Generalised measures of predictive uncertainty in online language processing

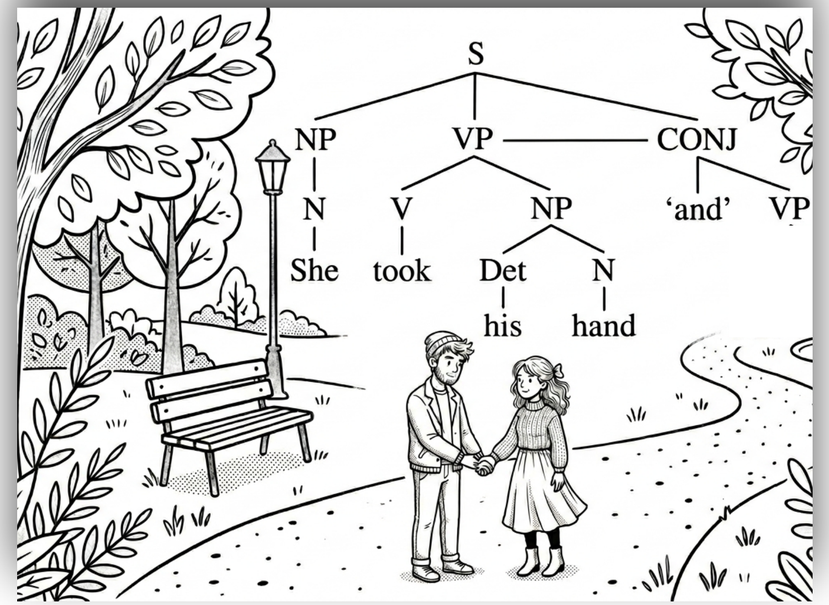
Mario Giulianelli

Computational Humanities Research Group
King's College London, 17 September 2026

What do language processors predict?

She took his hand and...

?



What do language processors predict?

She took his hand and...

?

next word
semantics

intents

sentence
structure

possible
worlds

exact next
word

exact n
next words

How are predictions updated during incremental processing?

She took his hand and

She took his hand and said...



She took his hand and

She took his hand and said…



A window into the information processing that occurs as a human (or a machine) **comprehends** a sentence

Measuring prediction error and/or update gives us a notion of **processing cost** (e.g., to estimate reading difficulty)

Helps explain **production choices** as well (we tend to choose what's more predictable under our own model)

Predicting what comes next is such a powerful **training objective** for artificial neural networks

She took his hand and

She took his hand and said…



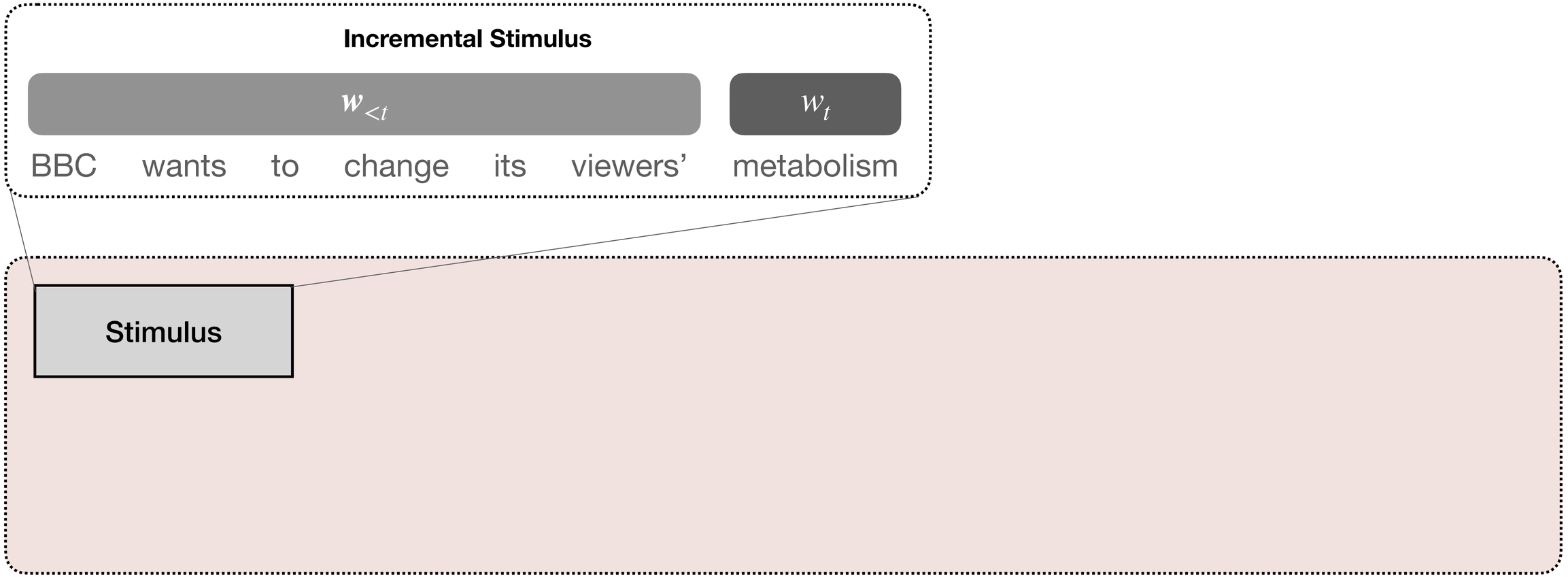
A window into the information processing that occurs as a human (or a machine) **comprehends** a sentence

Measuring prediction error and/or update gives us a notion of **processing cost** (e.g., to estimate reading difficulty)

Helps explain **production choices** as well (we tend to choose what's more predictable under our own model)

Predicting what comes next is such a powerful **training objective** for artificial neural networks

The incremental processing paradigm



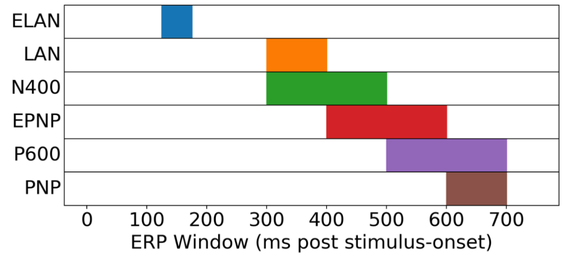
The incremental processing paradigm

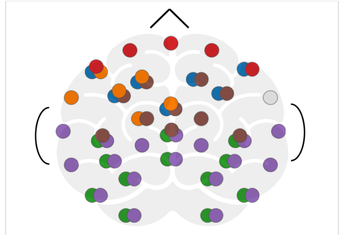
Incremental Stimulus

$w_{<t}$ w_t

BBC wants to change its viewers' metabolism

_E.g.,_ EEG response



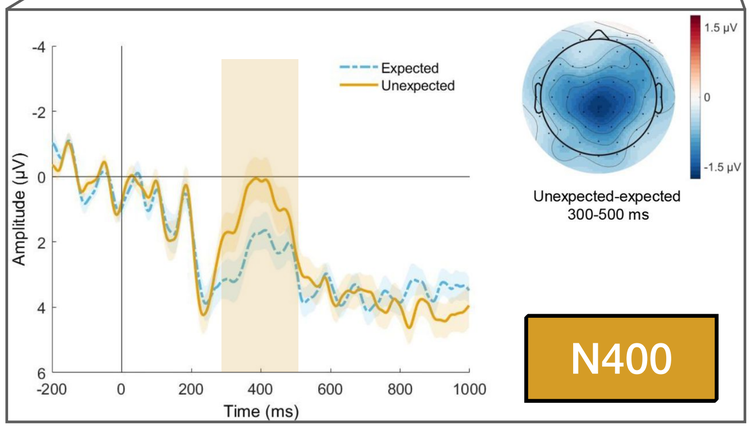


Stimulus

Response

Behavioural and neural responses provide a window into the **information processing tasks** carried out during **language processing**.

BBC wants to change its viewers' **metabolism**.



The incremental processing paradigm

Stimulus

$w_{<t}$ w_t

$t \in 1..T$

Response

e.g.



Brain imaging



Eye tracking

The incremental processing paradigm

Hypothesis

(Explanation)

Stimulus

$w_{<t}$ w_t

$t \in 1..T$

Response

e.g.

