Error-driven Learning in DFT:
A case study of structural priming

Zhenghao Herbert Zhou
Yale University
LSA Annual Meeting
Jan 10, 2025

Part of the Symposium Talk Series:
Dynamic Field Theory for unifying discrete and continuous aspects of linguistic representations

Structural Priming

Structural Priming: speakers tend to reuse the syntactic structures they have recently encountered during production or comprehension.

Our focus: Double Object (DO) vs. Prepositional Dative (PD) for ditransitive predicates.

- DO: *Alice sent Bob a letter.*
- PD: *Alice sent a letter to Bob.*

E.g. [Bock 1986, Chang 2012]

Lexical Boost Effect (LBE)

Lexical Boost Effect: structural priming effect is stronger when the word that heads the primed structures is repeated between prime and target sentences.

Carl <u>gave</u> Danis a letter.

Carl <u>showed</u> Danis a letter.   **PRIME**
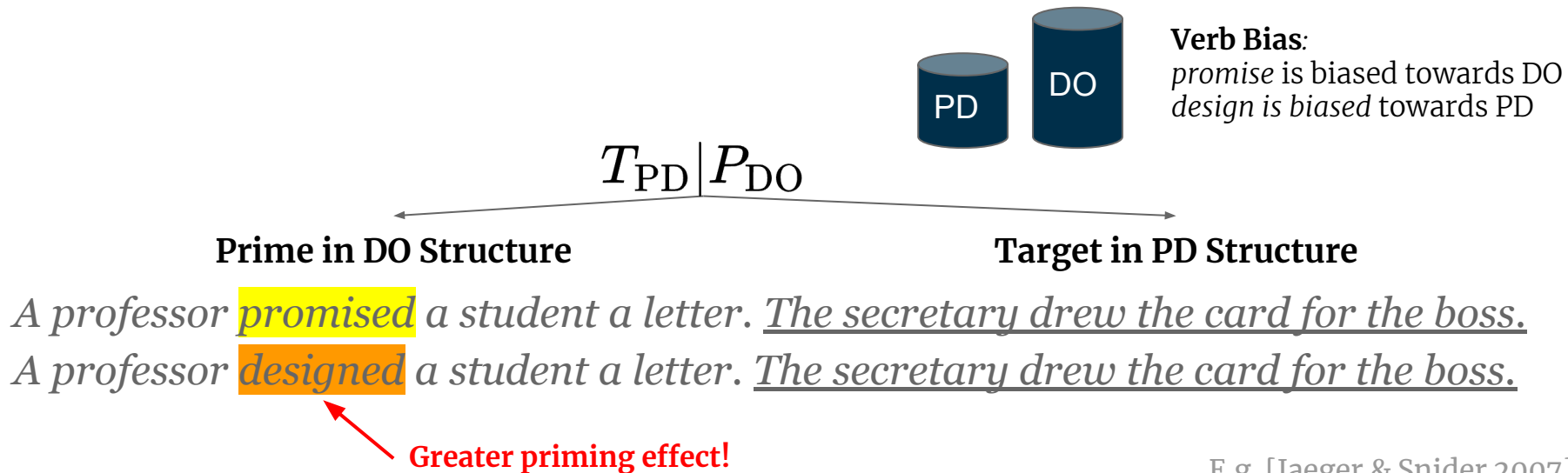
Alice <u>gave</u> Bob a book.   **TARGET**

E.g. [Pickering & Branigan 1998]

Inverse Frequency Effect (IFE)

Inverse Frequency Effect: the less preferred (lower frequency) syntactic structure causes a stronger priming effect than the more preferred (higher frequency) structural alternative.

Verb Bias:
promise is biased towards DO
design is biased towards PD

$$T_{\mathrm{PD}}|P_{\mathrm{DO}}$$

Prime in DO Structure   **Target in PD Structure**

A professor promised *a student a letter.* The secretary drew the card for the boss.
A professor designed *a student a letter.* The secretary drew the card for the boss.

Greater priming effect!

