailed analysis of complex dynamics, weak initial assumptions

Thank you!

Code: `https://github.com/coryshain/cdr` **Preprint:**

`https://psyarxiv.com/whvk5/`

References

- Brennan, Jonathan et al. (2016). "Abstract linguistic structure correlates with temporal activity during naturalistic comprehension". In: Brain and Language 157, pp. 81–94.
- Coltheart, Max et al. (2001). "DRC: a dual route cascaded model of visual word recognition and reading aloud.". In: Psychological review 108.1, p. 204.
- Federmeier, Kara D, Marta Kutas, and Rina Schul (2010). "Age-related and individual differences in the use of prediction during language comprehension". In: Brain and language 115.3, pp. 149–161.
- Frank, Stefan L and Rens Bod (2011). "Insensitivity of the Human Sentence-Processing System to Hierarchical Structure". In: Psychological Science 22.6, pp. 829–834. ISSN: 0956-7976. DOI: 10.1177/0956797611409589. URL: http://www.ncbi.nlm.nih.gov/pubmed/21586764http://journals.sagepub.com/doi/10.1177/0956797611409589.
- Frank, Stefan L and Morten H Christiansen (2018). "Hierarchical and sequential processing of language". In: Language, Cognition and Neuroscience 33.9, pp. 1213–1218.

References

- Futrell, Richard et al. (2018). "The Natural Stories Corpus". In: Proceedings of the Eleventh International Conference on Language Resources and Evaluation Ed. by Nicoletta Calzolari et al. Paris, France: European Language Resources Association (ELRA). ISBN: 979-10-95546-00-9.
- Gambi, Chiara et al. (2018). "The development of linguistic prediction: Predictions of sound and meaning in 2- to 5-year-olds". In: Journal of Experimental Child Psychology 173, pp. 351–370. ISSN: 0022-0965. DOI: https://doi.org/10.1016/j.jecp.2018.04.012. URL: http://www.sciencedirect.com/science/article/pii/S0022096517307531.
- Grodner, Daniel J and Edward Gibson (2005). "Consequences of the serial nature of linguistic input". In: Cognitive Science 29, pp. 261–291.
- Harm, Michael W and Mark S Seidenberg (2004). "Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes.". In: Psychological review 111.3, p. 662.
- Henderson, John M et al. (2015). "Neural correlates of fixation duration in natural reading: evidence from fixation-related fMRI". In: NeuroImage 119, pp. 390–397.

References

- Just, Marcel Adam and Patricia A Carpenter (1980). "A theory of reading: From eye fixations to comprehension". In: Psychological Review 87.4, pp. 329–354.
- Kaan, Edith and Tamara Y Swaab (2002). "The brain circuitry of syntactic comprehension". In: Trends in cognitive sciences 6.8, pp. 350–356.
- Kennedy, Alan, James Pynte, and Robin Hill (2003). "The Dundee corpus". In: Proceedings of the 12th European conference on eye movement.
- Levy, Roger (2008). "Expectation-based syntactic comprehension". In: Cognition 106.3, pp. 1126–1177.
- Levy, Roger, Evalina Fedorenko, and Edward Gibson (2013). "The syntactic complexity of Russian relative clauses". In: Journal of Memory and Language 69, pp. 461–495.
- Mitchell, Don C (1984). "An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading". In: New methods in reading comprehension research, pp. 69–89.

References

Mollica, Francis and Steve Piantadosi (2017). “An incremental information-theoretic buffer supports sentence processing”. In: Proceedings of the 39th Annual Cognitive Science Society Meeting.

Morton, John (1964). “The effects of context upon speed of reading, eye movements and eye-voice span”. In: Quarterly Journal of Experimental Psychology 16.4, pp. 340–354.

Norris, Dennis (2006). “The Bayesian Reader: Explaining word recognition as an optimal Bayesian decision process.”. In: Psychological review 113.2, p. 327.

Novick, Jared M, John C Trueswell, and Sharon L Thompson-Schill (2005). “Cognitive control and parsing: Reexamining the role of Broca’s area in sentence comprehension”. In: Cognitive, Affective, \& Behavioral Neuroscience 5.3, pp. 263–281.

Posner, Michael I (1980). “Orienting of attention”. In: The Quarterly Journal of Experimental Psychology 32.1, pp. 3–25.

Rasmussen, Nathan E and William Schuler (2018). “Left-Corner Parsing With Distributed Associative Memory Produces Surprisal and Locality Effects”. In: Cognitive Science 42, pp. 1009–1042.

References

- Reichle, Erik et al. (1998). "Toward a Model of Eye Movement Control in Reading". In: Psychological review 105, pp. 125–157. DOI: 10.1037/0033-295X.105.1.125.
- Seidenberg, Mark S and James L McClelland (1989). "A distributed, developmental model of word recognition and naming". In: Psychological review 96.4, p. 523.
- Shain, Cory (2019). "A large-scale study of the effects of word frequency and predictability in naturalistic reading". In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Cor pp. 4086–4094.
- Shain, Cory et al. (2016). "Memory access during incremental sentence processing causes reading time latency". In: Proceedings of the Computational Linguistics for Linguistic Complexity Workshop. Association for Computational Linguistics, pp. 49–58.
- Shain, Cory et al. (2019). "fMRI reveals language-specific predictive coding during naturalistic sentence comprehension". In: Neuropsychologia 138.

References

- van Schijndel, Marten and Tal Linzen (2018). “A Neural Model of Adaptation in Reading”. In: EMNLP 2018, pp. 4704–4710.
- van Schijndel, Marten and William Schuler (2015). “Hierarchic syntax improves reading time prediction”. In: Proceedings of NAACL-HLT 2015. Association for Computational Linguistics.