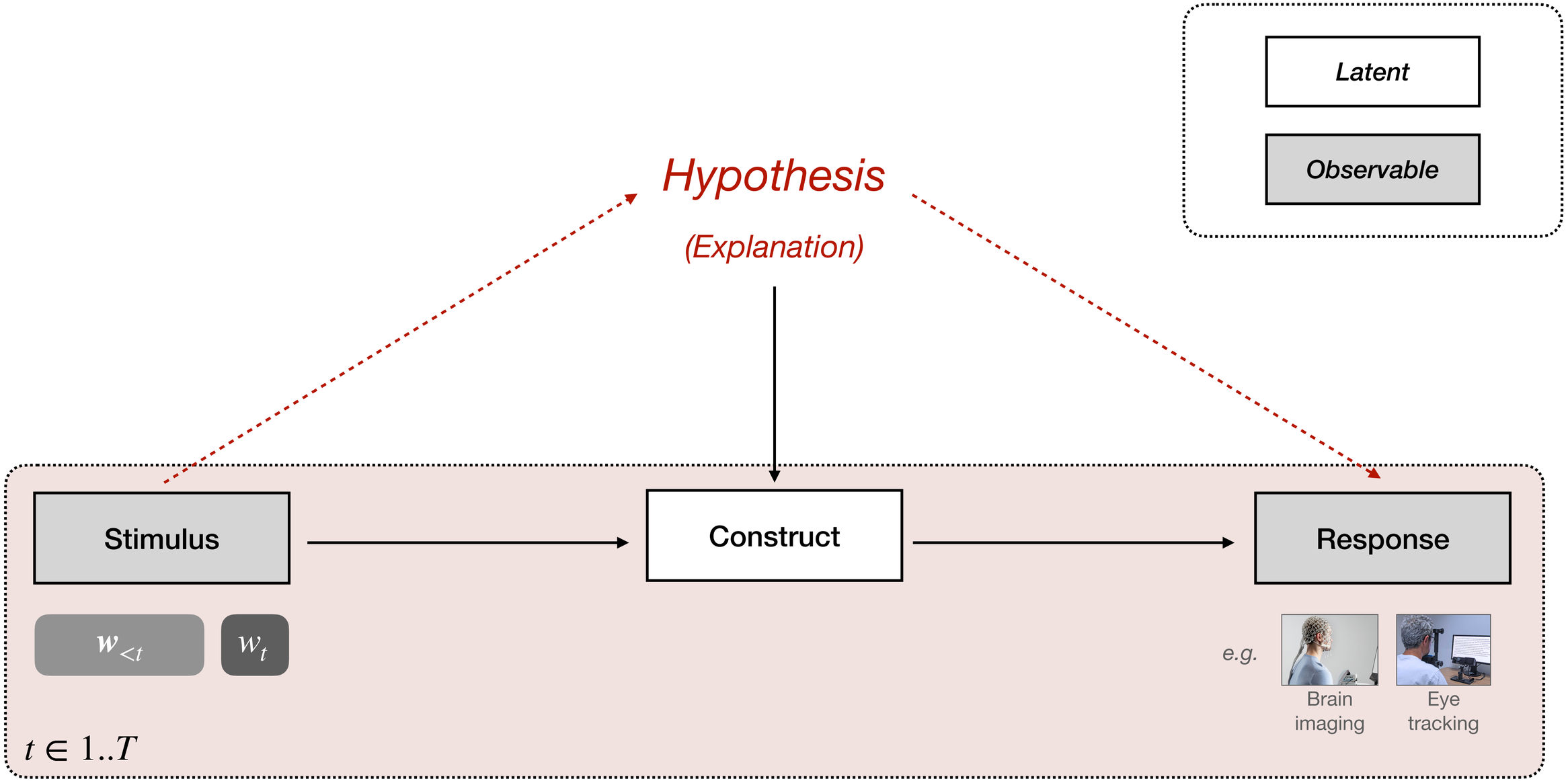


Brain imaging

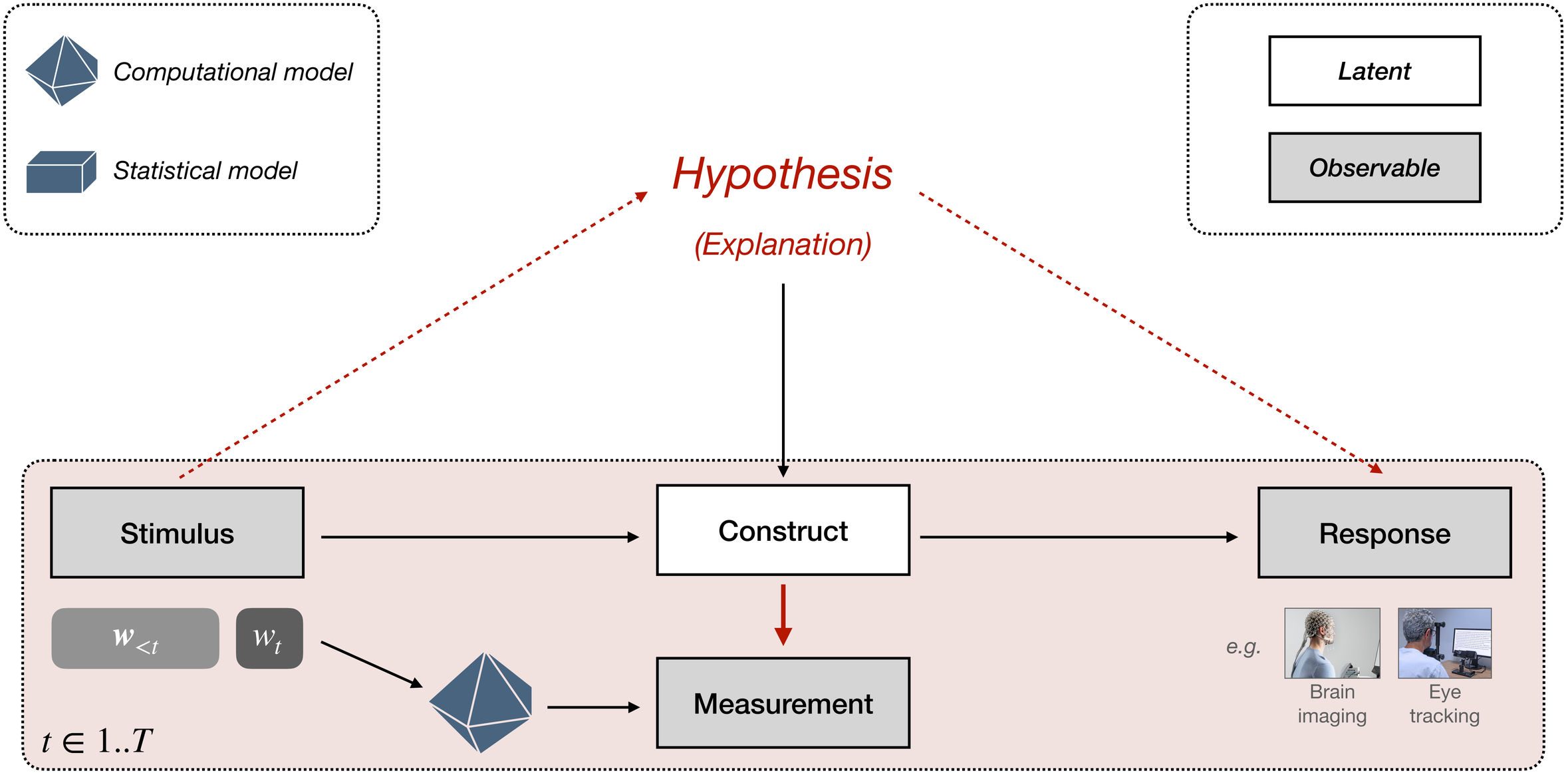


Eye tracking

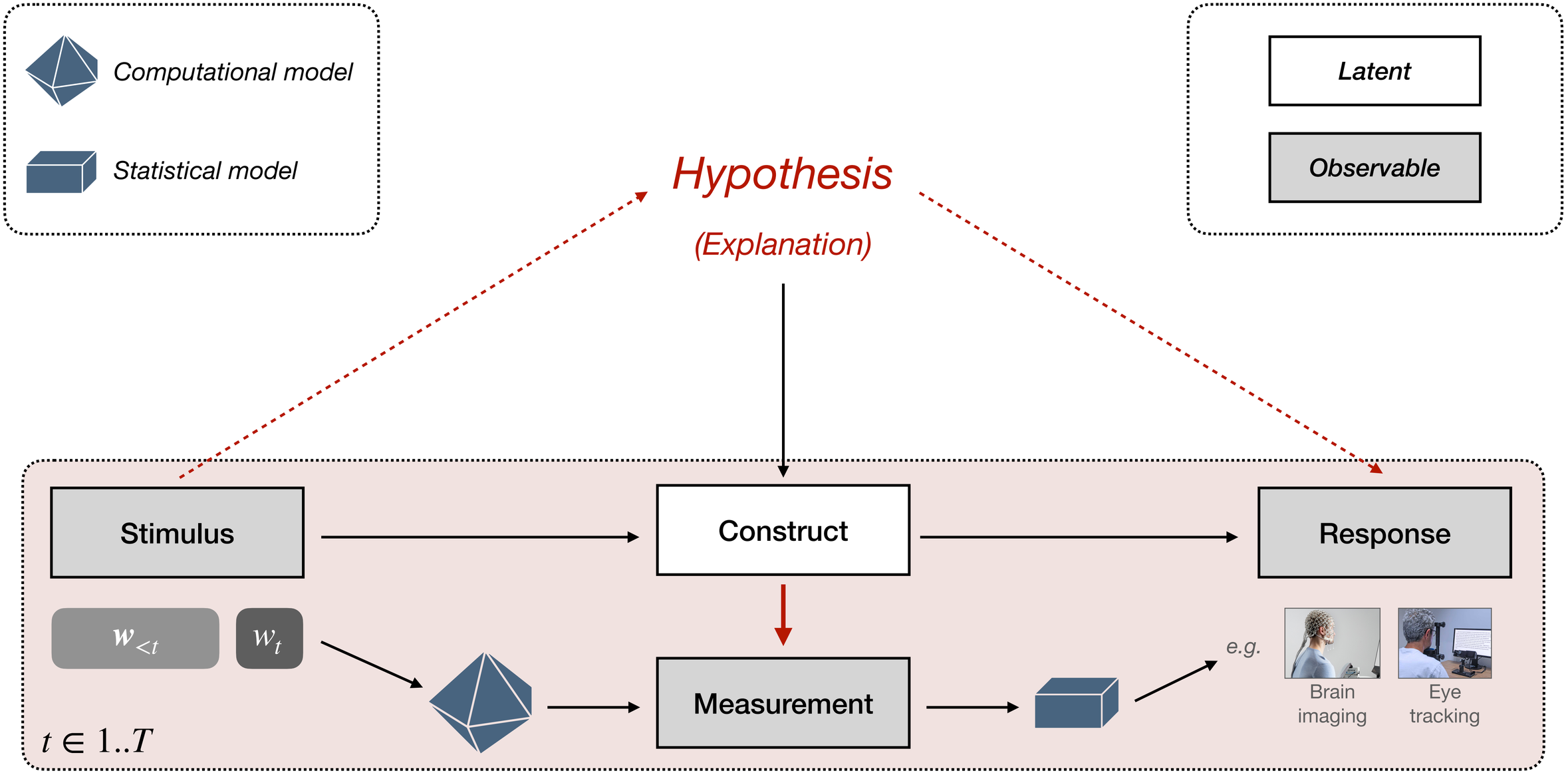
The incremental processing paradigm



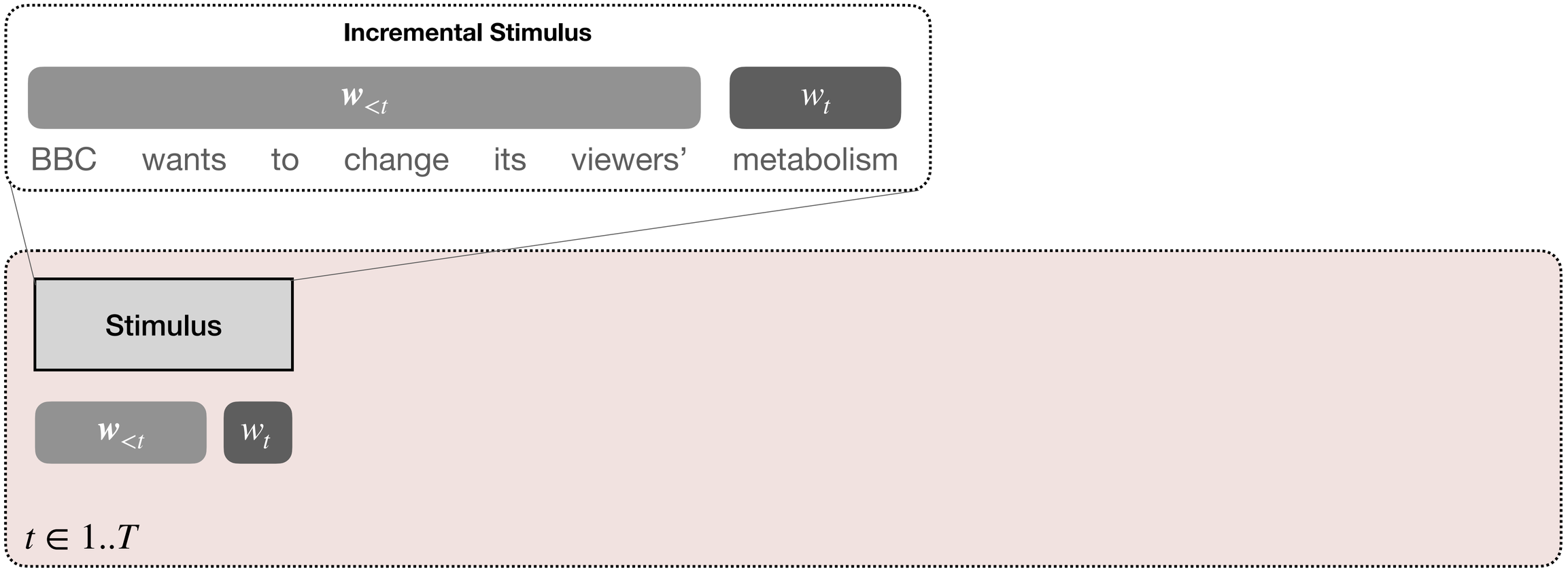
The incremental processing paradigm



The incremental processing paradigm

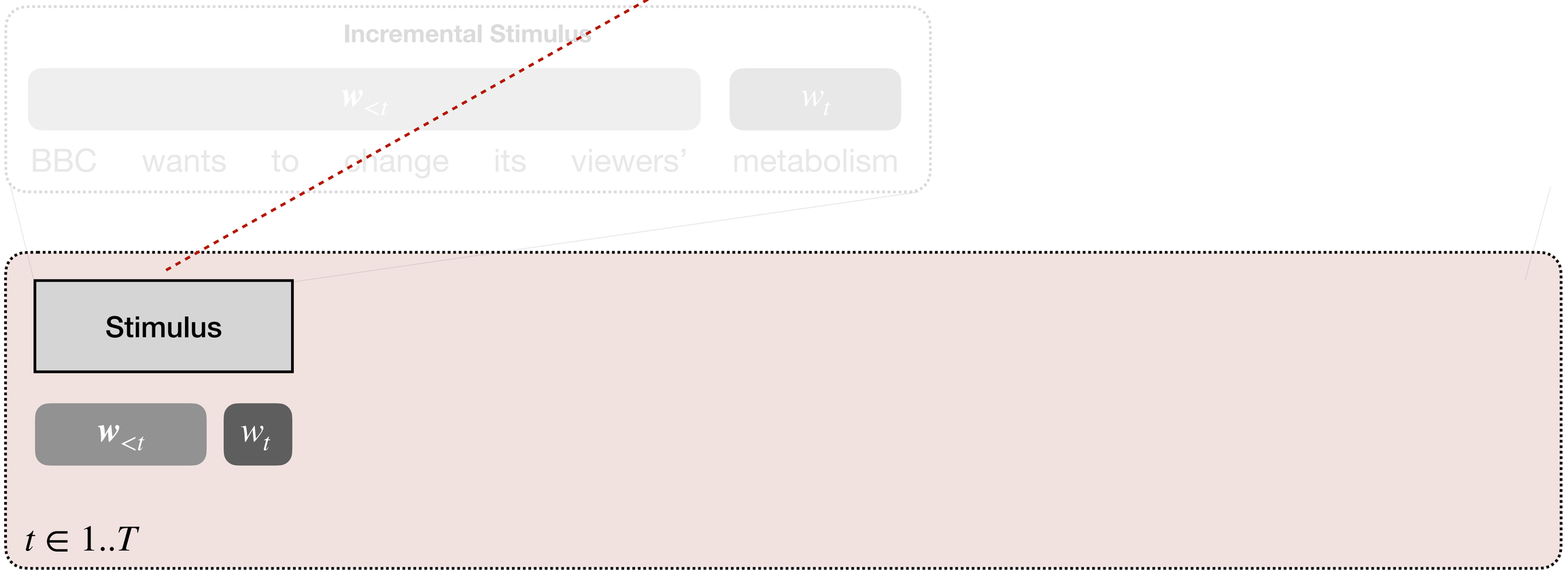


The Surprisal model of incremental comprehension



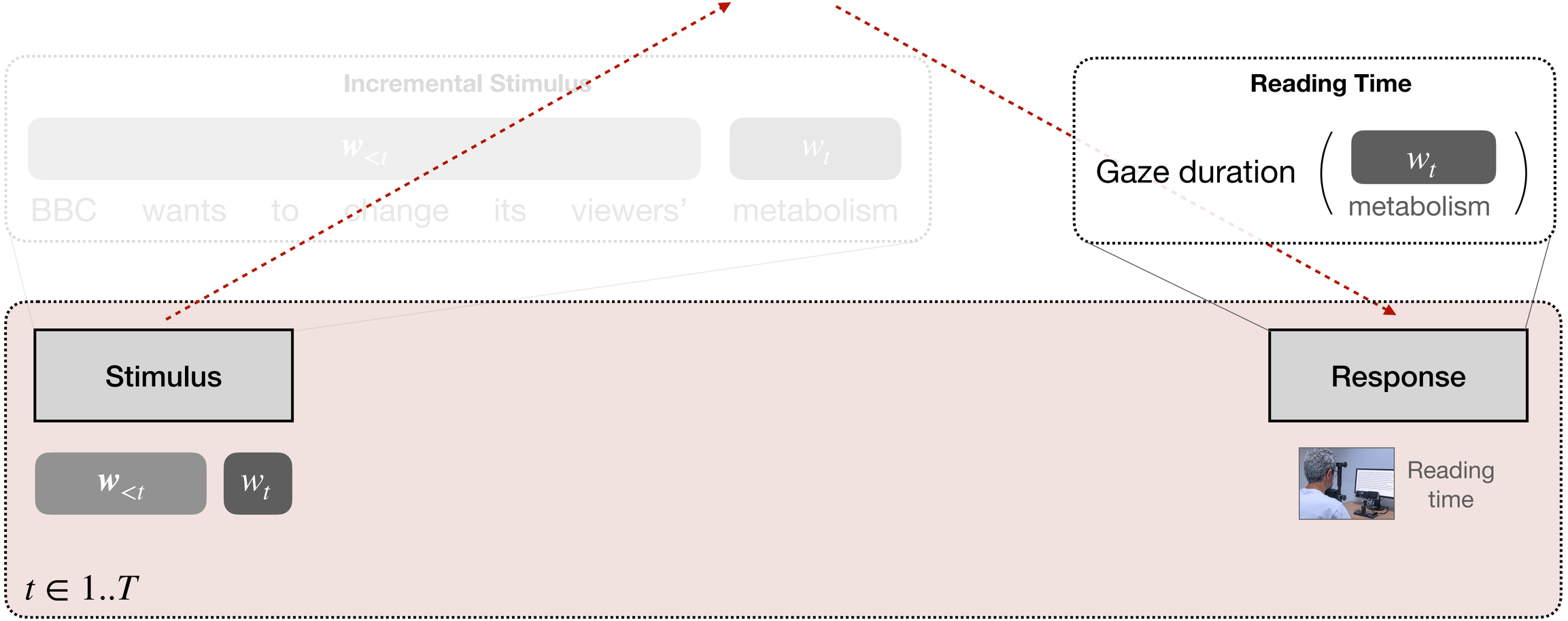
The Surprisal model of incremental comprehension

The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."



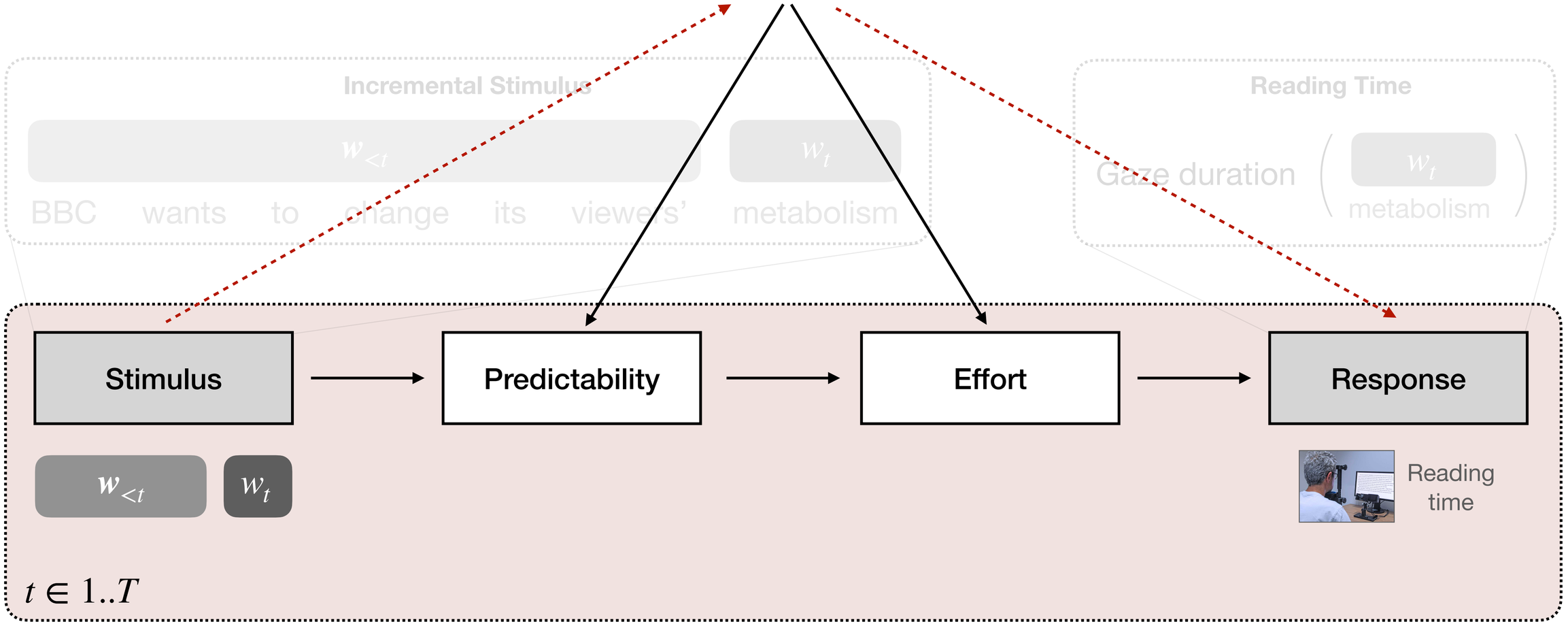
The Surprisal model of incremental comprehension

The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."



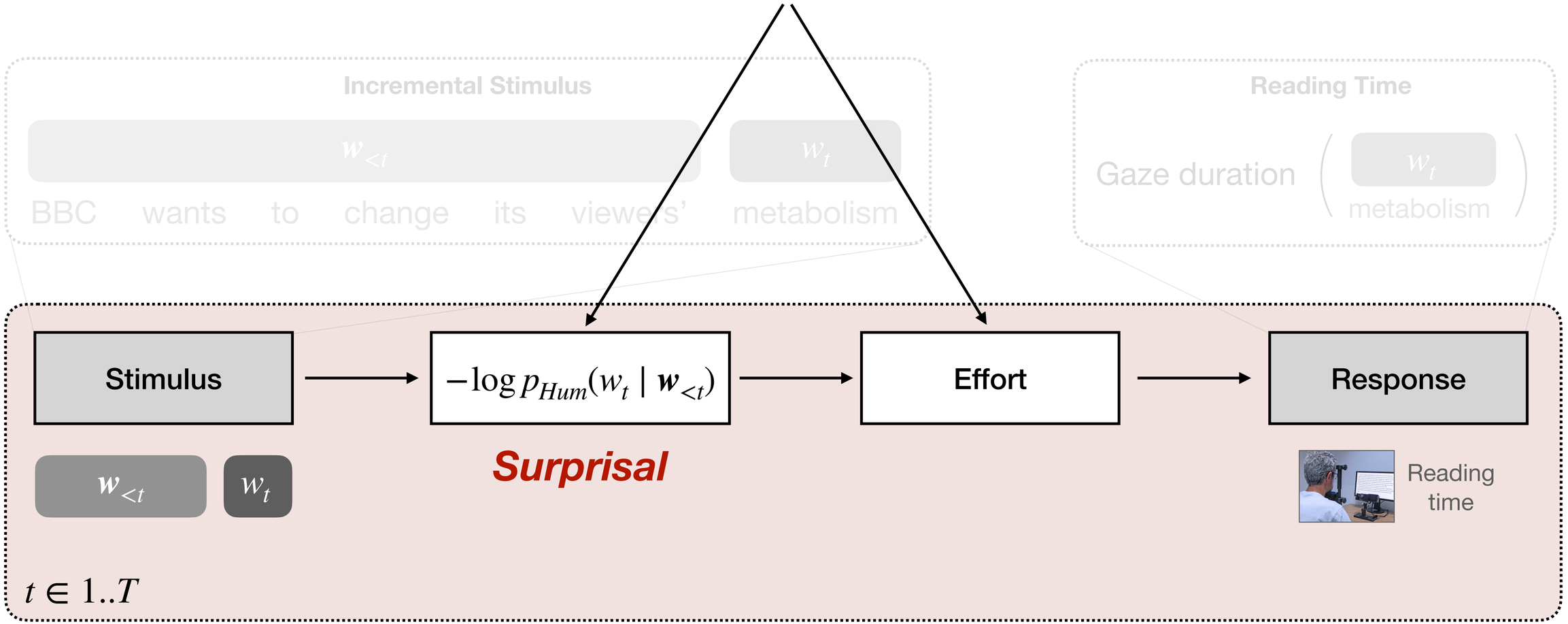
The Surprisal model of incremental comprehension

The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."



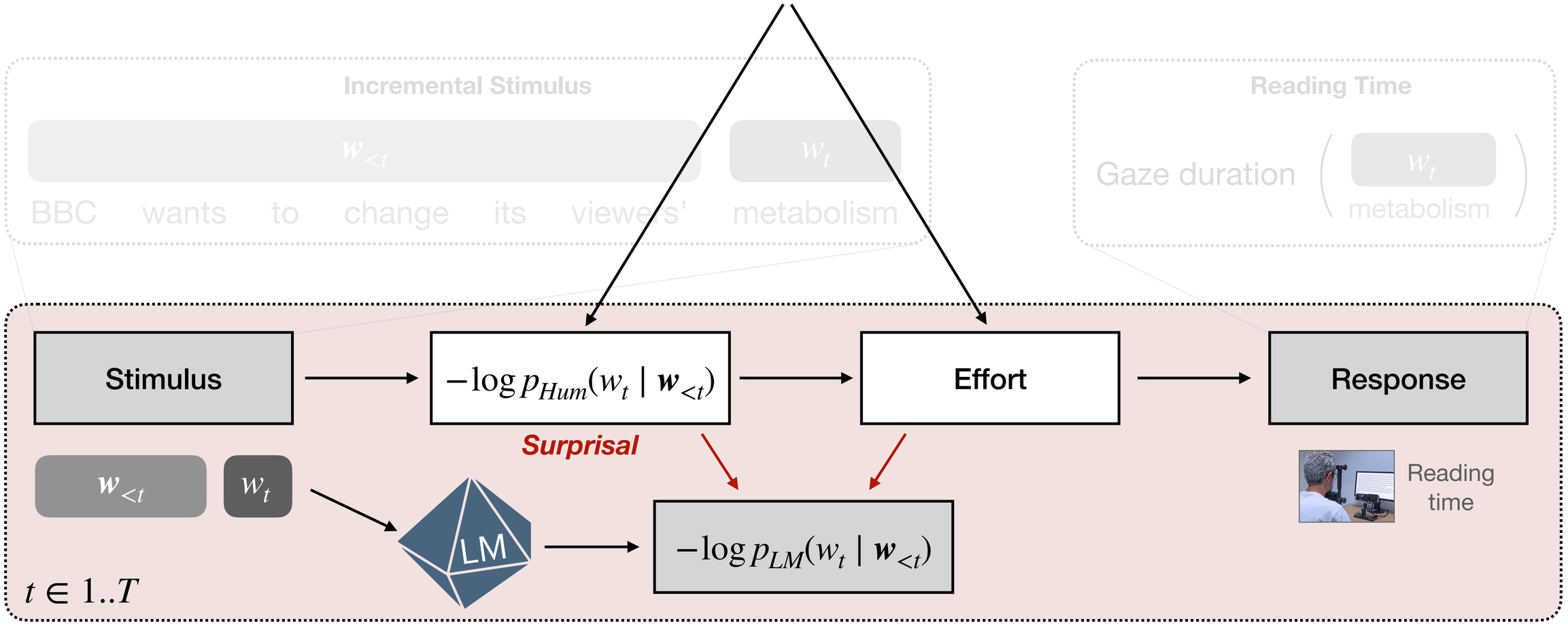
The Surprisal model of incremental comprehension

The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."



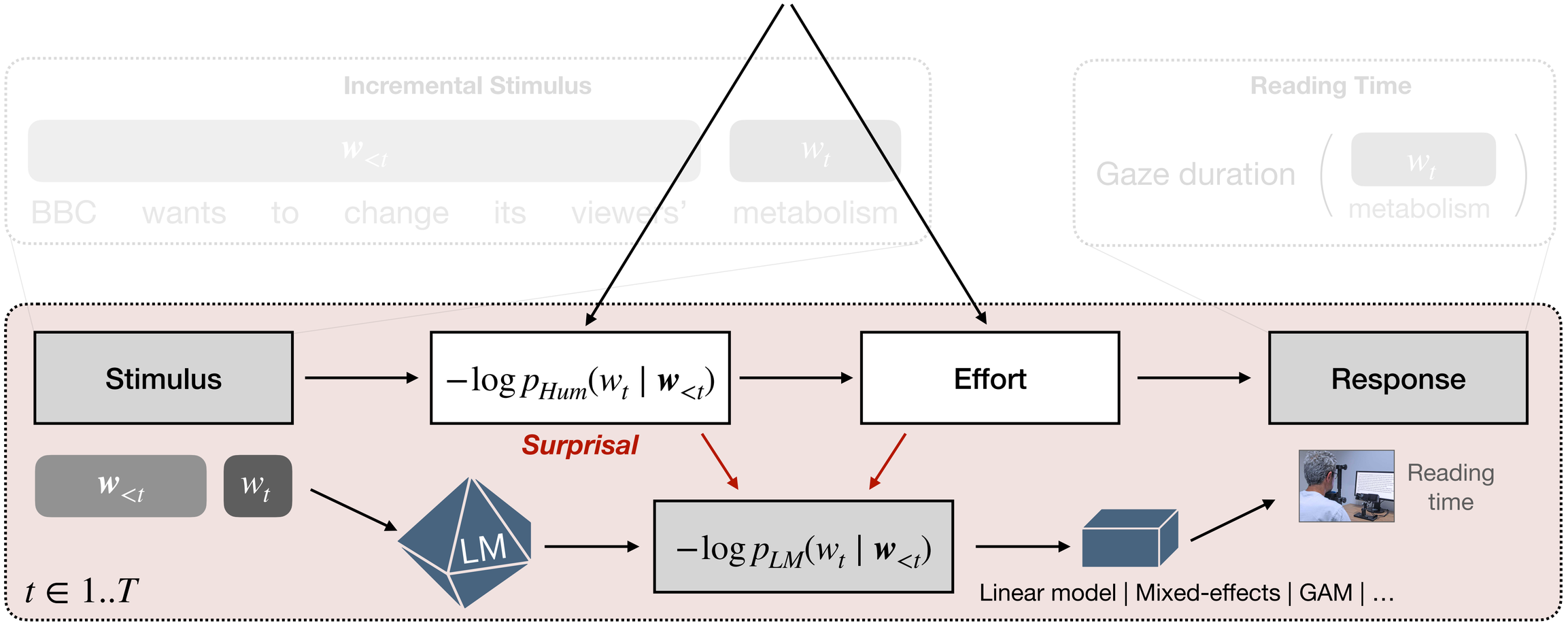
The Surprisal model of incremental comprehension

The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."

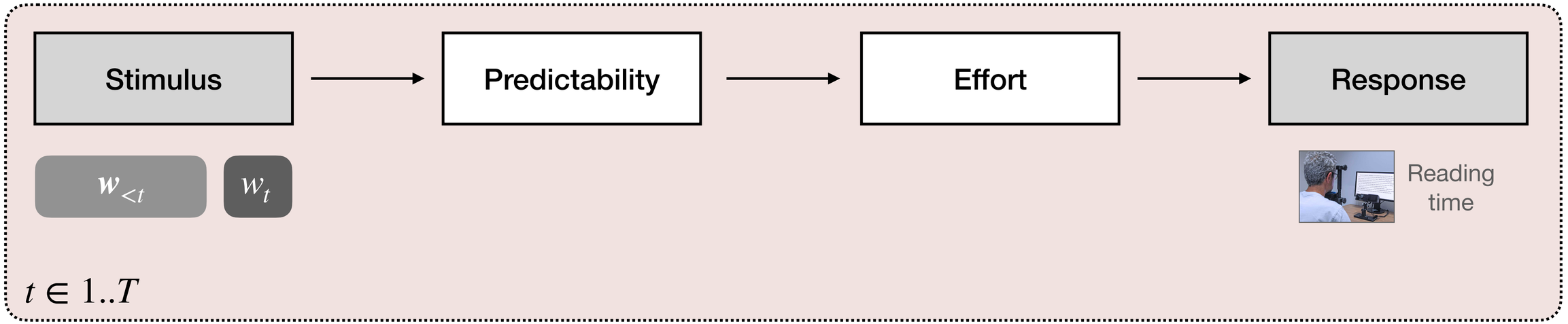


The Surprisal model of incremental comprehension

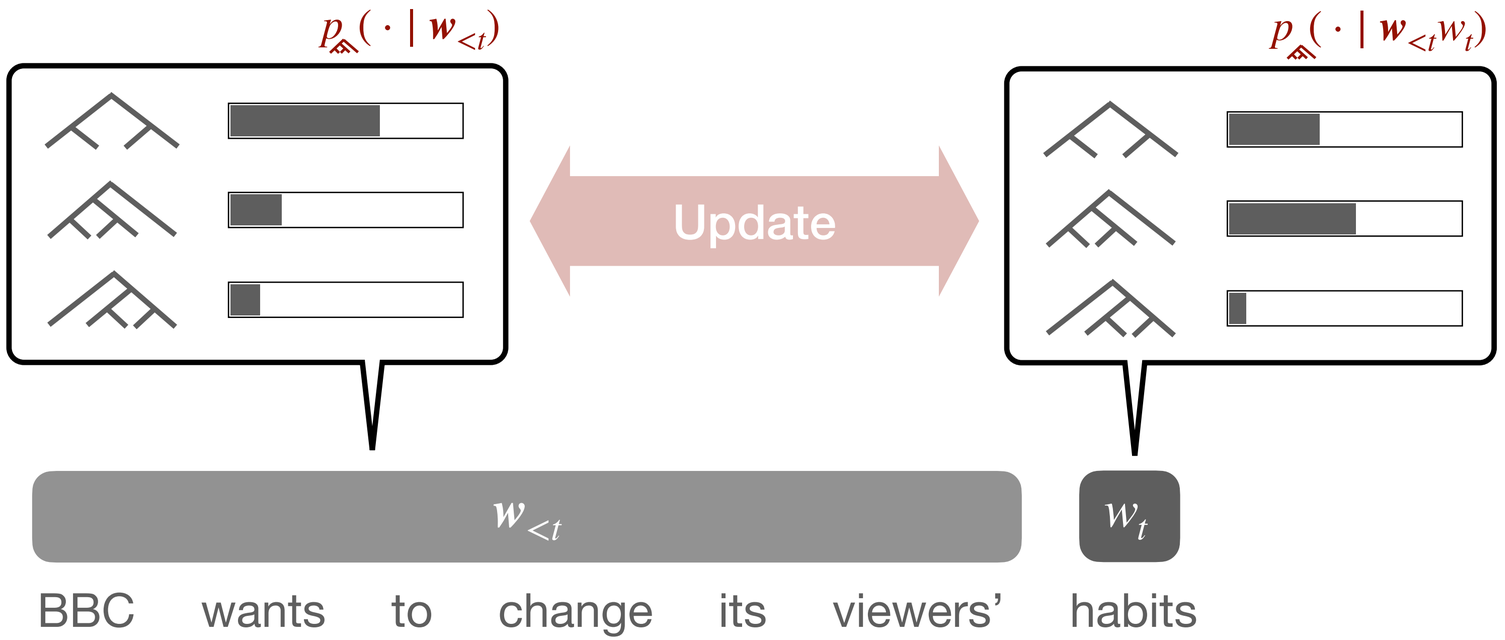
The Surprisal hypothesis: "The effort a reader must spend to comprehend a word is a function of the word's contextual predictability."