E.g. [Jaeger & Snider 2007]

Inverse Frequency Effect (IFE), cont.

$$T_{\mathrm{PD}}|P_{\mathrm{DO}}$$

Prime in DO Structure   **Target in PD Structure**

A professor <mark>promised</mark> *a student a letter.* <u>*The secretary drew the card for the boss.*</u>
A professor <mark>designed</mark> *a student a letter.* <u>*The secretary drew the card for the boss.*</u>

Prime Verb	**Verb PD Bias**	**Priming Effect (log scale)**
Promised	*0.22*	*0.5*
Buy	*0.36*	*1.1*
Find	*0.49*	*1.8*
Keep	*0.69*	*2.2*
Design	*0.72*	*2.9*

When priming in DO structure   *larger PD biases* ⟶ *a greater priming effect*

Structural Priming as Linguistic Adaptation

Linguistic Adaptation: the linguistic knowledge representations that are used for language processing **change** in response to language input.

- LBE is explained by transient activation theory (short-term);
- IFE is explained by implicit learning theory, an *error-driven mechanism.*
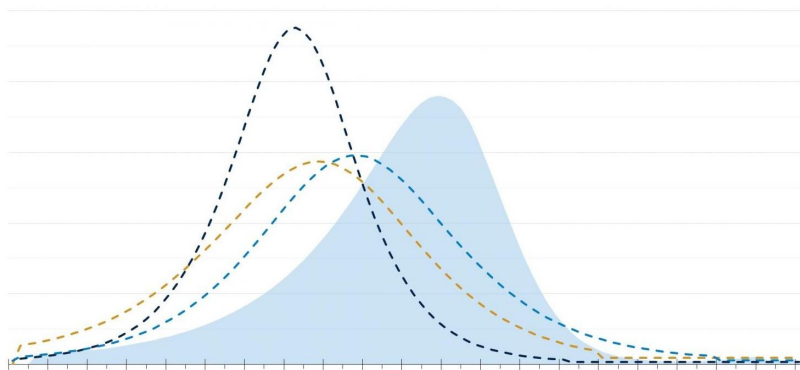
[E.g. Chang 2012]
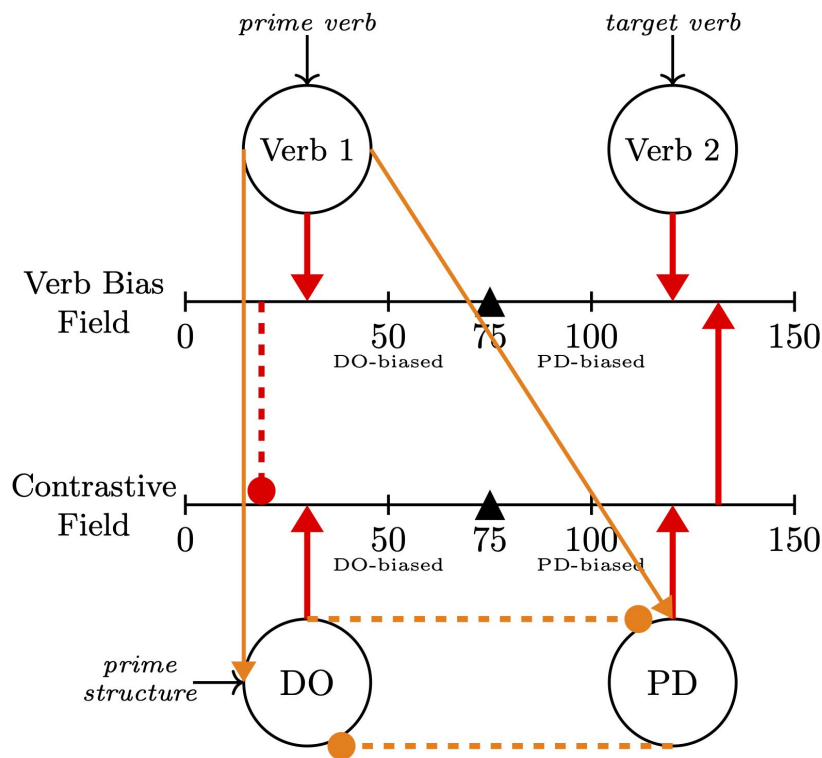
Model Architecture

Prime Verb	Verb PD Bias
Bring	*0.23*
Buy	*0.27*
Find	*0.41*
Draw	*0.52*
Design	*0.77*

- **Verb Bias Field** [*hosting both prediction and production*]: probabilistic information of producing one structure over the other;
- **Contrastive Field** [*hosting error signal computation*]: a space for computing the difference between expected and actual information;

🔴: to-field coupling, Gaussian stimuli;

🟠: to-node coupling, activation stimuli;

—: excitatory coupling;

- - -: inhibitive coupling

Processing Steps in DFT

Incrementally perceiving or processing the prime sentence:

- (i) perceiving the prime verb only;
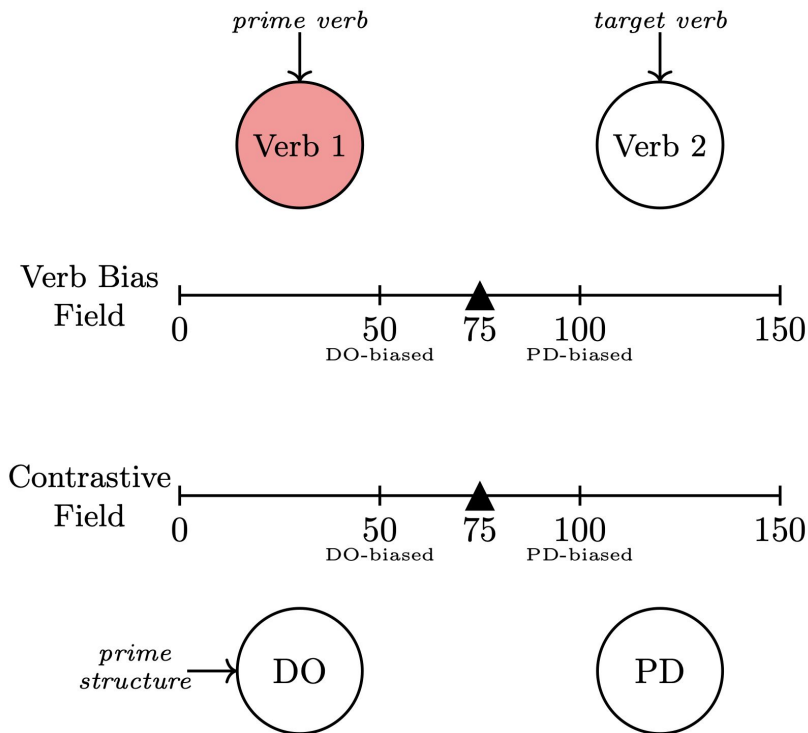
Processing Steps in DFT, cont.

**Incrementally perceiving or processing the prime sentence:**

- (i) perceiving the prime verb only;
- (ii) generating an expectation on the continuation of the sentence;
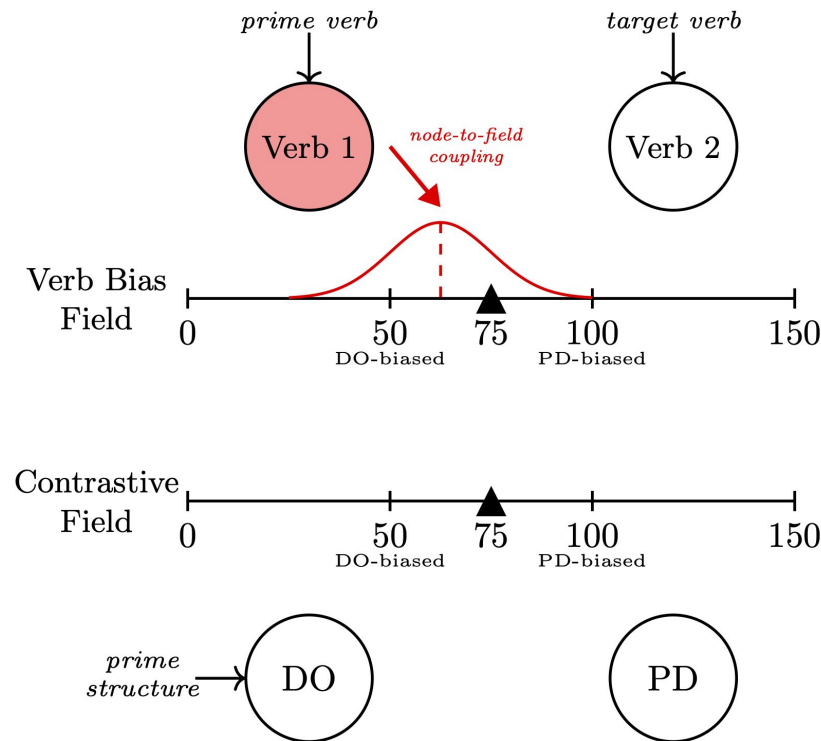
Processing Steps in DFT, cont.