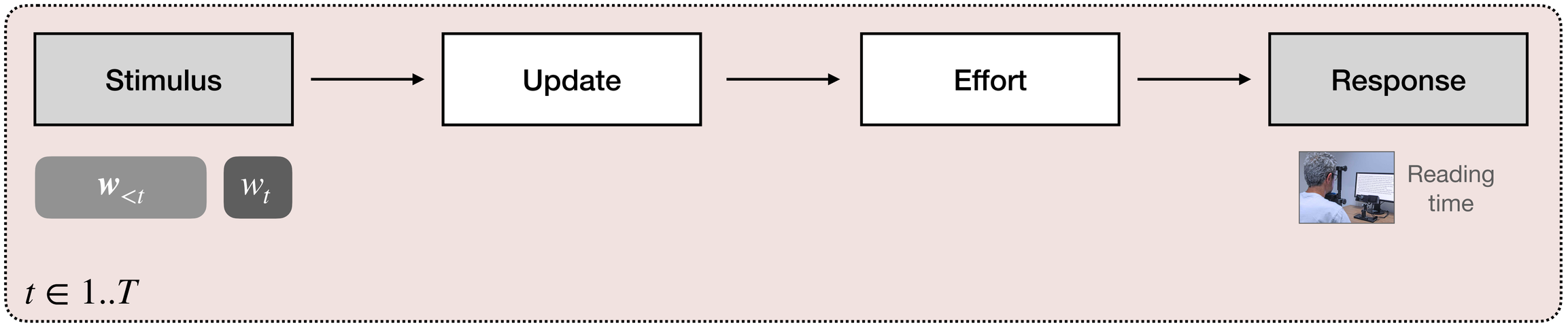


Interpretations of Surprisal

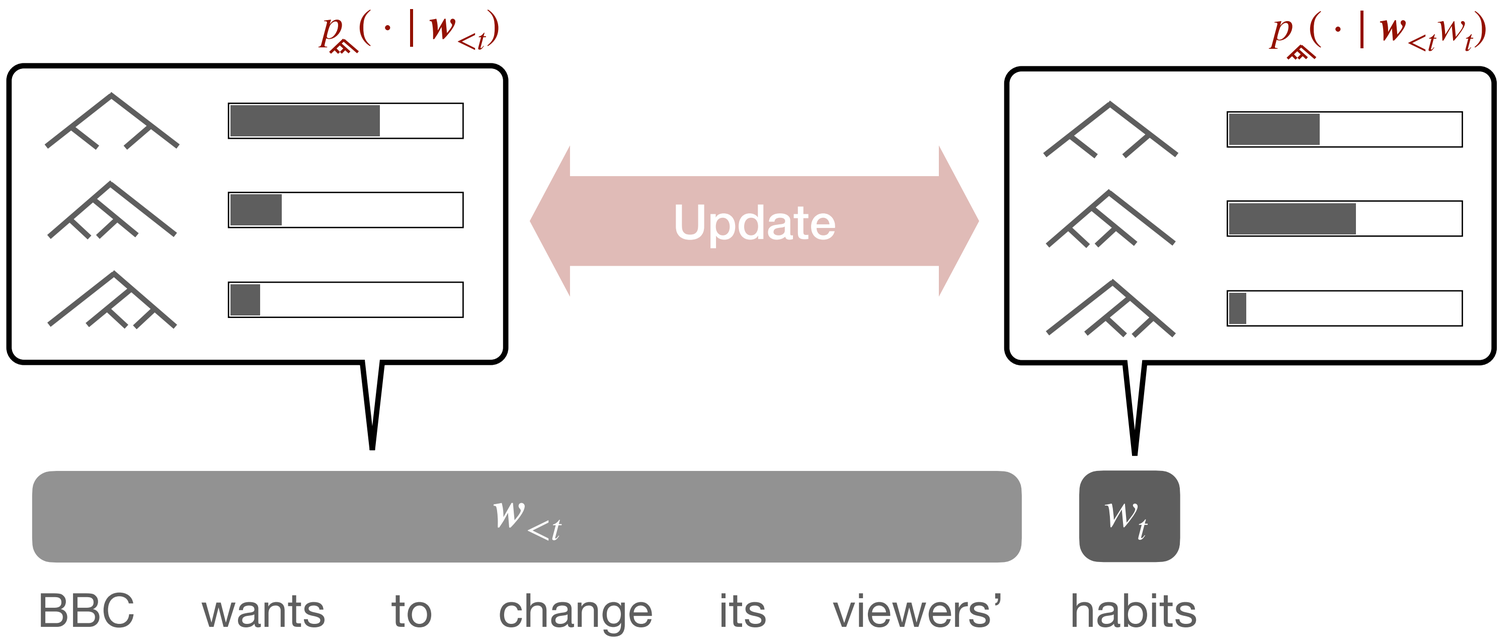


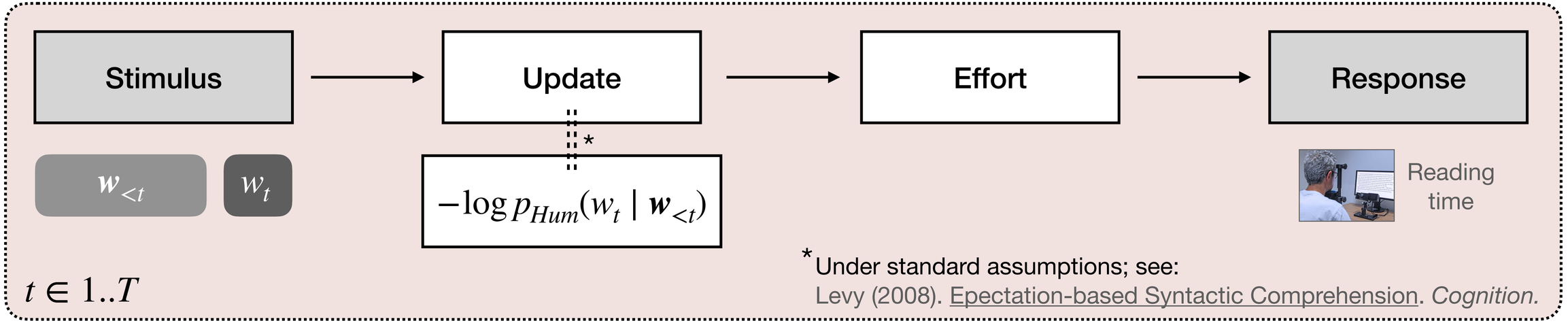
Interpretations of Surprisal





Interpretations of Surprisal

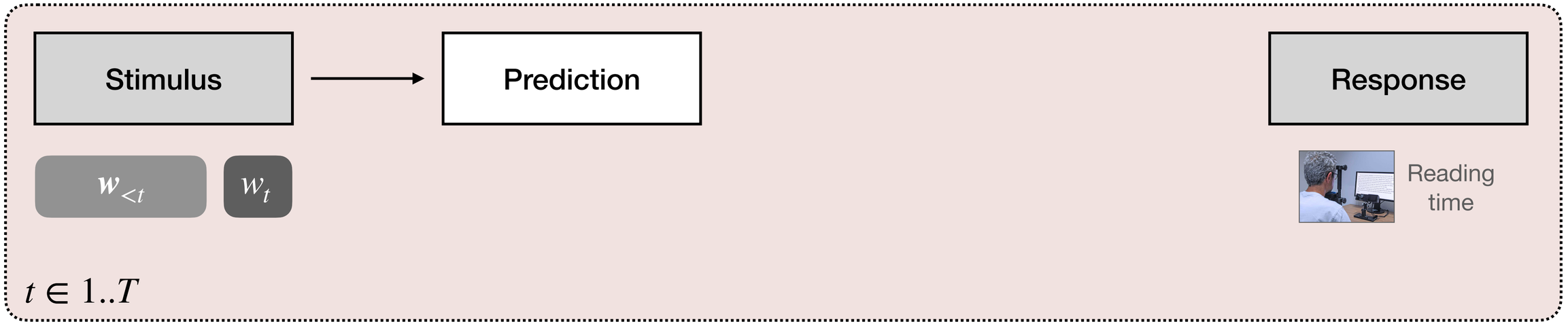




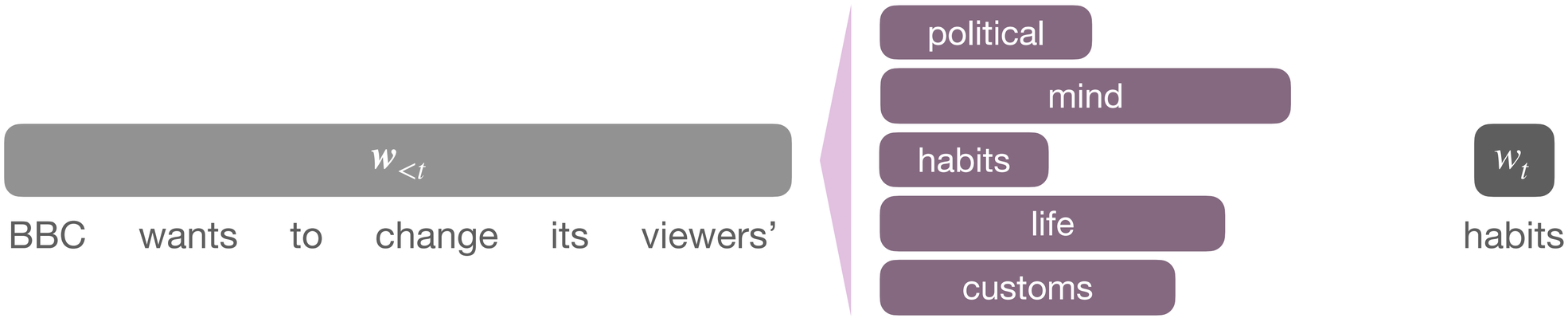
*Under standard assumptions; see:
Levy (2008). Epectation-based Syntactic Comprehension. *Cognition.*

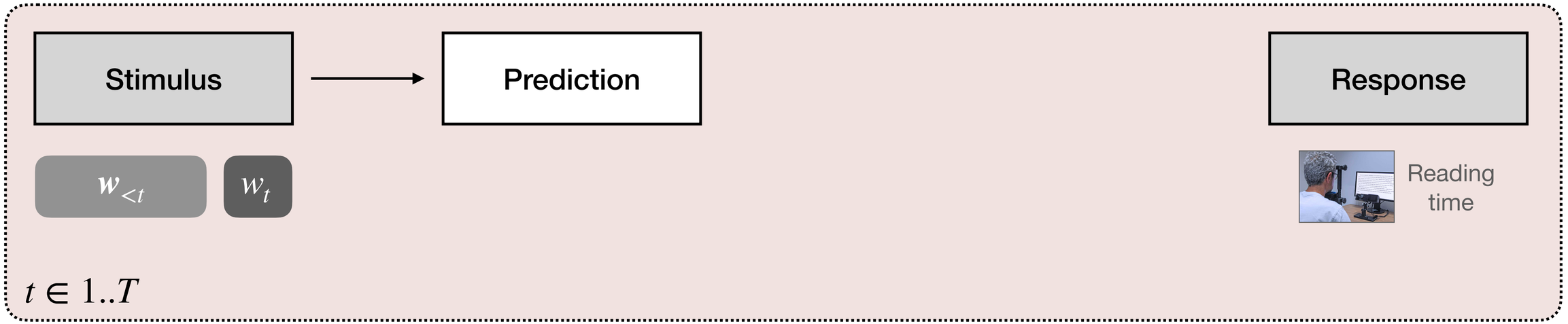
Interpretations of Surprisal



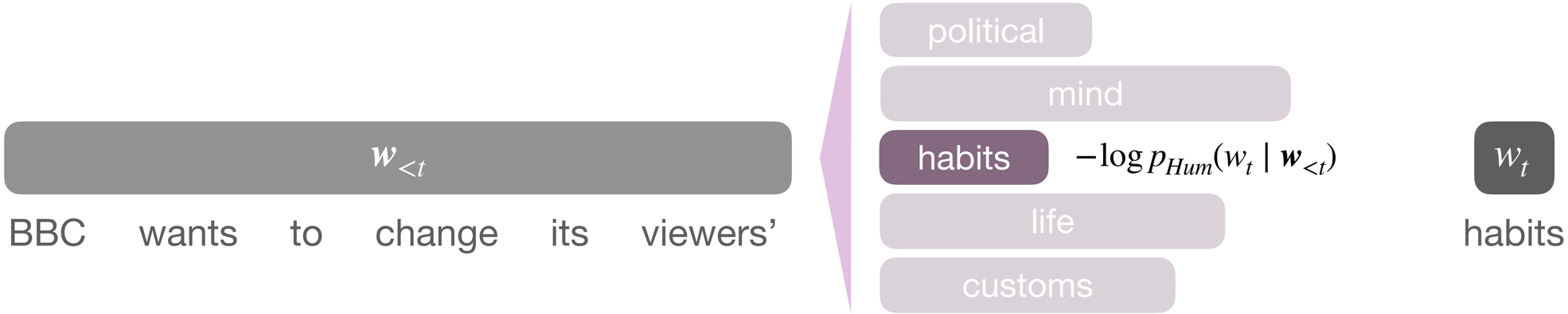


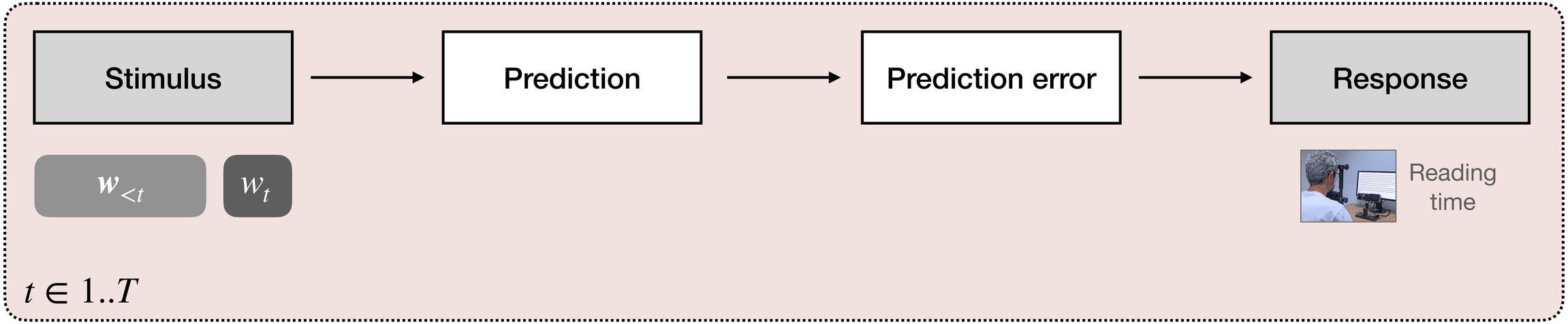
Interpretations of Surprisal



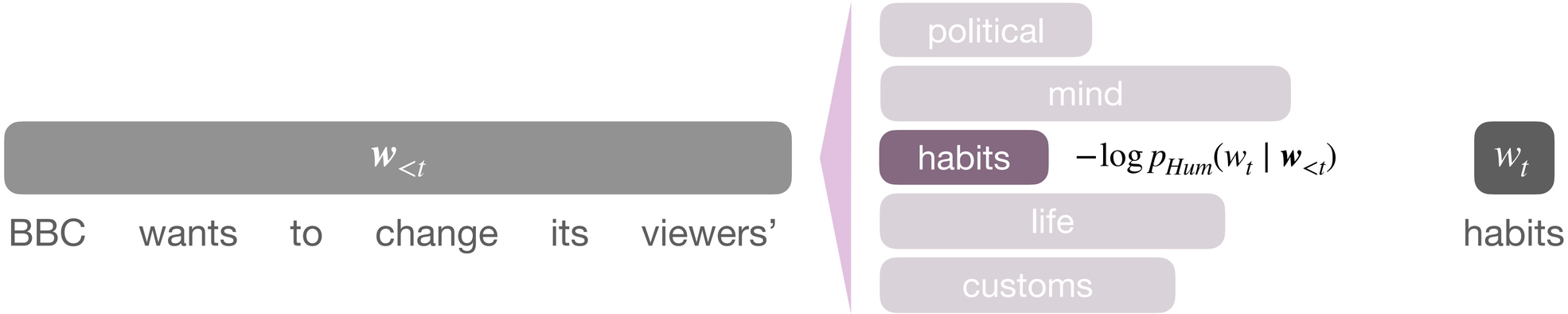


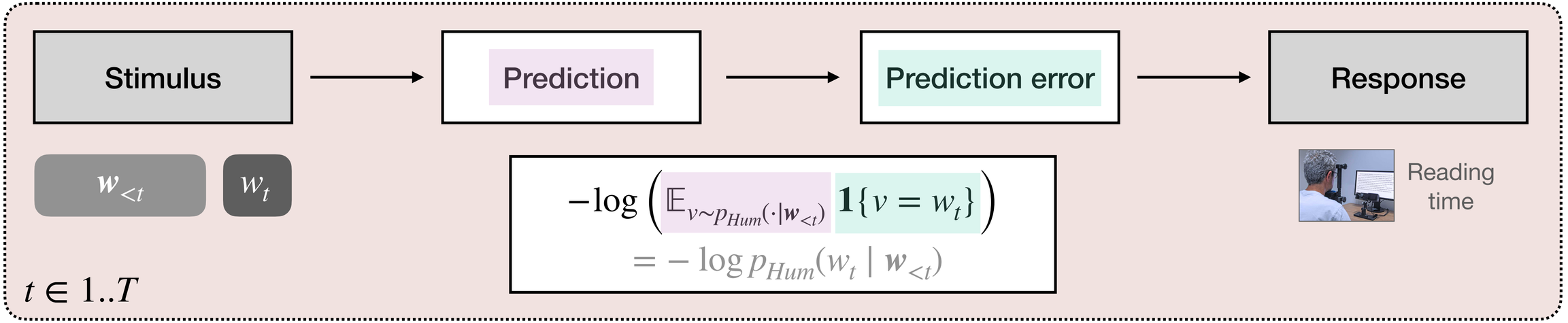
Interpretations of Surprisal



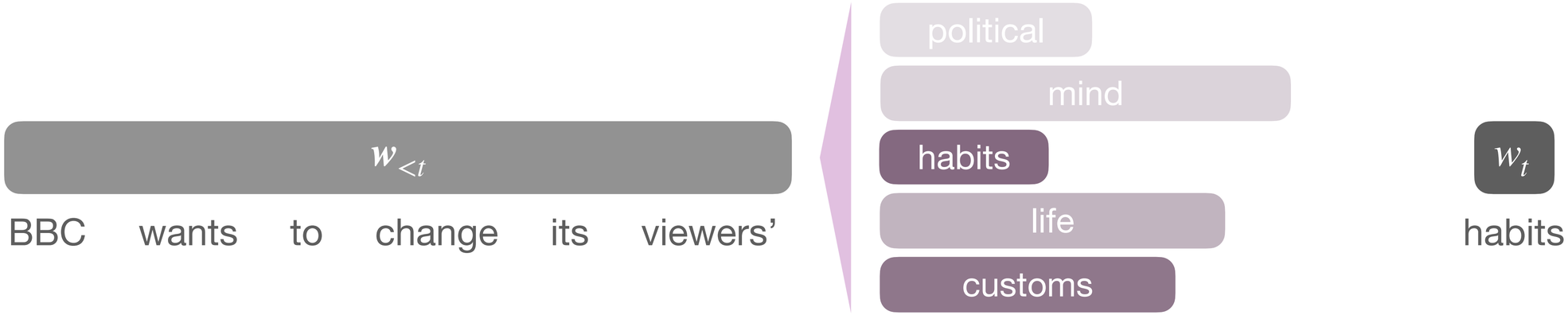


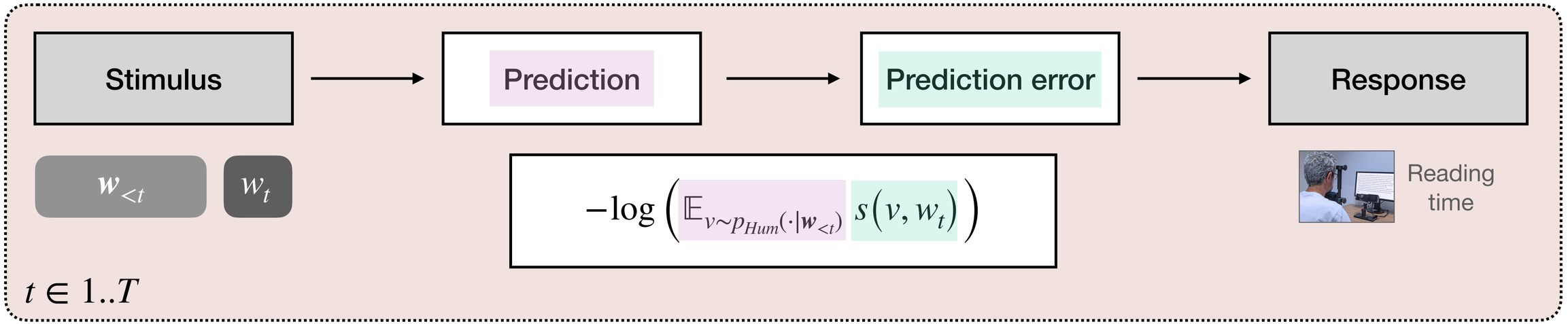
Interpretations of Surprisal



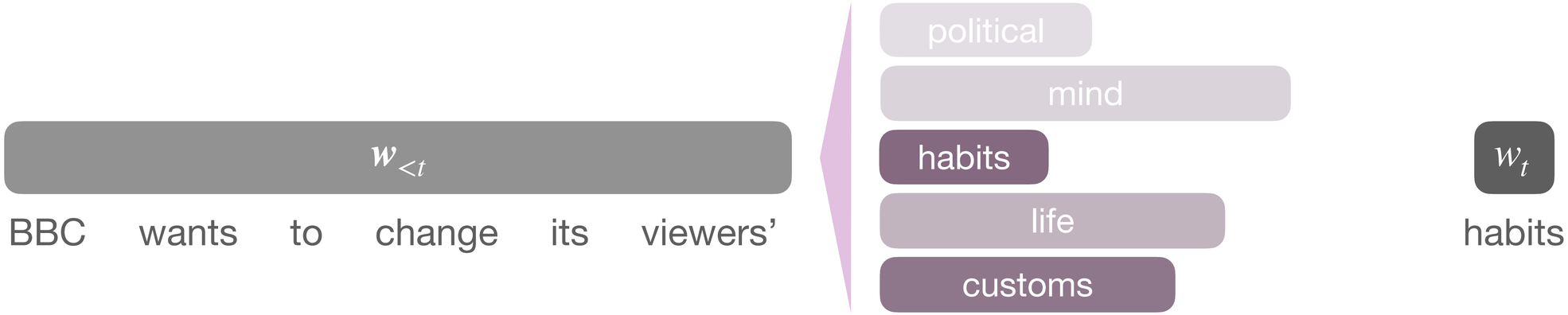


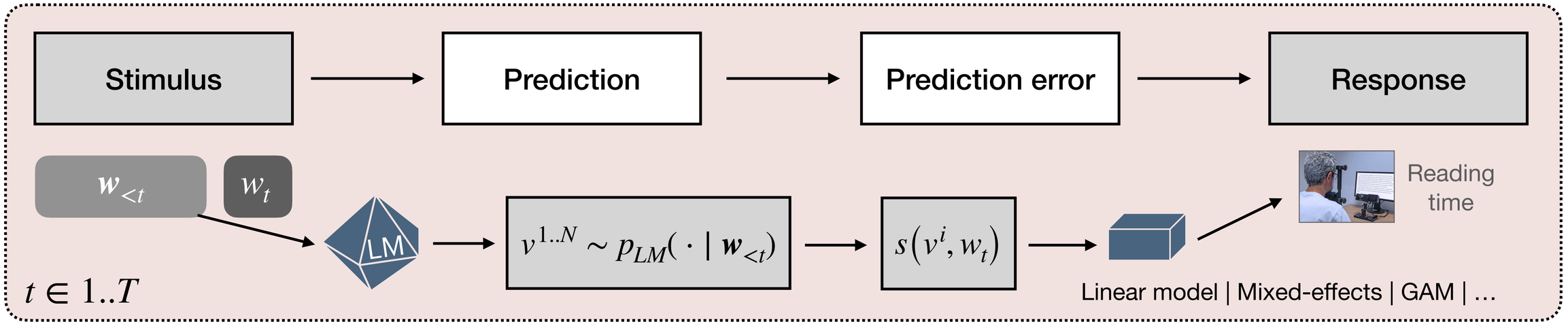
Beyond Surprisal



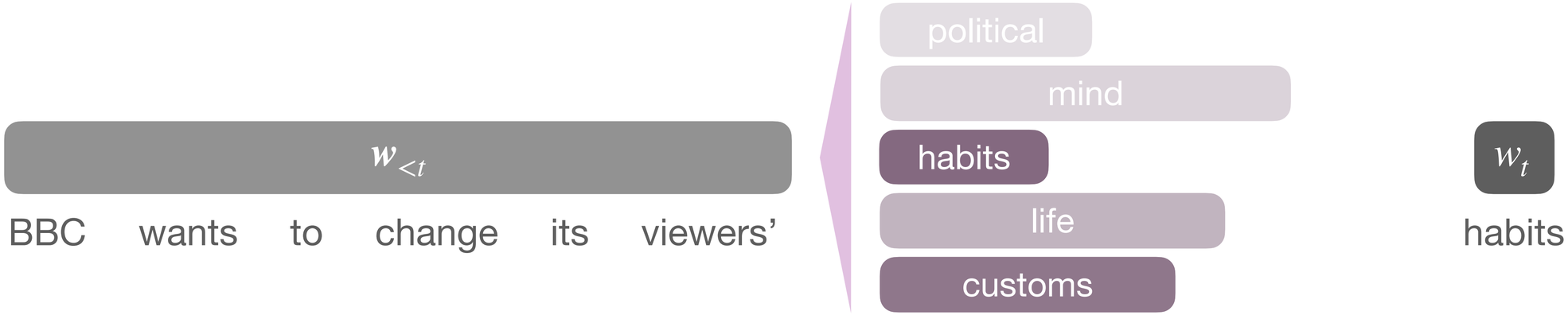


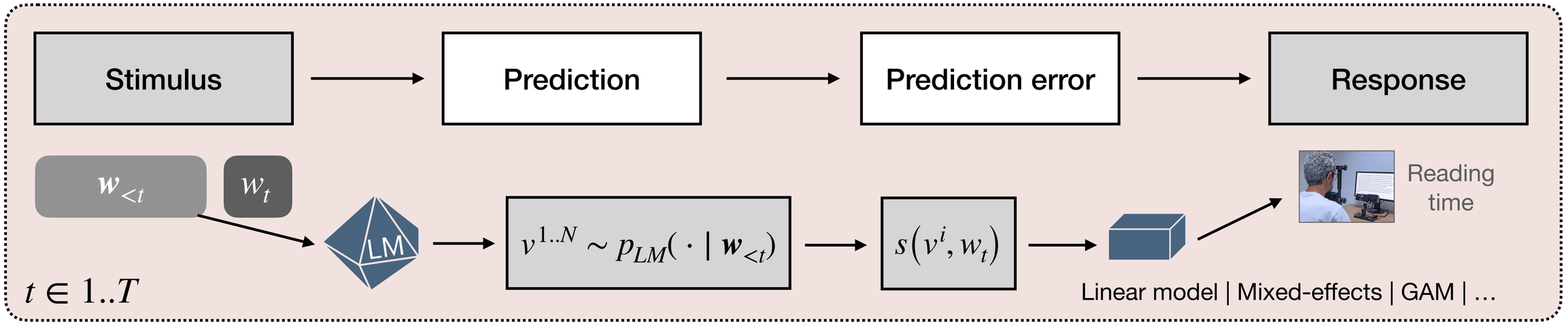
Beyond Surprisal





Beyond Surprisal





Beyond Surprisal

$s(v^i, w_t)$ Semantic

$s(v^i, w_t)$ Syntactic

$w_{<t}$

BBC wants to change its viewers'

political

mind

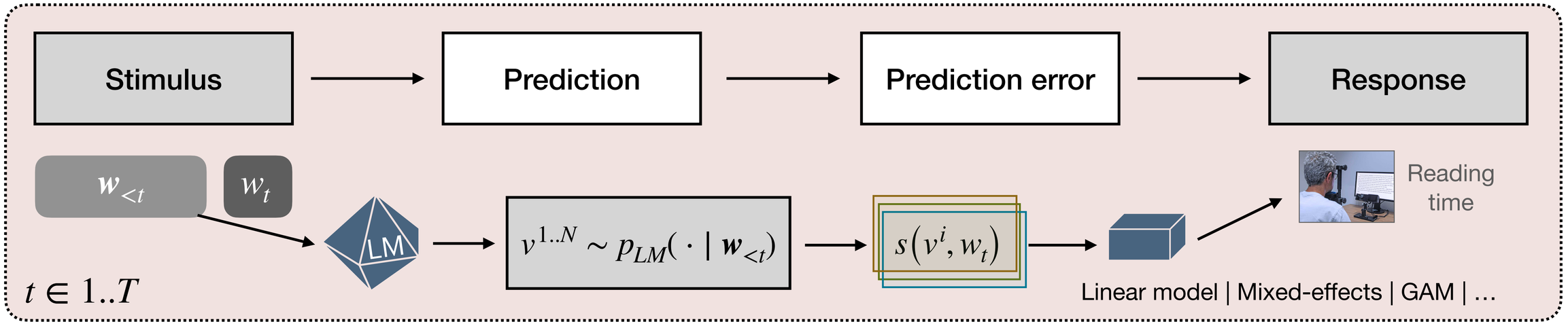
habits

life

customs

w_t

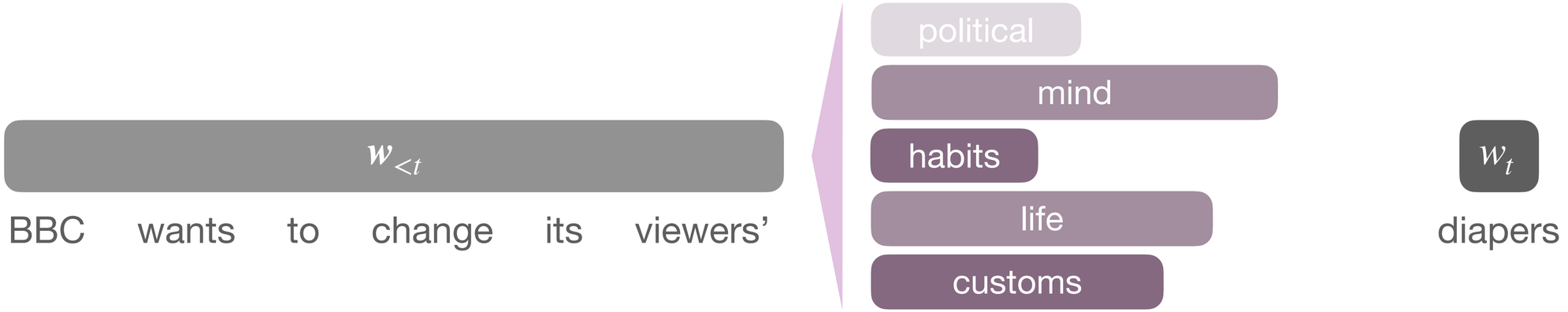
diapers

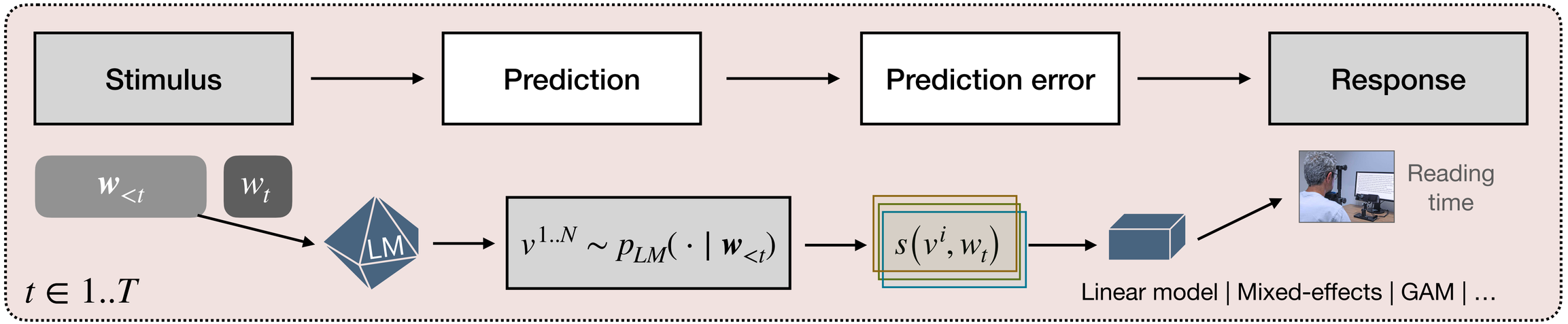


Beyond Surprisal

$s(v^i, w_t)$ Semantic

$s(v^i, w_t)$ Syntactic

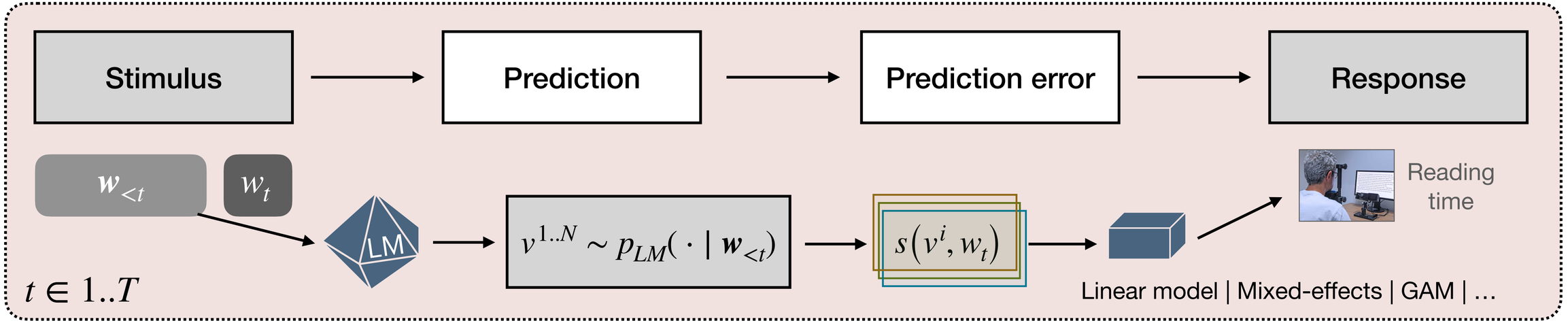




	Similarity-adjusted surprisal			
	Non-contextual	Contextual	POS	Orthographic
Brown	2.78***	-0.02	2.22	1.29*
Dundee	1.44***	0.01	0.69	0.64*
Natural Stories	3.18***	0.31**	1.04*	0.79*
Provo	1.34***	0.05	1.49	0.82

$s(v^i, w_t)$

Δ_{LogLik} with respect to baseline model; predictors for current and previous three words.



Giulianelli et al. (2023). Information Value: Measuring Utterance Predictability as Distance from Plausible Alternatives. *EMNLP.*

Giulianelli et al. (2024). Generalized Measures of Anticipation and Responsivity in Online Language Processing. *ACL Findings.*

Meister et al. (2024). Towards a Similarity-adjusted Surprisal Theory. *EMNLP.*

Giulianelli et al. (2026). Incremental alternative sampling as a lens into the temporal and representational resolution of linguistic prediction. *JML.*

Beyond Surprisal

$s(v^i, w_t)$ Semantic

$s(v^i, w_t)$ Syntactic

$w_{<t}$

BBC wants to change its viewers'

political

mind

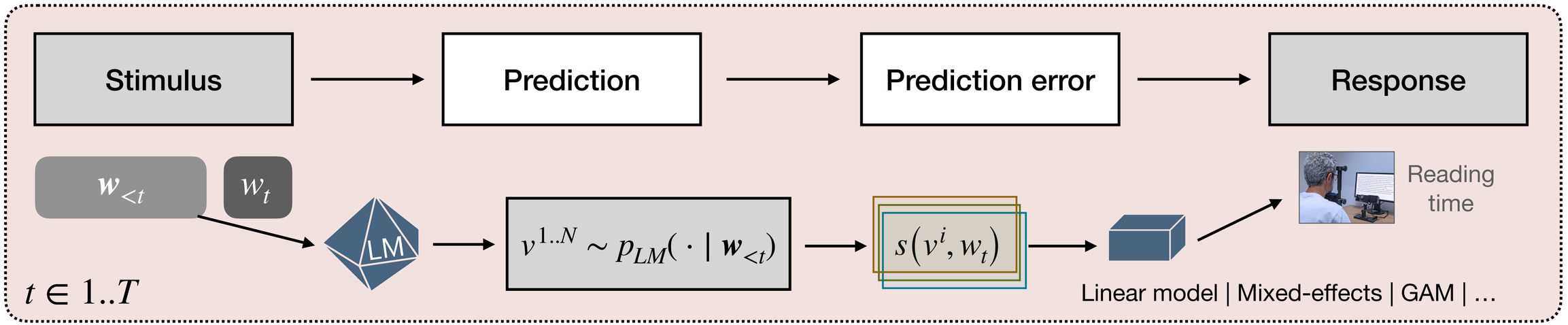
habits

life

customs

w_t

diapers



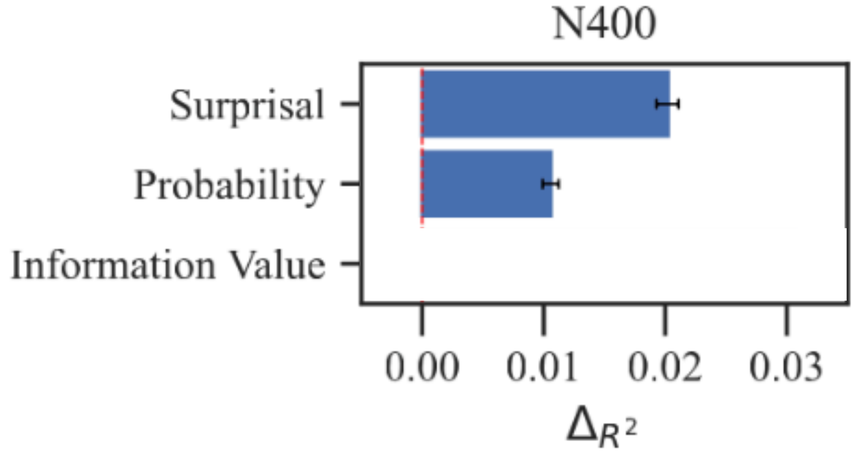
Generalised Surprisal

The **generalised surprisal** of a string $\boldsymbol{w}$ in a context $\boldsymbol{c}$ is

$$\iota_p^{(\mathrm{f},\mathrm{g})}(\boldsymbol{w};\boldsymbol{c}) := \mathrm{f}\left(\mathbb{E}_{\boldsymbol{v}\sim p(\cdot|\boldsymbol{c})}\ \mathrm{s}(\boldsymbol{v},\boldsymbol{w},\boldsymbol{c})\right)$$

where:

- $\mathrm{f} : \mathbb{R} \to \mathbb{R}$ is a **warping function**
- $\mathrm{s} : \Sigma^* \times \Sigma^* \times \Sigma^* \to \mathbb{R}$ is a **scoring function**



Language model: GPT-2 Small
Stimuli: *N* = 1726 from English novels (de Varda et al. 2023)
Response: N400 (avg. across participants)

Next-word Probability	Next-word Surprisal	Next-word Information Value
$\mathrm{f}(x) = x$	$\mathrm{f}(x) = -\log(x)$	$\mathrm{f}(x) = x$
$\mathrm{g}(v, w_t, \boldsymbol{w}_{<t}) = \mathbf{1}\{v = w_t\}$	$\mathrm{g}(v, w_t, \boldsymbol{w}_{<t}) = \mathbf{1}\{v = w_t\}$	$\mathrm{g}(v, w_t, \boldsymbol{w}_{<t}) = d_{\boldsymbol{w}_{<t}}(v, w_t)$

Generalised Surprisal

Most predictive uncertainty measure varies by response type

ELAN	LAN	N400	EPNP	P600	PNP	First-fixation RT	First-pass RT	Right-bounded RT	Self-paced RT
	Information value	Information value	Surprisal	Surprisal	Probability	Surprisal	Surprisal	Surprisal	Surprisal
(Sequence) Entropy	Information value	Information value	(Sequence) Entropy	(Sequence) Entropy	Exp. Next-symbol Probability	Exp. Next-symbol Information Value	Exp. Next-symbol Information Value	Exp. Next-symbol Information Value	(Sequence) Entropy

$\mathcal{M}$ = (p_{LM}, sampling procedure, **warping function**, **scoring function**, **anticipatory/responsive**)

What do language processors predict?

She took his hand and...

?

next word semantics

intents

sentence structure

possible worlds

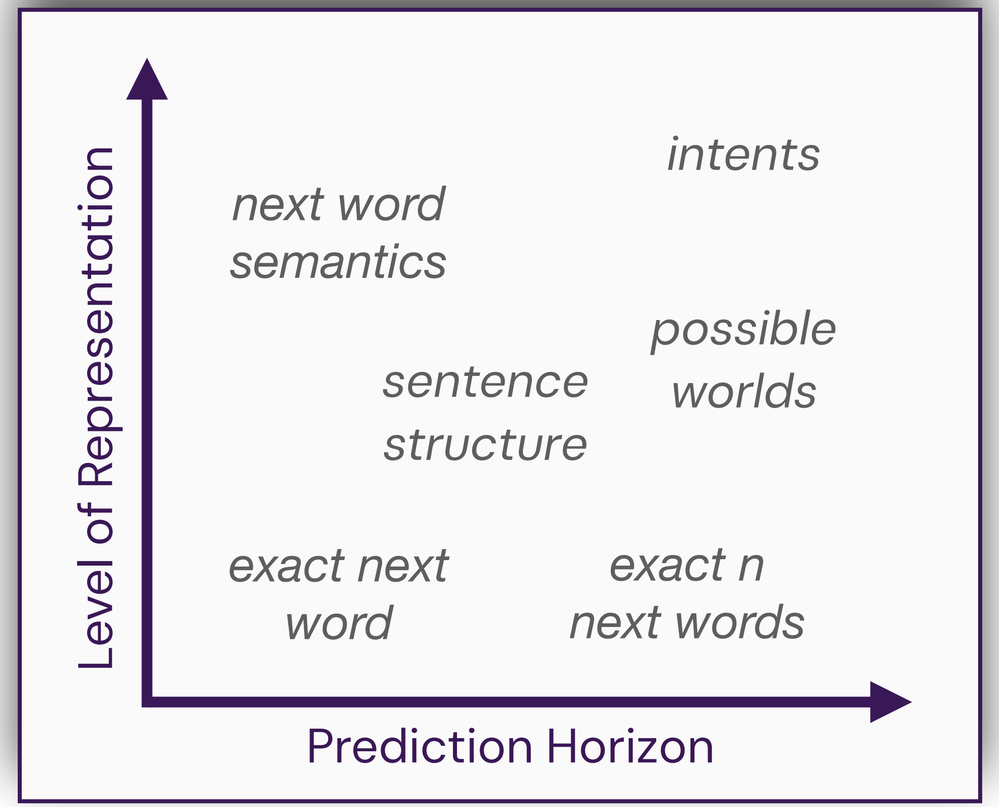
exact next word

exact n next words

What do language processors predict?

She took his hand and...

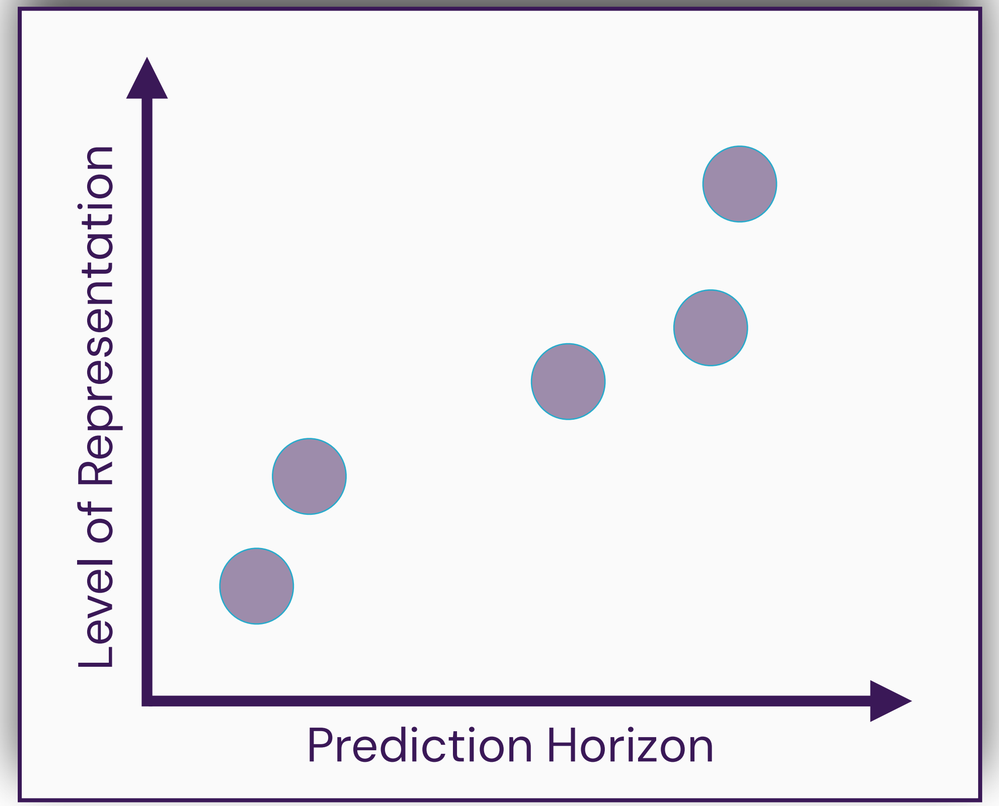
?