**Incrementally perceiving or processing the prime sentence:**

- (i) perceiving the prime verb only;
- (ii) generating an expectation on the continuation of the sentence;
- (iii) perceiving the structure of the sentence;
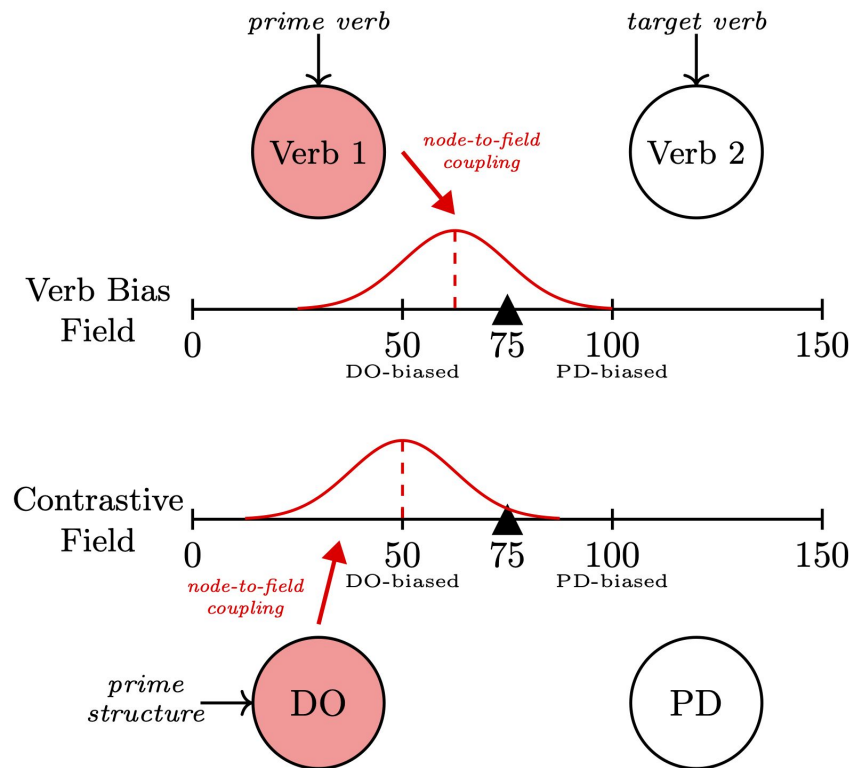
Processing Steps in DFT, cont.