What do language processors predict?

She took his hand and...

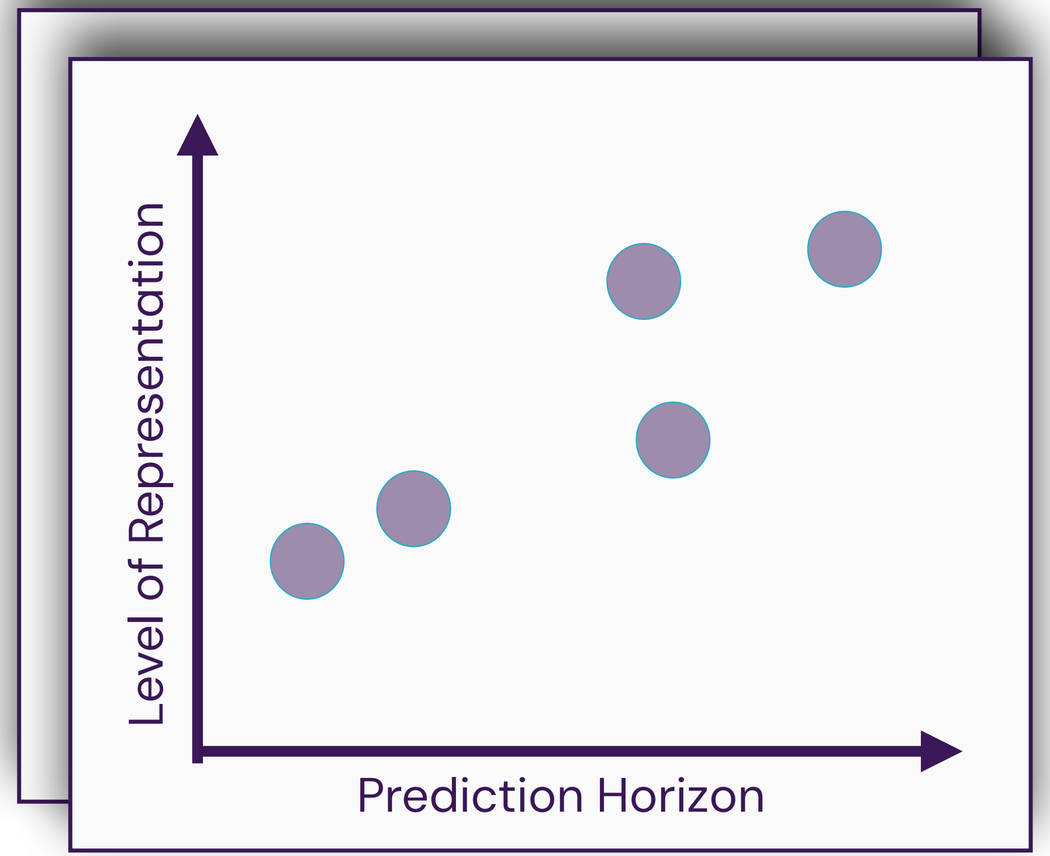
?



How are predictions updated during incremental processing?

She took his hand and

She took his hand and said…



Incremental alternative sampling

She took his hand and...



shook it

kissed it fervently

he tried not to shudder in relief

LM as generator

of contextual alternatives

Incremental alternative sampling

She took his hand and...



shook it

kissed it fervently

he tried not to shudder in relief

She took his hand and said...



it would be alright

that she loved him

nothing for a moment

LM as generator

of contextual alternatives

Incremental alternative sampling

$h = 1$

She took his hand and...



shook t
kissed it fervently
he tried not to shudder in relief



Forecast horizon

$h \in 1..10$

She took his hand and said...



it would be alright
that she loved him
nothing for a moment

LM as generator

of contextual alternatives

Incremental alternative sampling

Forecast horizon

$h \in 1..10$

She took his hand and...



shook it

kissed it fervently

he tried not to shudder in relief



She took his hand and said...



it would be alright

that she loved him

nothing for a moment

LM as generator

of contextual alternatives

LM as representation function

(activations at layer ℓ, prediction horizon h)

Incremental alternative sampling

Forecast horizon

$h \in 1..10$

She took his hand and...

LM

shook it

kissed it fervently

he tried not to shudder in relief

She took his hand and said...

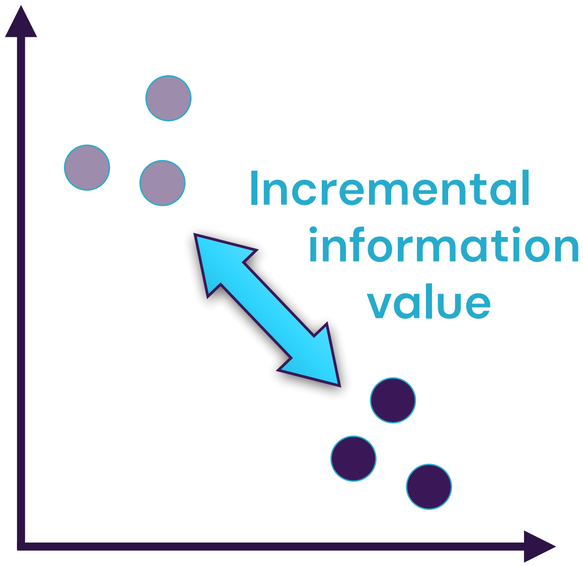
LM

it would be alright

that she loved him

nothing for a moment

LM



LM as generator

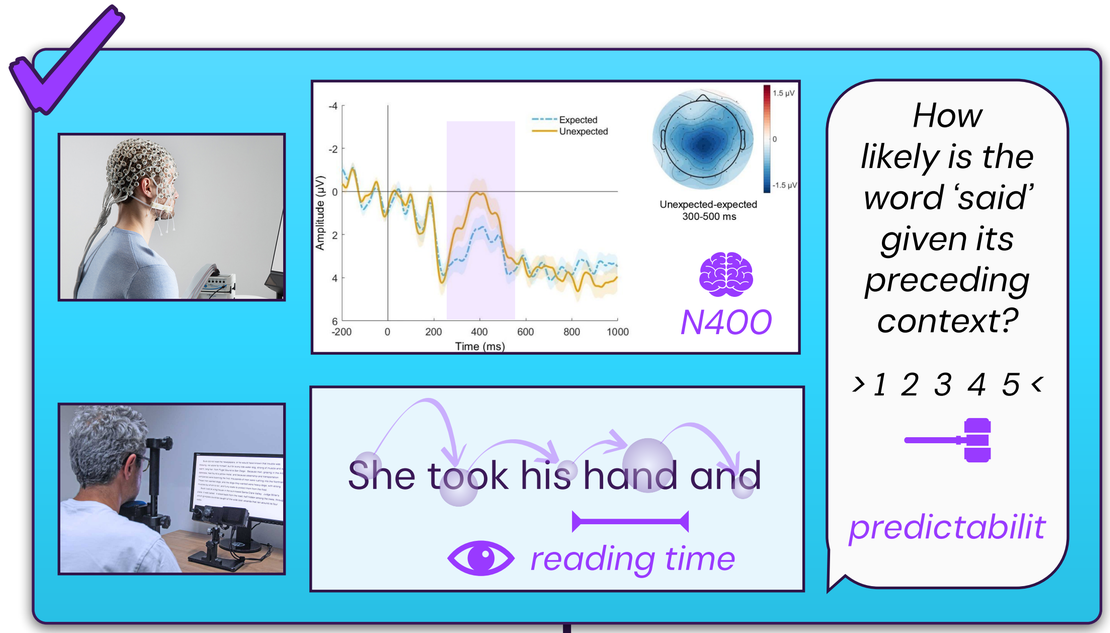
of contextual alternatives

LM as representation function

(activations at layer ℓ, prediction horizon h)

Incremental alternative sampling

Giulianelli, Wallbridge, Cotterell, Fernández. (2026). Incremental alternative sampling as a lens into the temporal and representational resolution of linguistic prediction. *Journal of Memory and Language.*



She took his hand and...

LM

shook it

kissed it fervently

he tried not to shudder in relief

She took his hand and said...

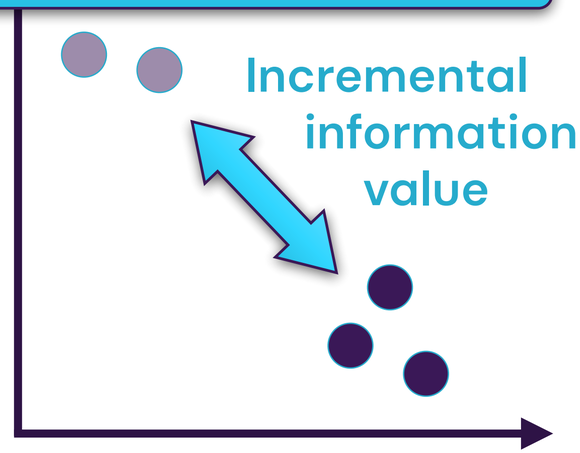
LM

it would be alright

that she loved him

nothing for a moment

LM



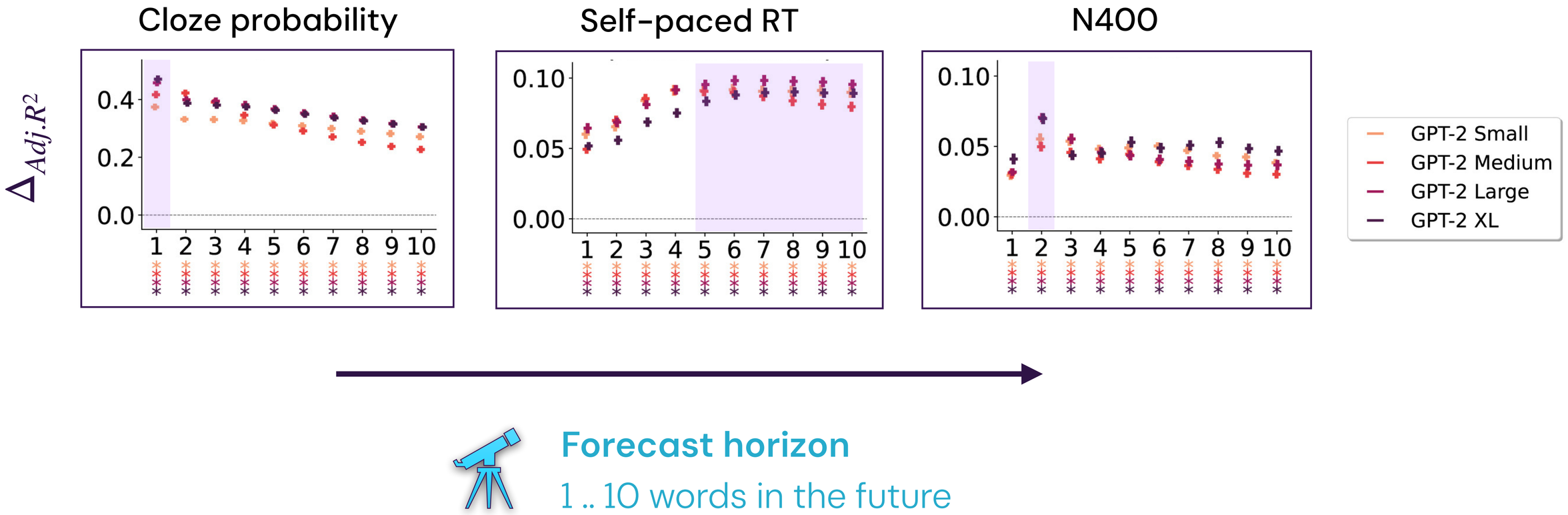
LM as generator

of contextual alternatives

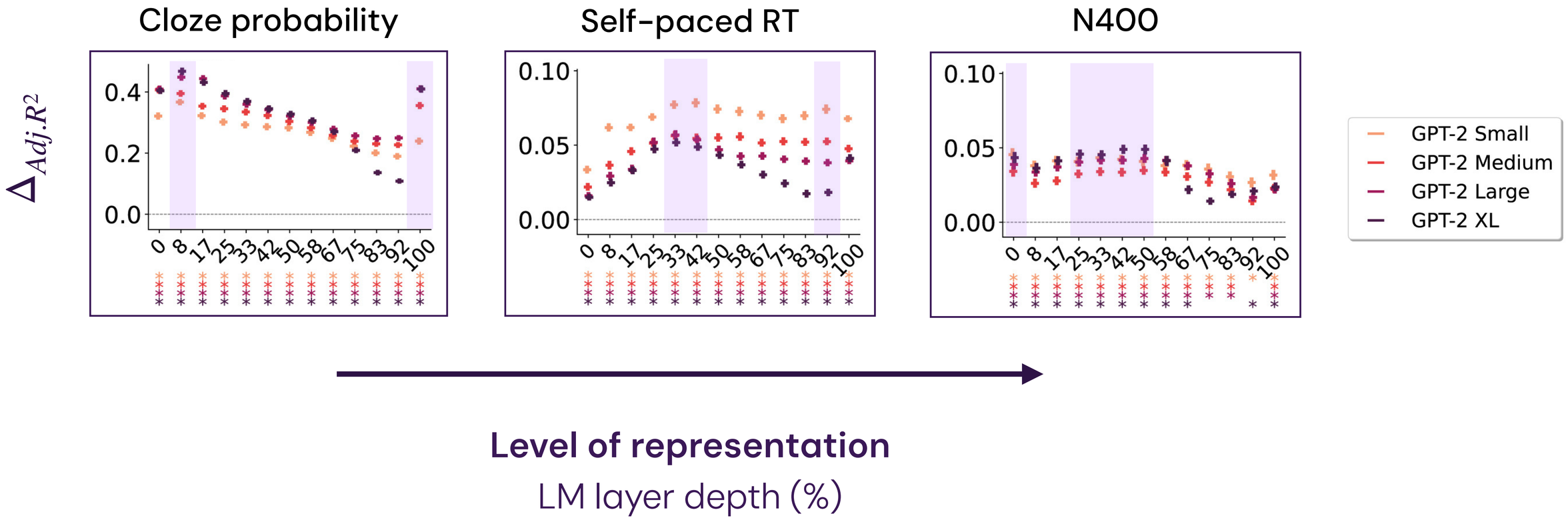
LM as representation function

(activations at layer ℓ, prediction horizon h)

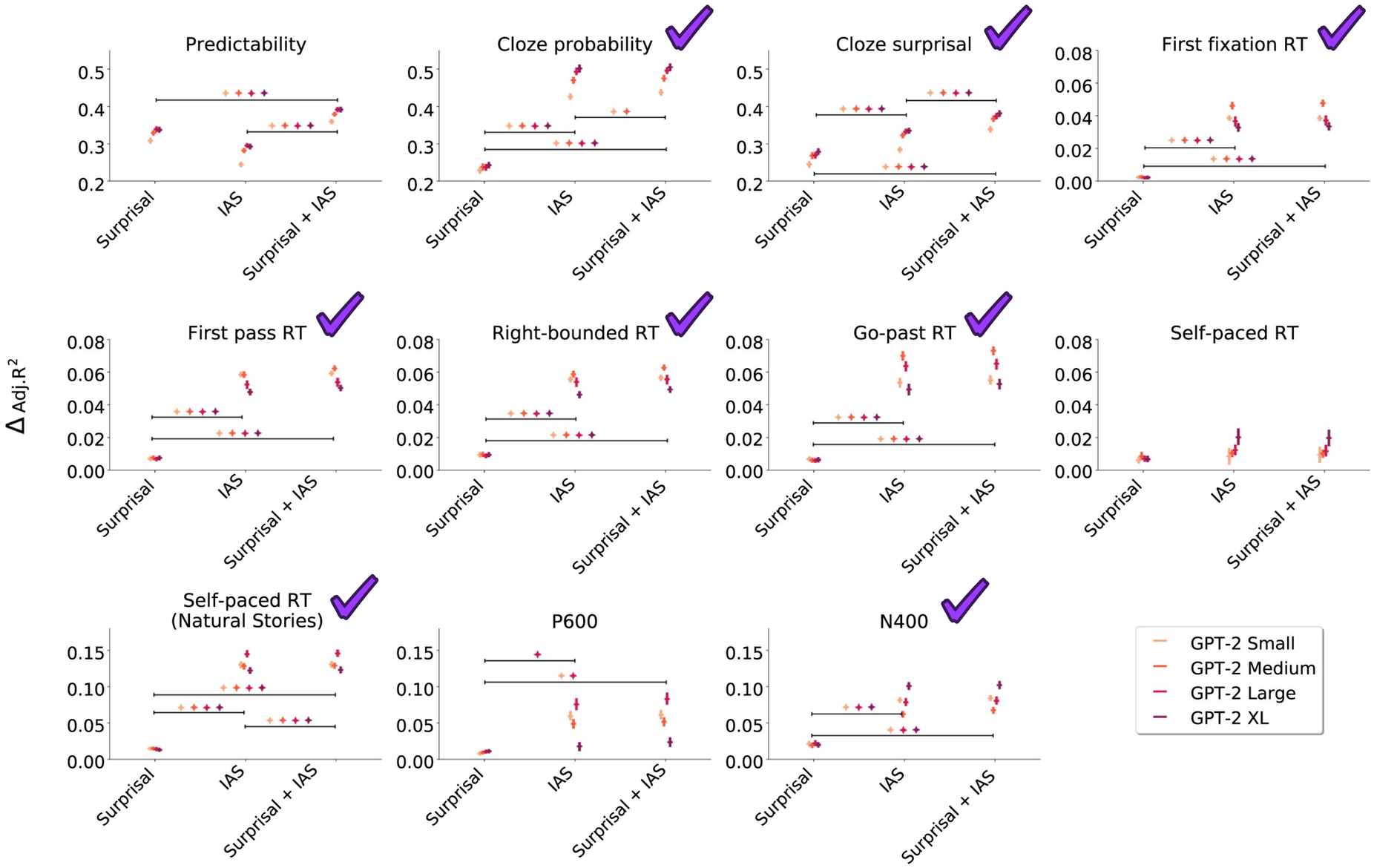
IAS predicts behavioural and neural responses to incremental input



IAS predicts behavioural and neural responses to incremental input



Stronger predictive power than Surprisal

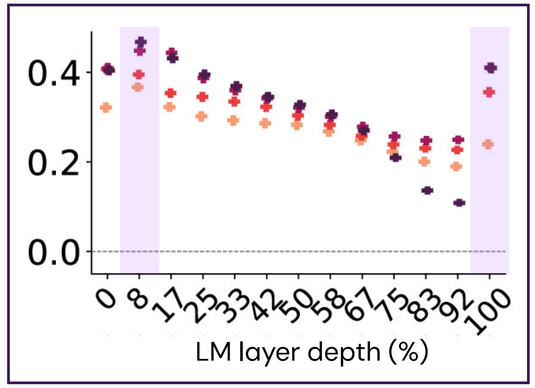


Incremental language comprehension engages prediction across **multiple temporal and representational resolutions**.

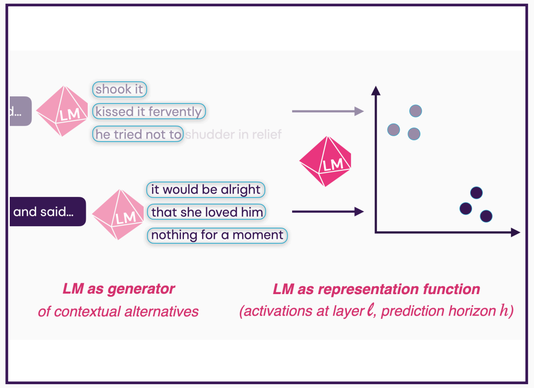
Recap - Incremental alternative sampling



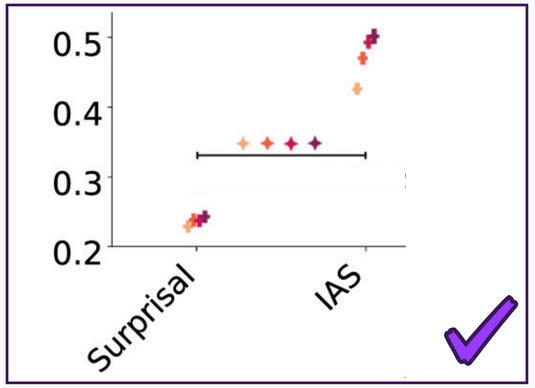
Incremental Alternative Sampling is a **computational-level model** of language comprehension.



IAS links specific measurements to the **representational levels** and **forecast horizons** that best explain them.

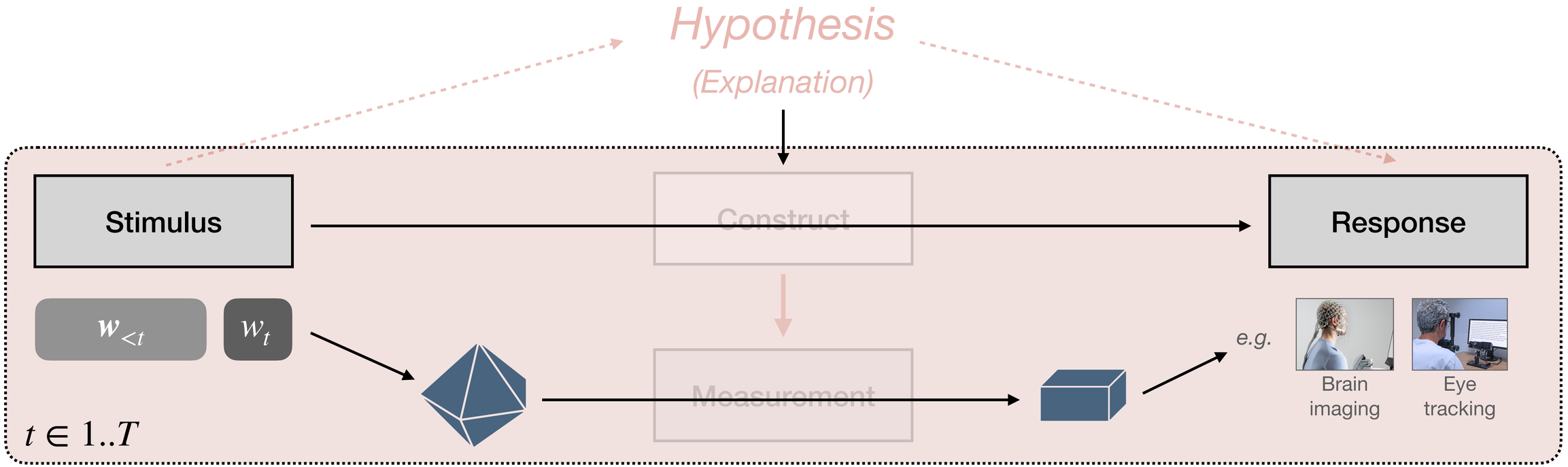


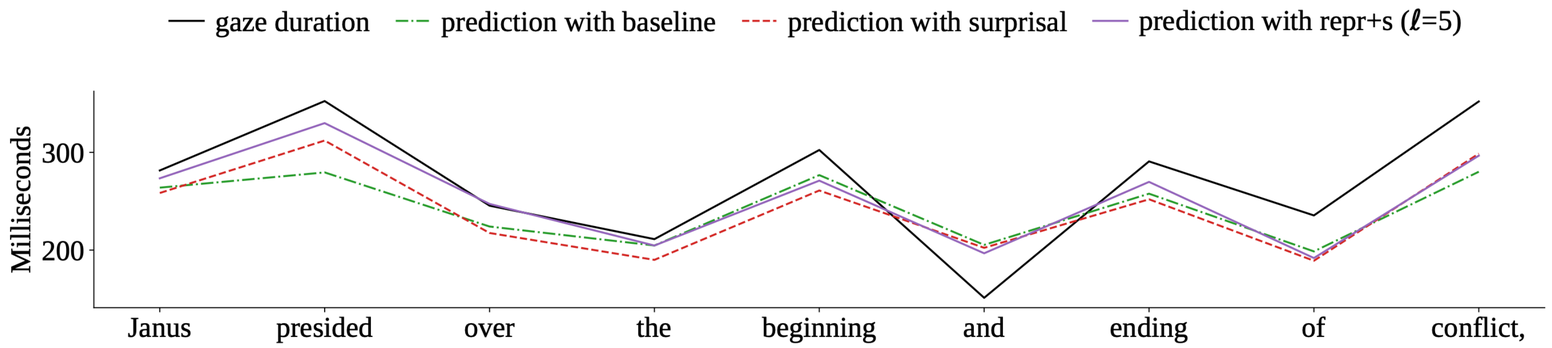
IAS can be implemented using **LMs as both generators** and proxy **representation functions**.



IAS predicts cloze response probability, N400, eye-tracking, and self-paced reading data **better than Surprisal**.

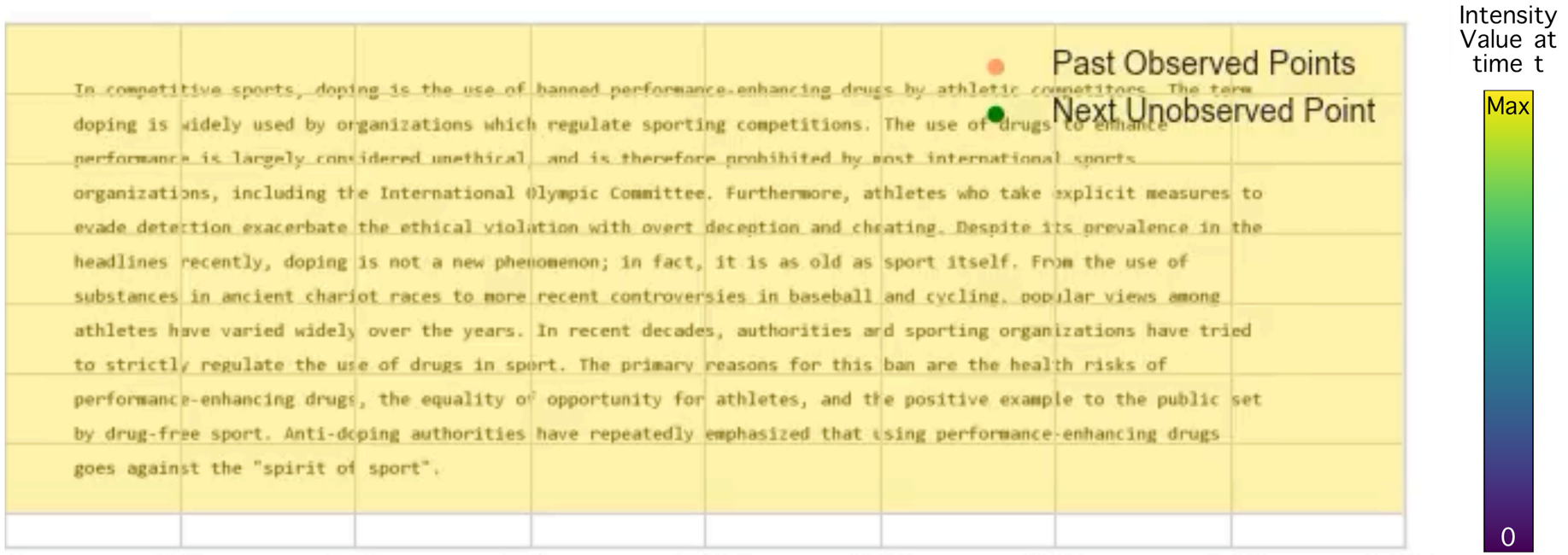
Predicting directly from LM embeddings





Tsipidi, Kiegeland, Re, Xu, Giulianelli, Stańczak, Cotterell. *ACL 2026*.

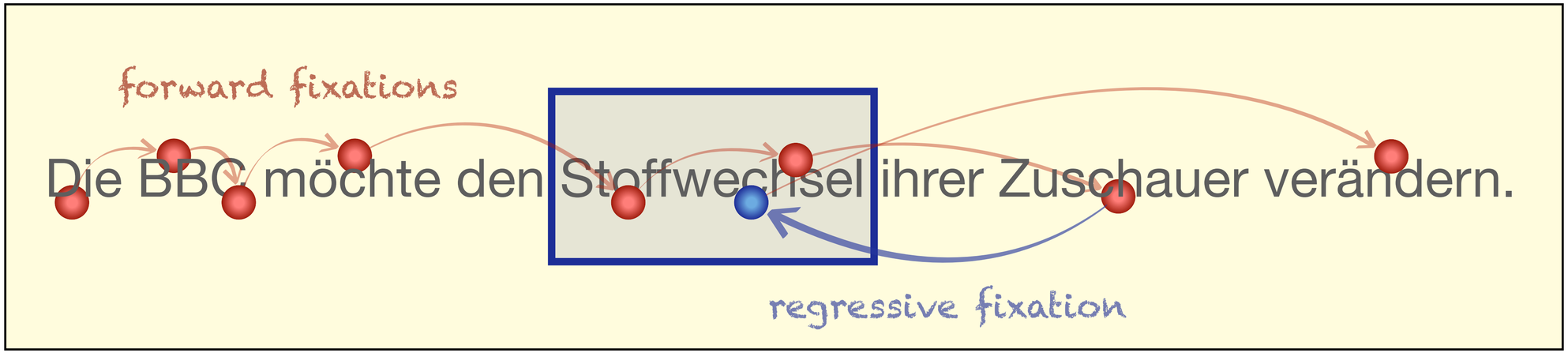
Fine-grained spatio-temporal modeling of reading behaviour



Predicted Gaze Intensity map of a **spatio-temporal Hawkes process** with type-writer effect, previous fixation surprisal, and reader-specific coefficients.

Re, Opedal, Manaiev, Giulianelli, Cotterell. *ACL 2025*.

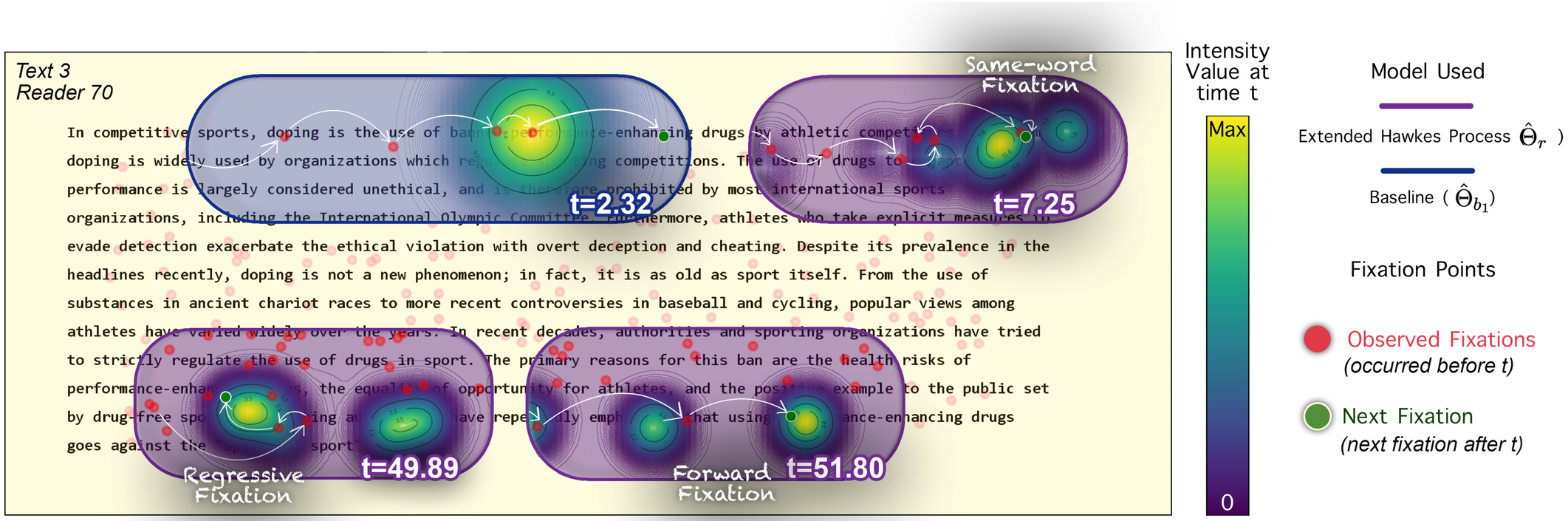
Lossy treatment of spatio-temporal data



Multiple **raw data points** *(e.g., individual fixations)* **aggregated** into a single measurement *(e.g., total fixation time)*

→ Obscures **distinct underlying cognitive processes**

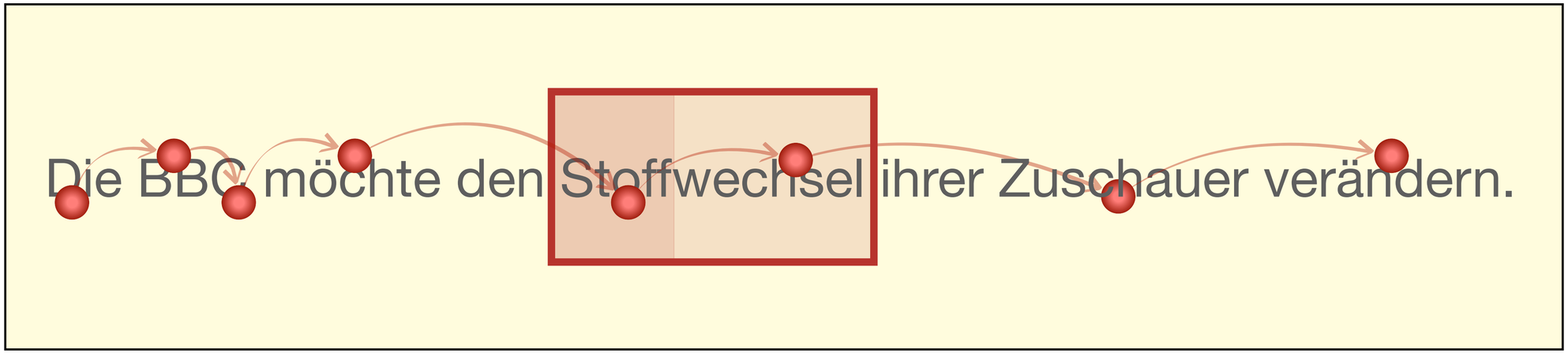
Fine-grained spatio-temporal modeling of reading behaviour



Predicted Gaze Intensity map of a **spatio-temporal Hawkes process** with type-writer effect, previous fixation surprisal, and reader-specific coefficients.

Re, Opedal, Manaiev, Giulianelli, Cotterell. *ACL 2025*.

Lossy treatment of spatio-temporal data



Words assumed as the fundamental unit of incremental processing

→ Language models defined over arbitrary vocabulary of “**tokens**”

→ Ignores **other plausible regions of interest** *(e.g., characters or morphemes)*

Token-level LMs for character-level problems

Anne␣lost␣control␣and␣laughed.

Stimulus

Skip Rate (control␣|Anne␣lost␣)

Measurement

$-\log p(\text{con}|\text{Anne␣lost␣})$

Predictor

Token-level LM



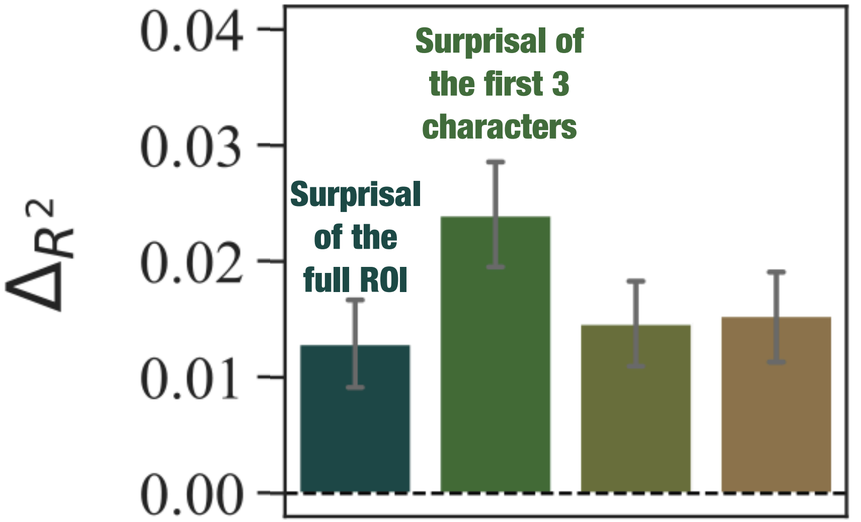
realphabetization



Character-level LM

Vieira, LeBrun, Giulianelli, Gastaldi, DuSell, Terilla, O'Donnel, Cotterell. ICML 2025.

Skip Rate



Full ROI	⟨lost␣, control␣, and␣, laughed.⟩
Fixed	⟨los, con, and, lau⟩
Dynamic (7)	⟨lost␣, contr, and, laugh⟩
Dynamic (8)	⟨lost␣, contro, and␣, laughe⟩

Giulianelli, Malagutti, Gastaldi, DuSell, Vieira, Cotterell. EMNLP 2024.

She took his hand and

She took his hand and said…



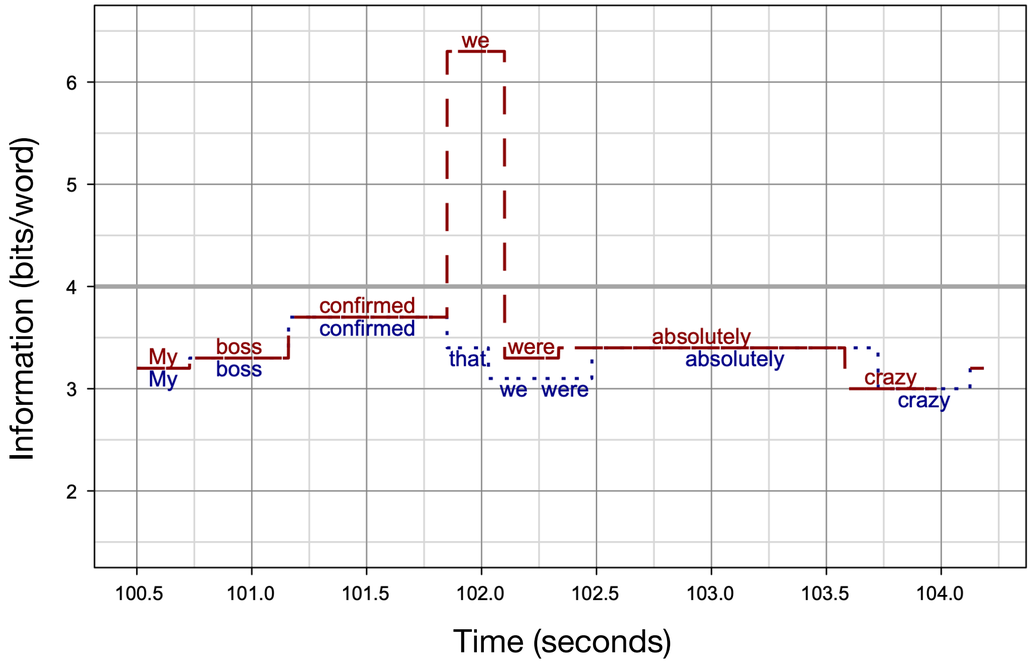
A window into the information processing that occurs as a human (or a machine) **comprehends** a sentence

Measuring prediction error and/or update gives us a notion of **processing cost** (e.g., to estimate reading difficulty)

Helps explain **production choices** as well (we tend to choose what's more predictable under our own model)

Predicting what comes next is such a powerful **training objective** for artificial neural networks

Information contours in texts and dialogues

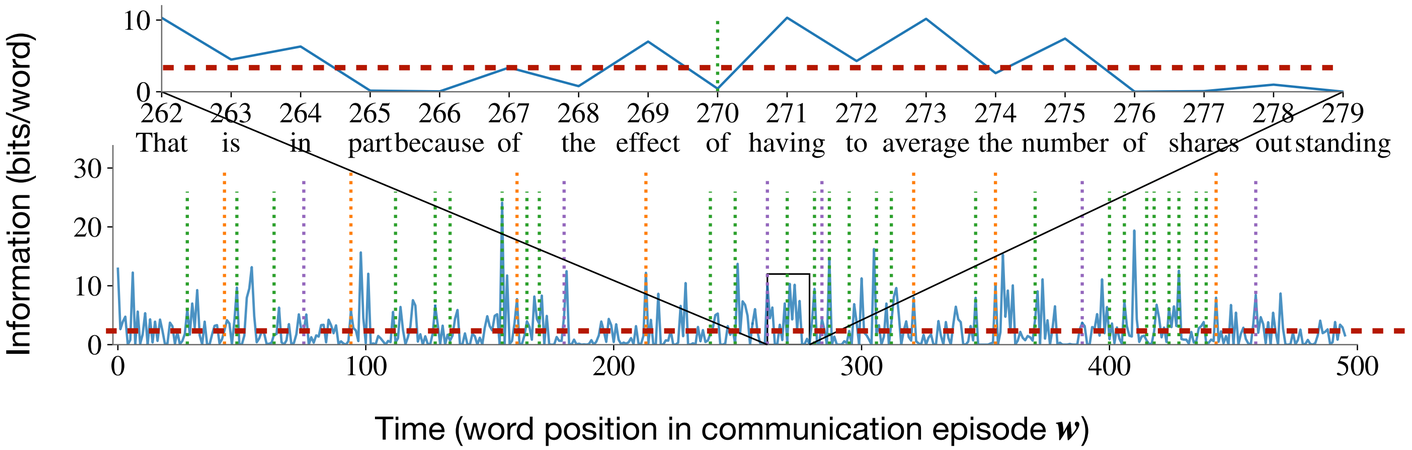


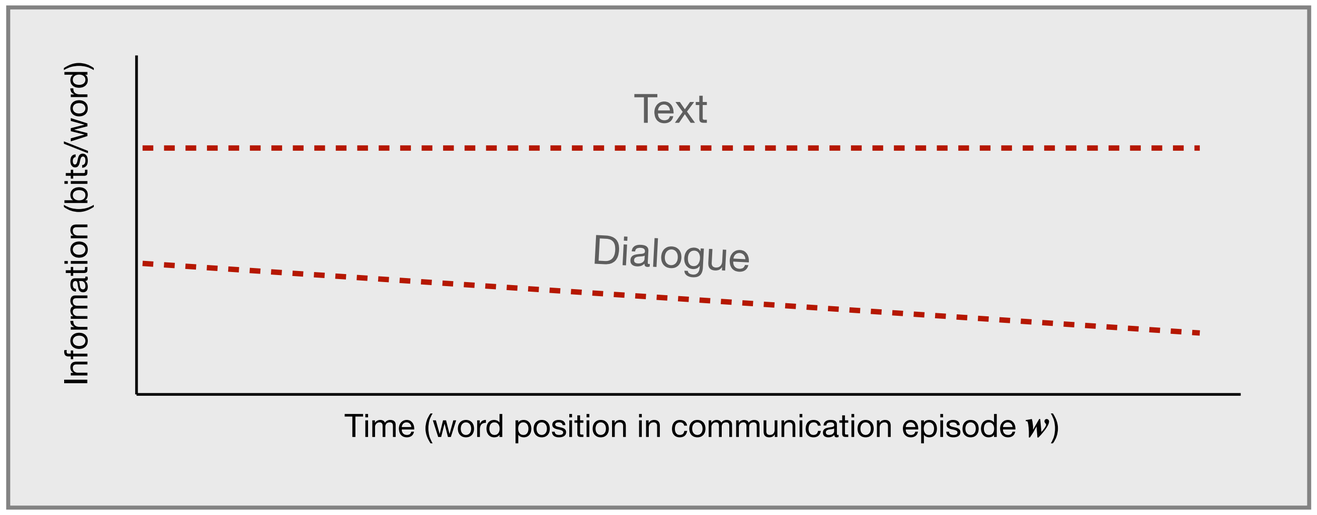
Jaeger. *Cognitive Psychology* (2010).

Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals ι_w are distributed as uniformly as possible throughout a communication episode w.

Information contours in texts and dialogues





Hypothesis 1: Uniform Information Density

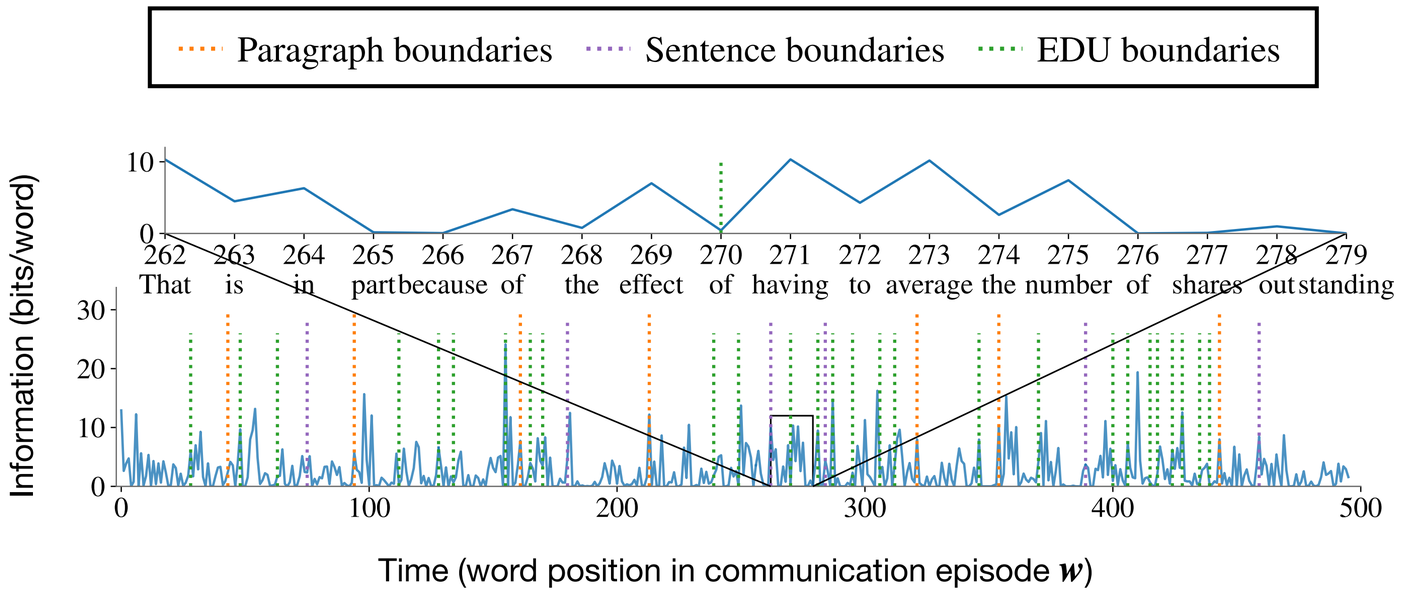
Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals ι_w are distributed as uniformly as possible throughout a communication episode w.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Information contours in texts and dialogues



Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals $\iota_{\boldsymbol{w}}$ are distributed as uniformly as possible throughout a communication episode $\boldsymbol{w}$.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Hypothesis 2: Structured Context

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ are (partially) determined by the position of w_t within the hierarchy of $\boldsymbol{w}$'s constituent structural units.