Error-signal computation:

- (iv) compare the expectation of the probabilities of DO and PD versus the actual perceived prime structure.
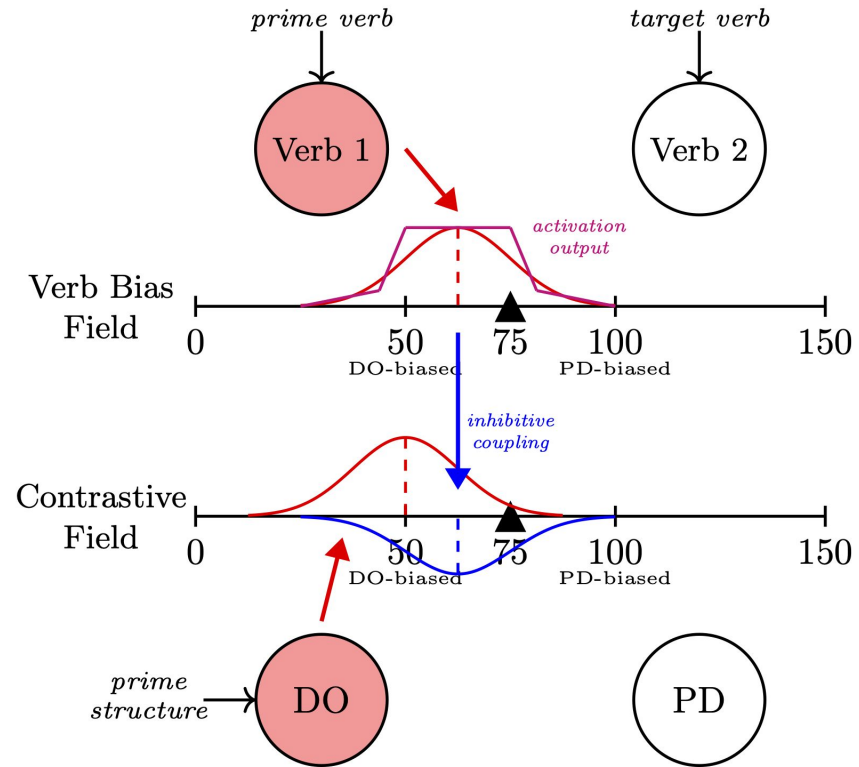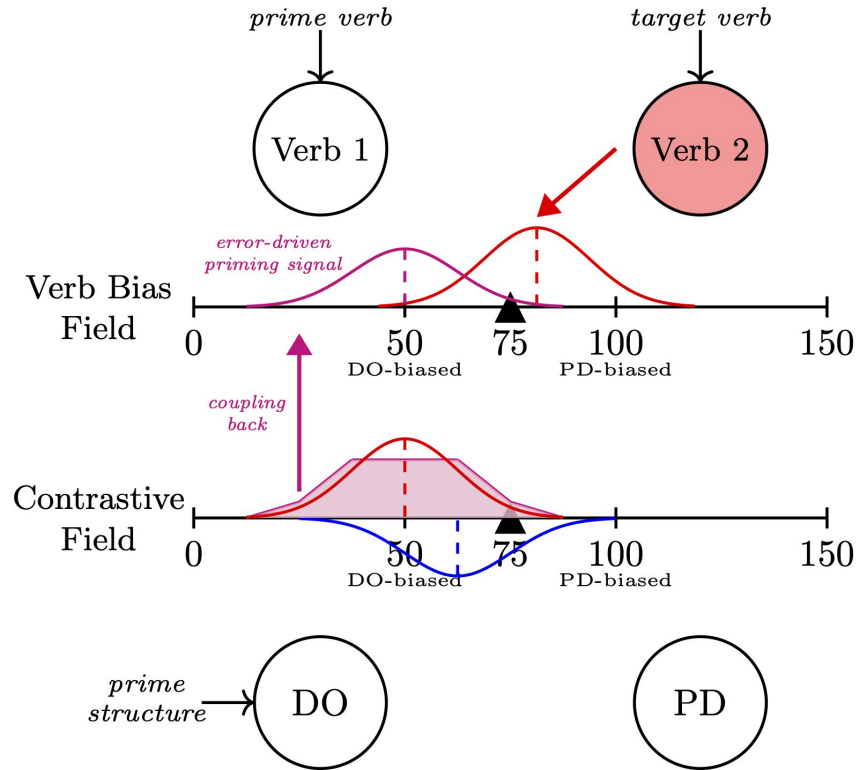
Processing Steps in DFT, cont.

Error-signal computation:

- (iv) compare the expectation of the probabilities of DO and PD versus the actual perceived prime structure.