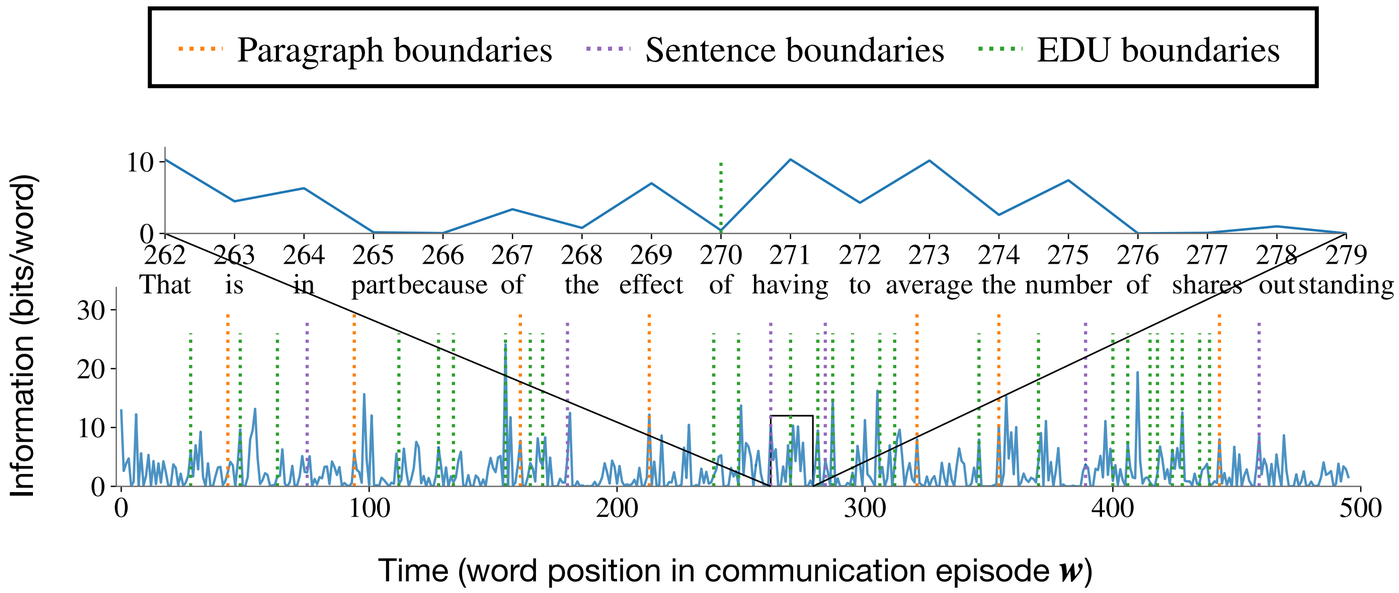
→ a unit's position within **contextual structure** predicts its surprisal

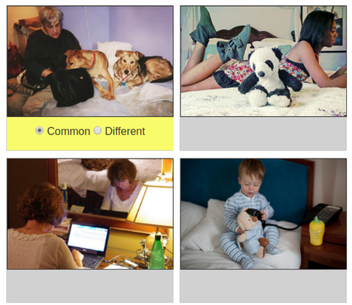
→ RST discourse units in texts

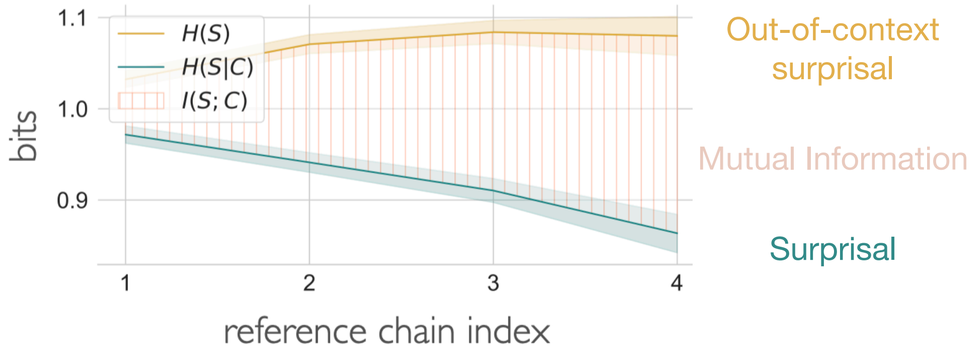
→ task-specific contextual units in dialogues

Giulianelli, Sinclair, Fernández. EMNLP 2021.
Tsipidi, Nowak, Cotterell, Wilcox, Giulianelli, Warstadt. EMNLP 2024.

Information contours in texts and dialogues







Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals $\iota_{\boldsymbol{w}}$ are distributed as uniformly as possible throughout a communication episode $\boldsymbol{w}$.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Hypothesis 2: Structured Context

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ are (partially) determined by the position of w_t within the hierarchy of $\boldsymbol{w}$'s constituent structural units.

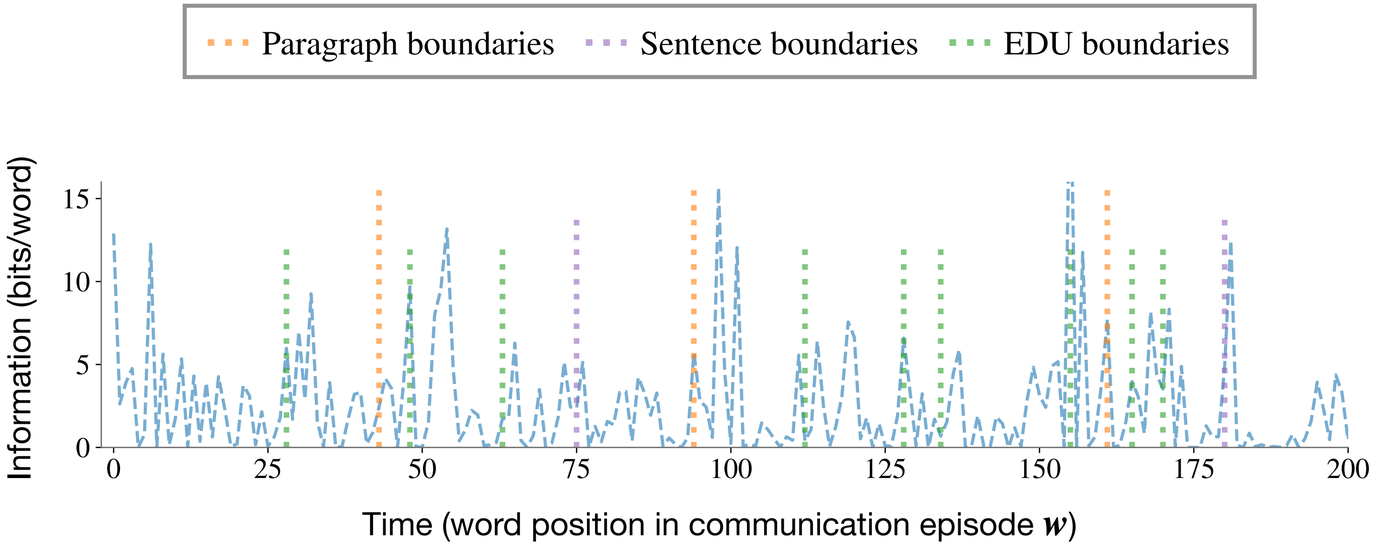
→ a unit's position within **contextual structure** predicts its surprisal

→ RST discourse units in texts

→ task-specific contextual units in dialogues

Giulianelli, Sinclair, Fernández. EMNLP 2021.
Tsipidi, Nowak, Cotterell, Wilcox, Giulianelli, Warstadt. EMNLP 2024.

Information contours in texts and dialogues



Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals ι_w are distributed as uniformly as possible throughout a communication episode w.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Hypothesis 2: Structured Context

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour ι_w of a communication episode w are (partially) determined by the position of w_t within the hierarchy of w's constituent structural units.

→ a unit's position within **contextual structure** predicts its surprisal

→ RST discourse units in texts

→ task-specific contextual units in dialogues

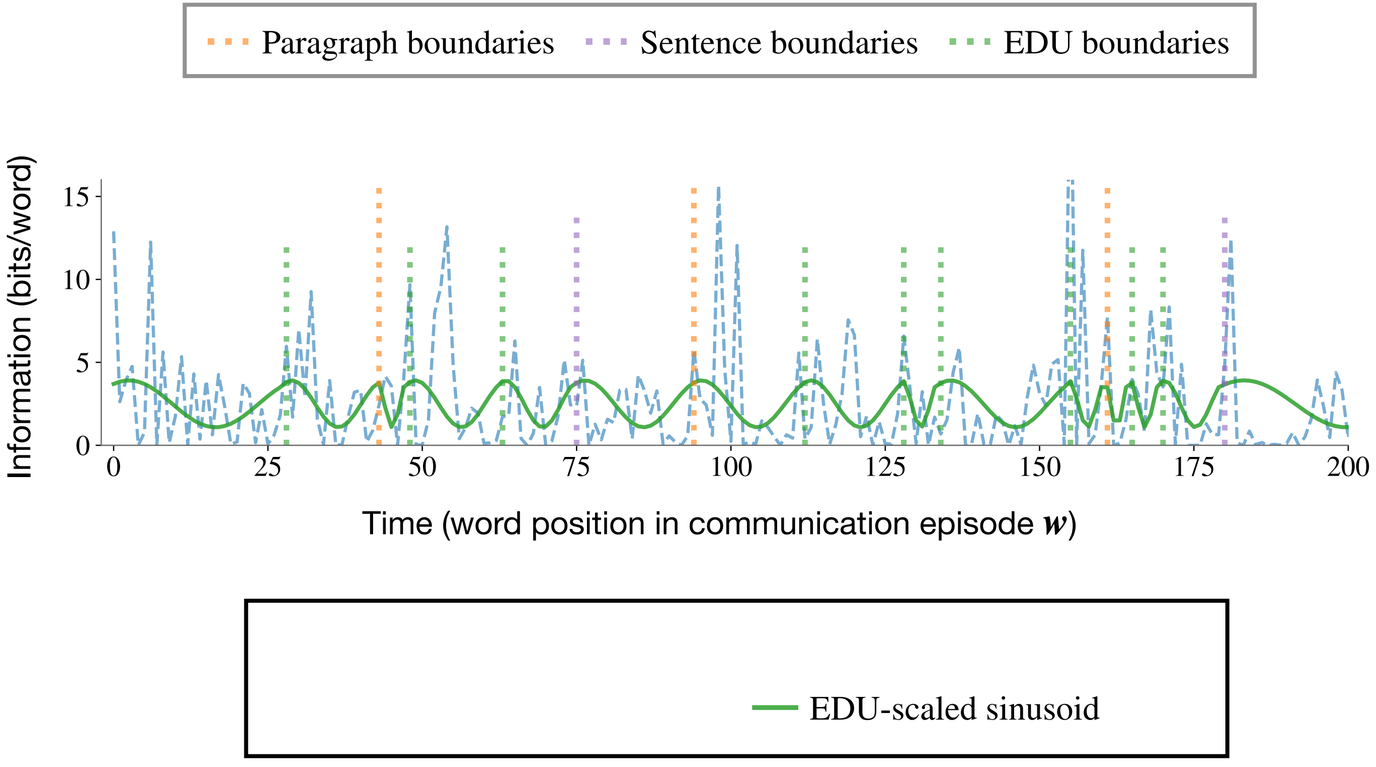
Giulianelli, Sinclair, Fernández. EMNLP 2021.
Tsipidi, Nowak, Cotterell, Wilcox, Giulianelli, Warstadt. EMNLP 2024.

Hypothesis 3: Harmonic Surprisal

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour ι_w of a communication episode w vary periodically, with periods that correspond to the boundaries of structural units within ι_W .

Tsipidi, Kiegeland, Nowak, Xu, Wilcox, Warstadt, Cotterell, Giulianelli. ACL 2025.

Information contours in texts and dialogues



Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals $\iota_{\boldsymbol{w}}$ are distributed as uniformly as possible throughout a communication episode $\boldsymbol{w}$.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Hypothesis 2: Structured Context

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ are (partially) determined by the position of w_t within the hierarchy of $\boldsymbol{w}$'s constituent structural units.

→ a unit's position within **contextual structure** predicts its surprisal

→ RST discourse units in texts

→ task-specific contextual units in dialogues

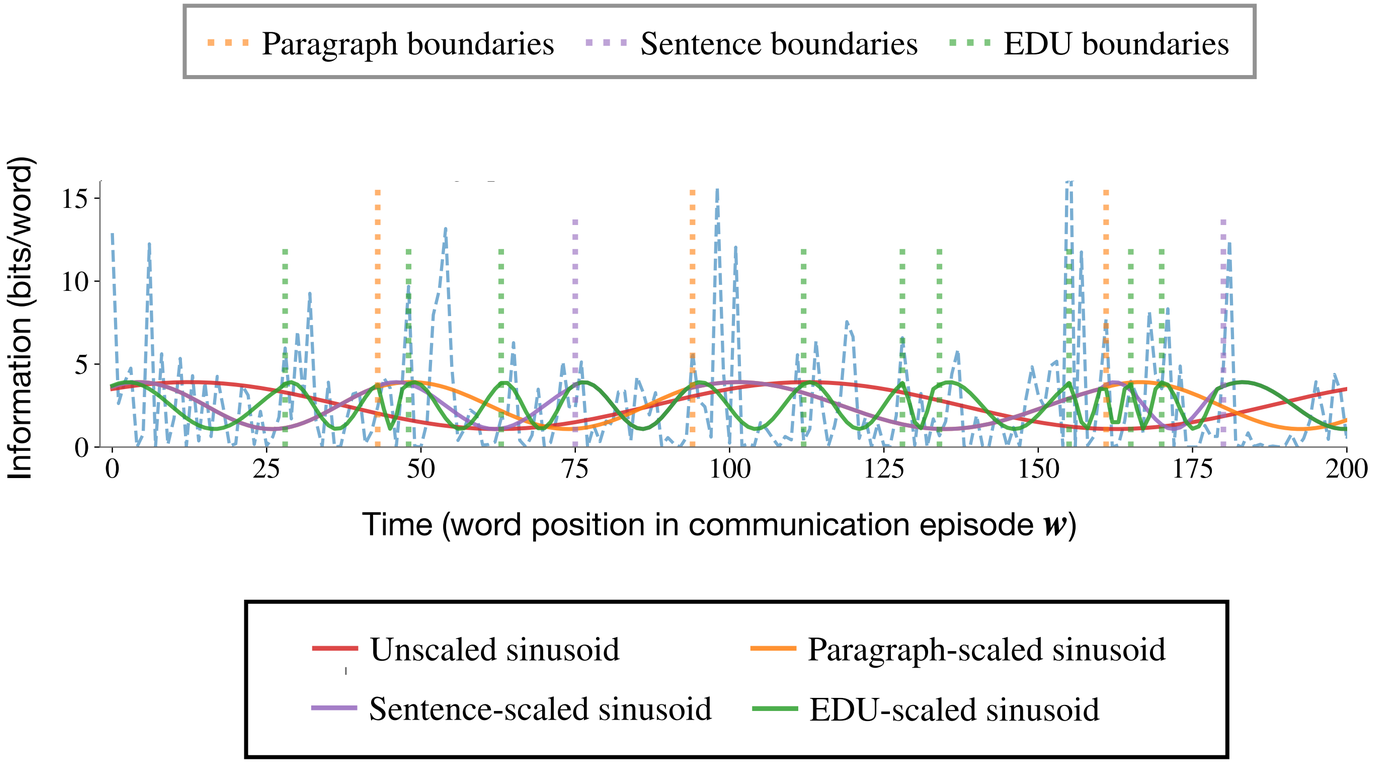
Giulianelli, Sinclair, Fernández. EMNLP 2021.
Tsipidi, Nowak, Cotterell, Wilcox, Giulianelli, Warstadt. EMNLP 2024.

Hypothesis 3: Harmonic Surprisal

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ vary periodically, with periods that correspond to the boundaries of structural units within $\iota_{\boldsymbol{W}}$.

Tsipidi, Kiegeland, Nowak, Xu, Wilcox, Warstadt, Cotterell, Giulianelli. ACL 2025.

Information contours in texts and dialogues



Hypothesis 1: Uniform Information Density

Subject to the constraints of the grammar, speakers optimise their linguistic signals such that the surprisals $\iota_{\boldsymbol{w}}$ are distributed as uniformly as possible throughout a communication episode $\boldsymbol{w}$.

→ no evidence of local uniformity, pressure toward a global mean

→ information rate decreases in dialogues

Giulianelli & Fernández. CoNLL 2021.
Giulianelli, Sinclair, Fernández. AACL 2021.

Hypothesis 2: Structured Context

Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ are (partially) determined by the position of w_t within the hierarchy of $\boldsymbol{w}$'s constituent structural units.

→ a unit's position within **contextual structure** predicts its surprisal

→ RST discourse units in texts

→ task-specific contextual units in dialogues

Giulianelli, Sinclair, Fernández. EMNLP 2021.
Tsipidi, Nowak, Cotterell, Wilcox, Giulianelli, Warstadt. EMNLP 2024.

Hypothesis 3: Harmonic Surprisal

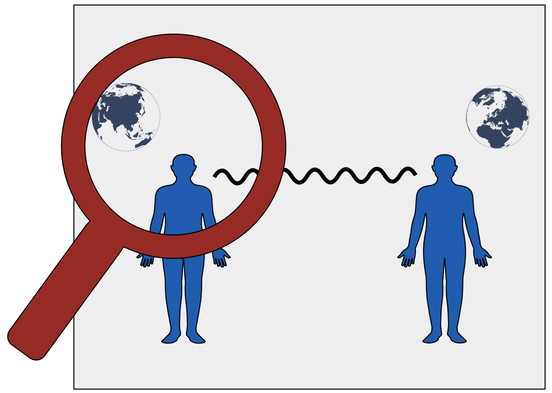
Values $\iota\left(w_t; \boldsymbol{w}_{<t}\right)$ in the surprisal contour $\iota_{\boldsymbol{w}}$ of a communication episode $\boldsymbol{w}$ vary periodically, with periods that correspond to the boundaries of structural units within $\iota_{\boldsymbol{W}}$.

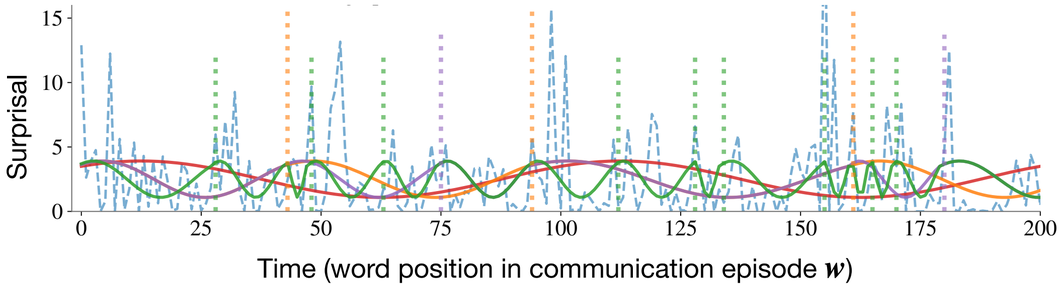
Tsipidi, Kiegeland, Nowak, Xu, Wilcox, Warstadt, Cotterell, Giulianelli. ACL 2025.

Information contours in texts and dialogues

Producers' communicative strategies through the lens of information rate modulation

- facilitating production (e.g., repetitions)
  Giulianelli, Sinclair, Fernández. AACL 2022.
- enhancing coordination in dialogue
  Yee, Giulianelli, Sinclair. LREC-COLING 2024.
- style, genre, and writing quality
- human-generated vs. model-generated texts
- facilitating comprehension
- in multimodal contexts
  Gay, Haley, Giulianelli, Ponto. EACL 2026.

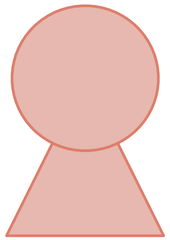




Experimental Setup

Human dialogue context:

My boss confirmed..



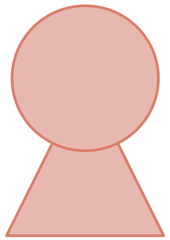
that we were absolutely crazy

Switchboard Dialogue Act Corpus
(Stolcke et al., 2000) Dialogue act modeling for automatic tagging and recognition of conversational speech, Computational Linguistics.

Experimental Setup

Human dialogue context:

My boss confirmed..



that we were absolutely crazy

Paraphrases

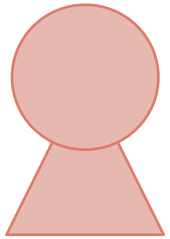
LM

Continuations

Experimental Setup

Human dialogue context:

My boss confirmed..



that we were absolutely crazy

Paraphrases

- we were absolutely crazy.
- that we were absolutely crazy.
- that we were completely irrational.
- that she thought we were insane.

LM

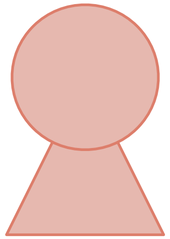
Continuations

Goal-directed
Goal-agnostic

Experimental Setup

Human dialogue context:

My boss confirmed..



that we were absolutely crazy

LM

Paraphrases

- we were absolutely crazy.
- that we were absolutely crazy.
- that we were completely irrational.
- that she thought we were insane.

Continuations

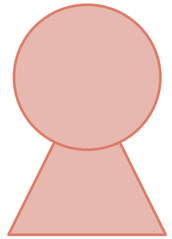
- the meeting has been moved.
- we have the ok on this project.
- my request for time off.
- that we have her full support.
- that we were crazy.

Goal-directed
Goal-agnostic

Experimental Setup

Human dialogue context:

My boss confirmed..



that we were absolutely crazy

Human

LM

Paraphrases

we were absolutely crazy.
that we were absolutely crazy.
that we were completely irrational.
that she thought we were insane.

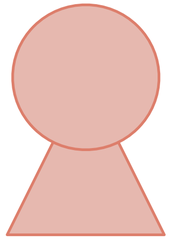
Goal-directed

Continuations

the meeting has been moved.
we have the ok on this project.
my request for time off.
that we have her full support.
that we were crazy.