  ***The magnitude of this difference is proportional to the final priming strength.
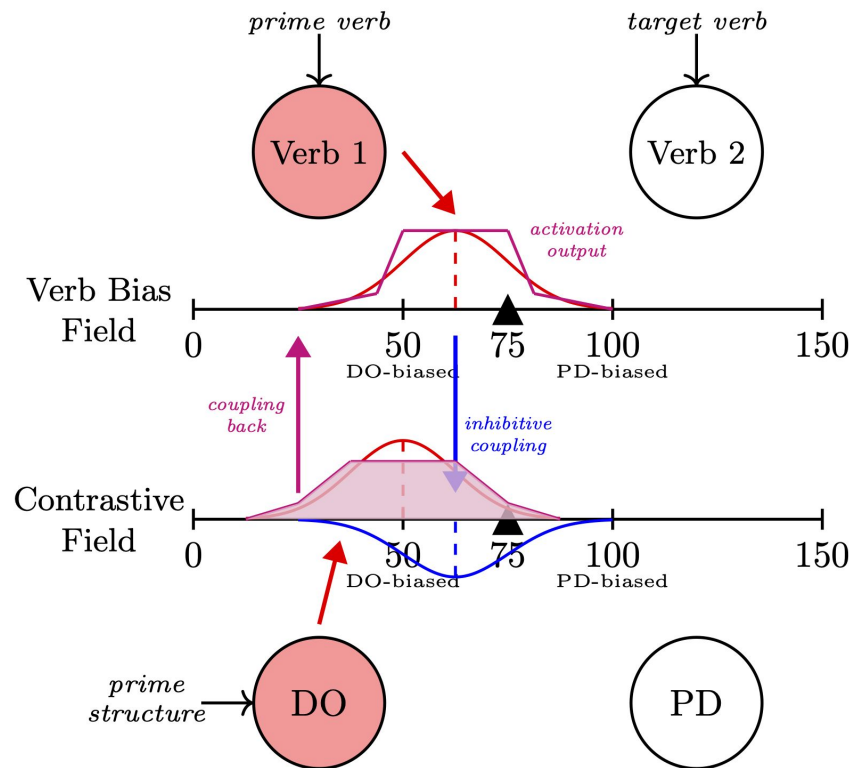
Processing Steps in DFT, cont.

**Producing the target sentence, as is affected by the prime:**

- (v) generating an expected relative frequency between DO and PD according to the target verb's verb bias;
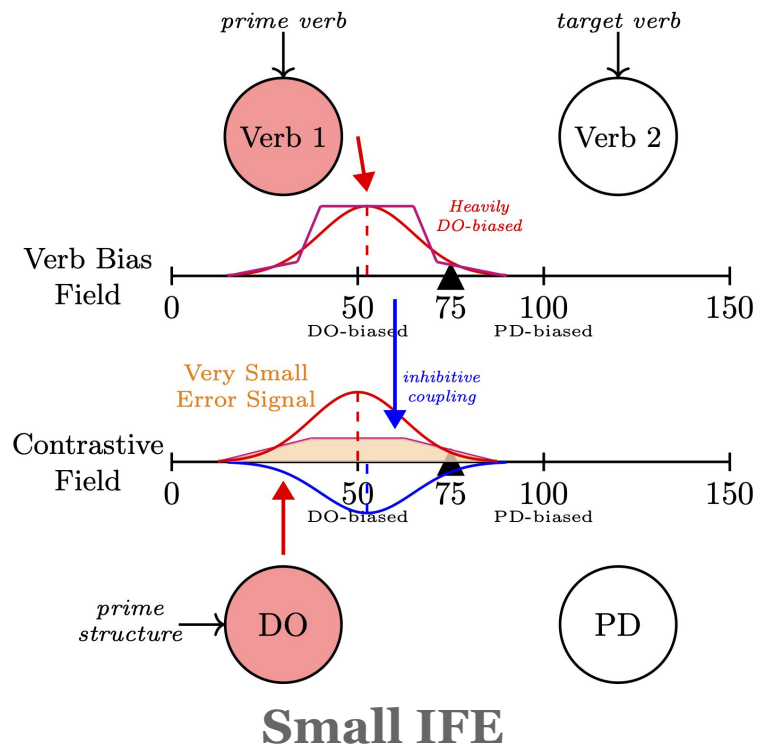
Processing Steps in DFT, cont.

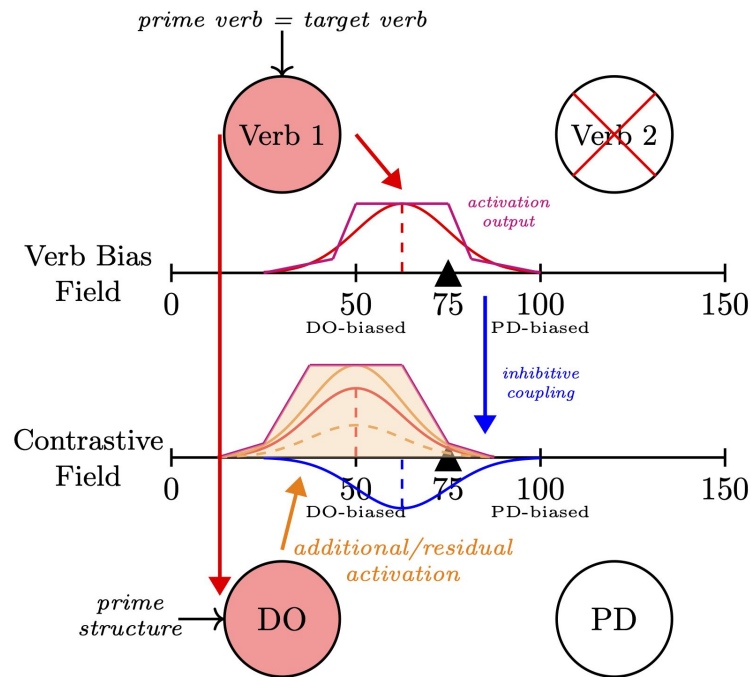**Producing the target sentence, as is affected by the prime:**