Goal-agnostic

Experimental Setup



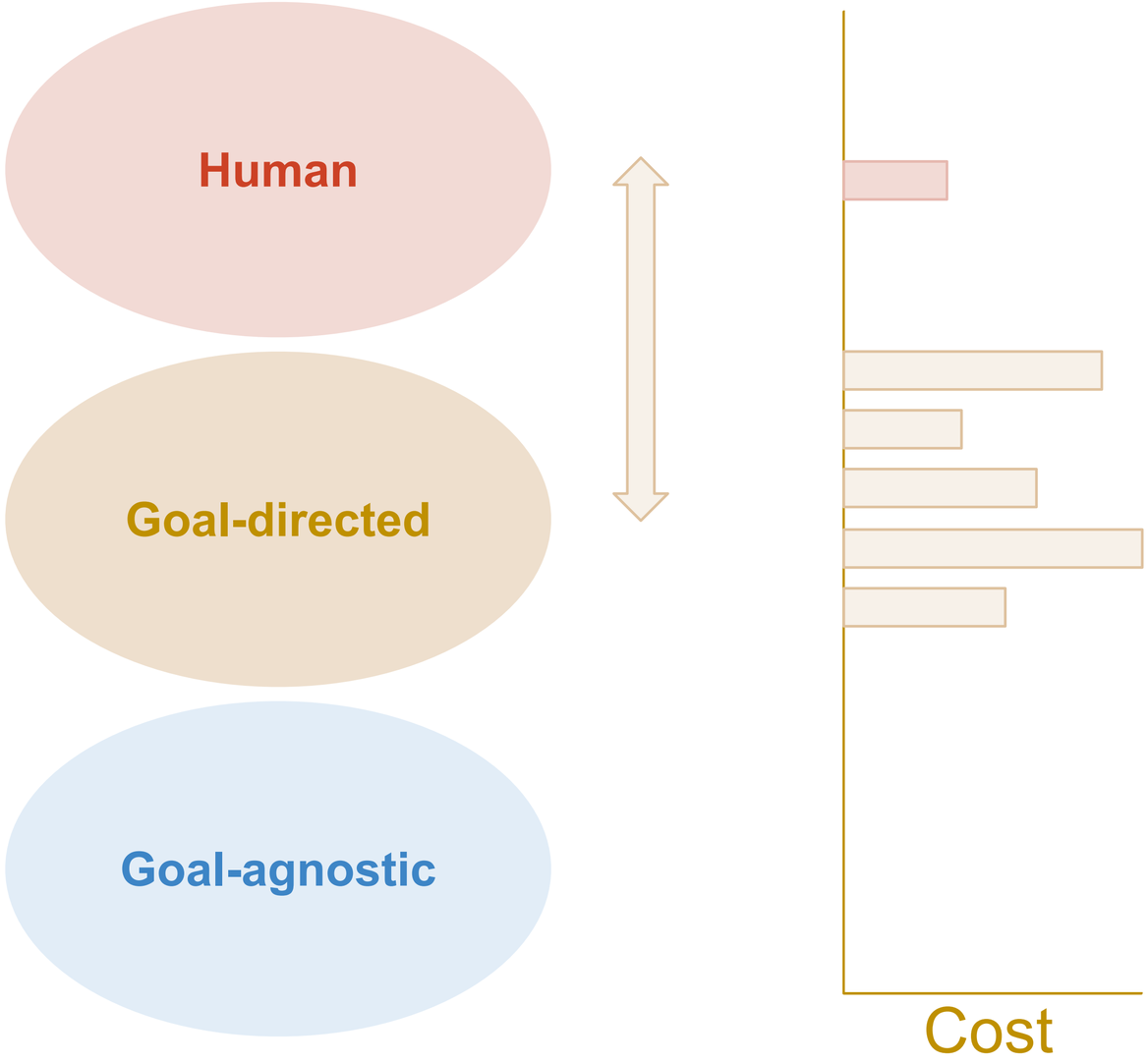
My boss confirmed..

that we were absolutely crazy

LM

we were absolutely crazy.
that we were absolutely crazy.
that we were completely irrational.
that she thought we were insane.
that we were crazy.

the meeting has been moved.
we have the ok on this project.
my request for time off.
that we have her full support.
that we were crazy



Experimental Setup

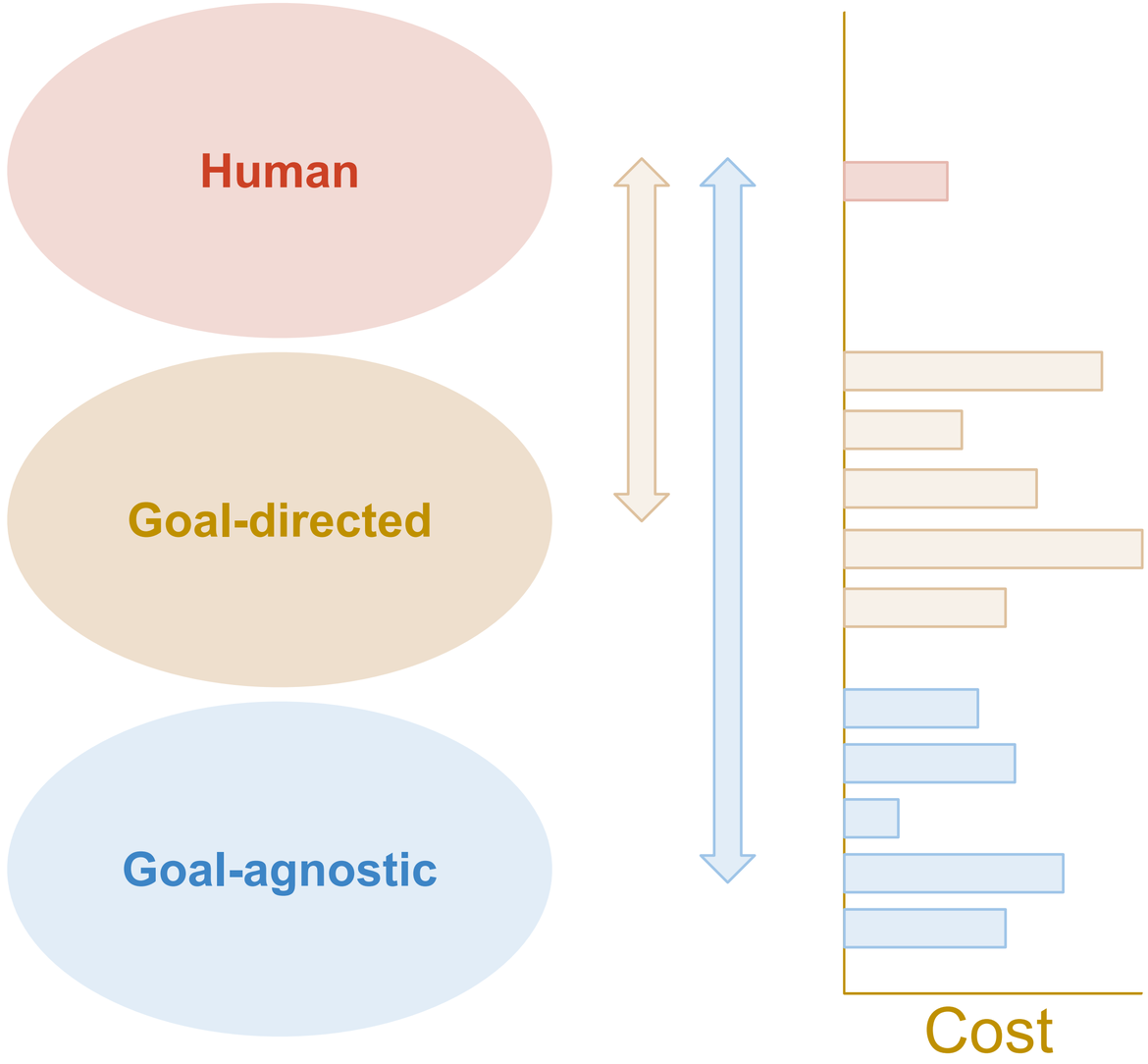
My boss confirmed..

that we were absolutely crazy

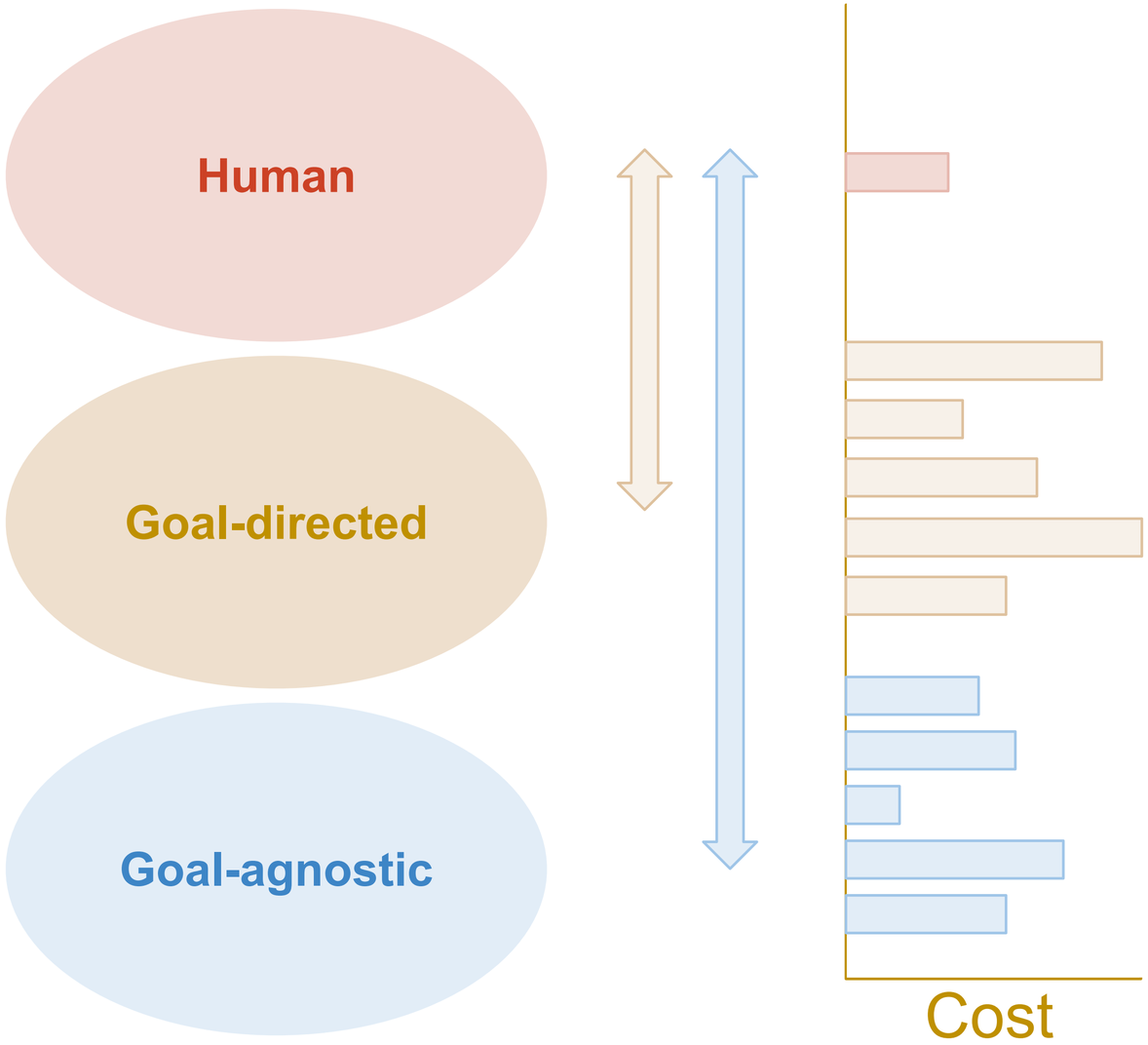
LM

we were absolutely crazy.
that we were absolutely crazy.
that we were completely irrational.
that she thought we were insane.
that we were crazy.

the meeting has been moved.
we have the ok on this project.
my request for time off.
that we have her full support.
that we were crazy

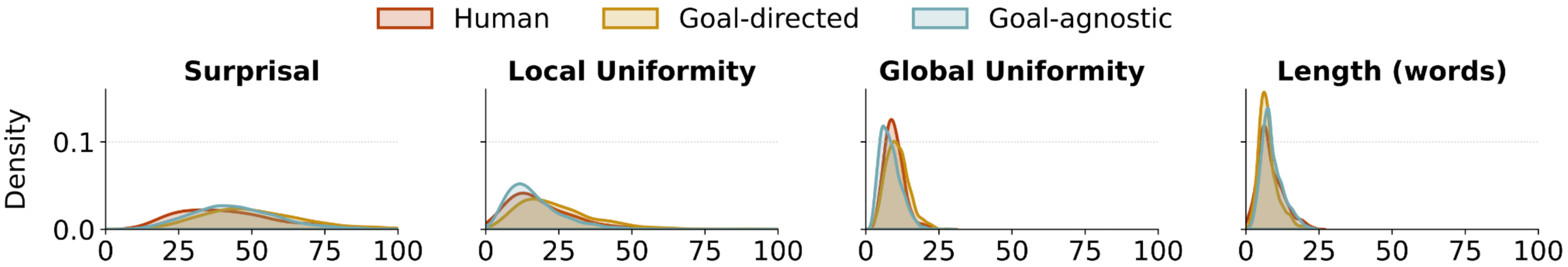


Measures of Production Cost



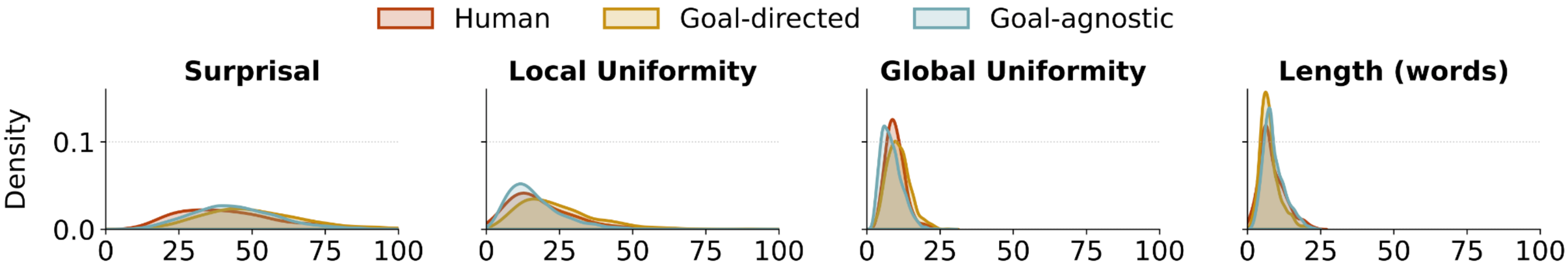
- **Surprisal –** How unlikely a continuation is, given a context.
- **Uniformity of Information Density (UID) –** how unevenly token surprisal is distributed across an utterance.
  - **Local UID –** The average difference between the surprisal of each adjacent token in the continuation.
  - **Global UID –** The variance of surprisal over the full utterance.
- **Length** - The number of words in the continuation.

Results - Global Distributions



Global distributions of **human**, **goal-directed** and **goal-agnostic** continuations for different cost measures

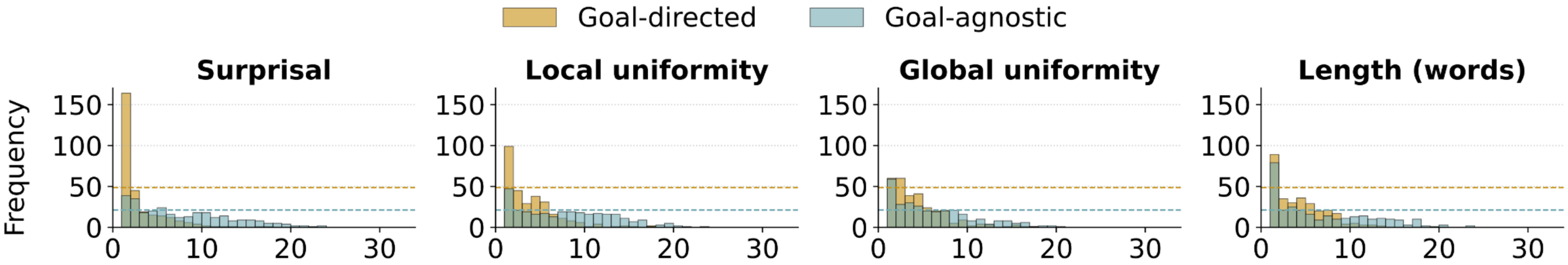
Results - Global Distributions



Global distributions of **human**, **goal-directed** and **goal-agnostic** continuations for different cost measures

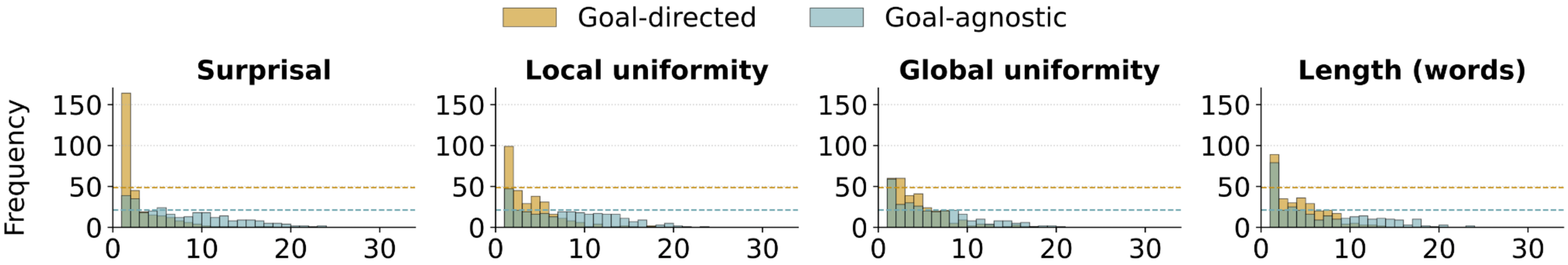
→ Stratified sampling ensures that generated alternatives do not systematically differ from human continuations *in their overall cost distributions.*

Results - Deterministic Cost Minimisation



Rank distributions of human continuations under different cost measures, evaluated against **goal-directed** and **goal-agnostic** alternative sets.

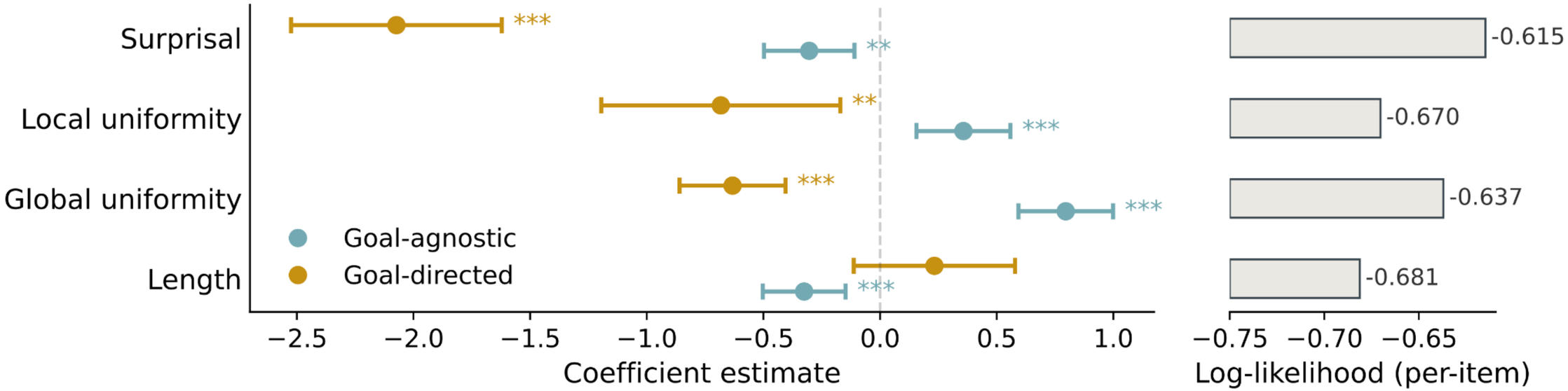
Results - Deterministic Cost Minimisation



Rank distributions of human continuations under different cost measures, evaluated against **goal-directed** and **goal-agnostic** alternative sets.

→ Human continuations deterministically minimise cost (rank 1) significantly more often than chance.

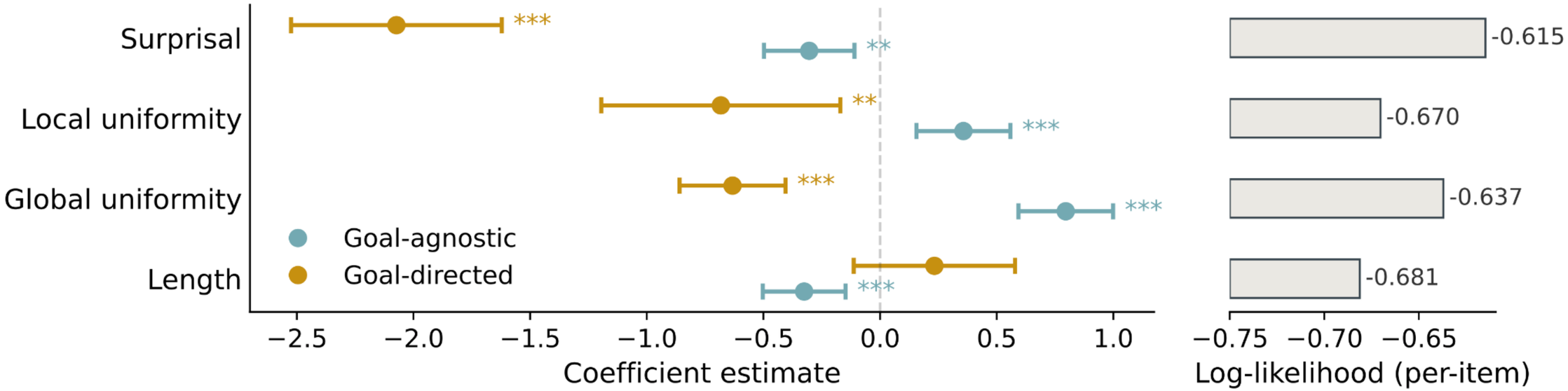
Results - Graded Cost Sensitivity



Pairwise logistic choice model predicting whether a continuation is selected over an alternative as a function of their difference in cost.

Coefficient estimate on log scale, bars indicate 95% CI, $p < 0.05$* $p < 0.01$** $p < 0.001$***

Results - Graded Cost Sensitivity



Pairwise logistic choice model predicting whether a continuation is selected over an alternative as a function of their difference in cost.

→ Probabilistic minimisation of surprisal over goal-directed alternatives provides the best explanation of production choices.

Coefficient estimate on log scale, bars indicate 95% CI, $p < 0.05$* $p < 0.01$** $p < 0.001$***