- (v) generating an expected relative frequency between DO and PD according to the target verb's verb bias;
- (vi) shifting the relative frequency towards the direction of the prime structure;
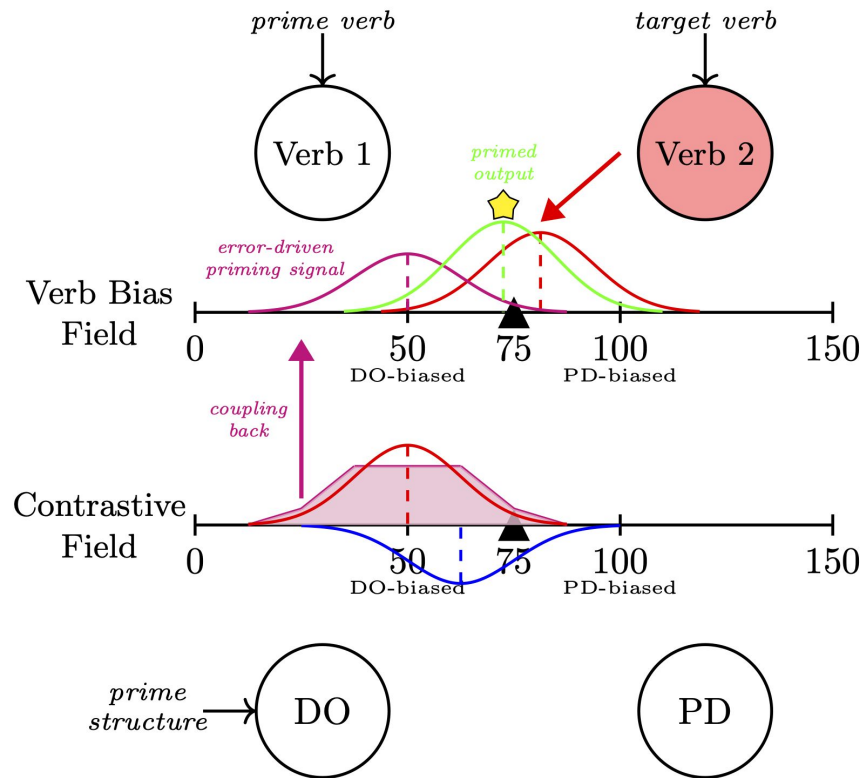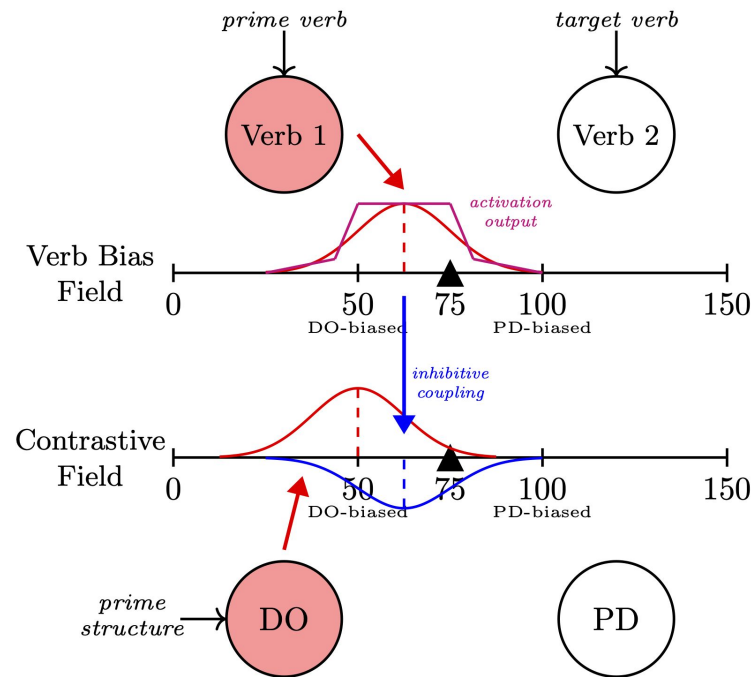
Capturing the Inverse Frequency Effect
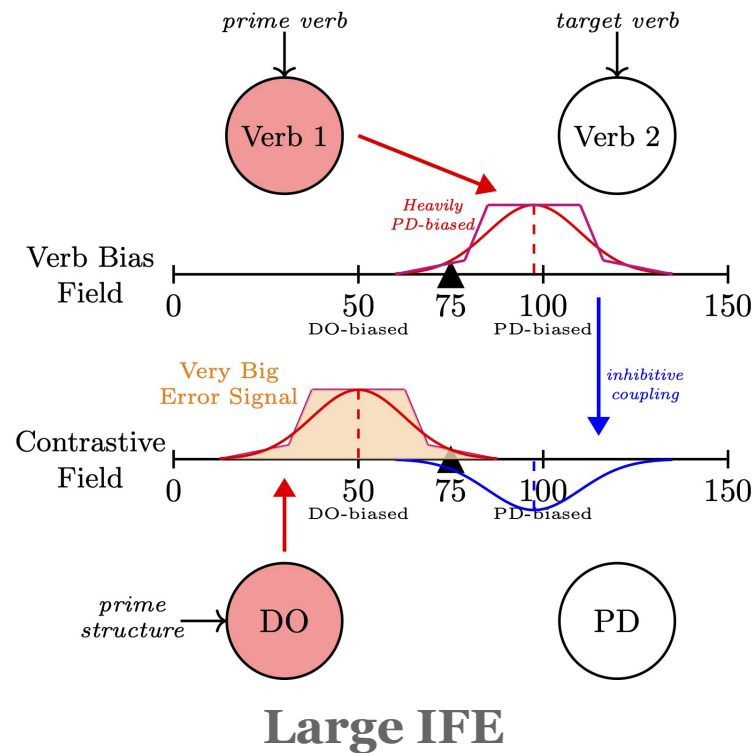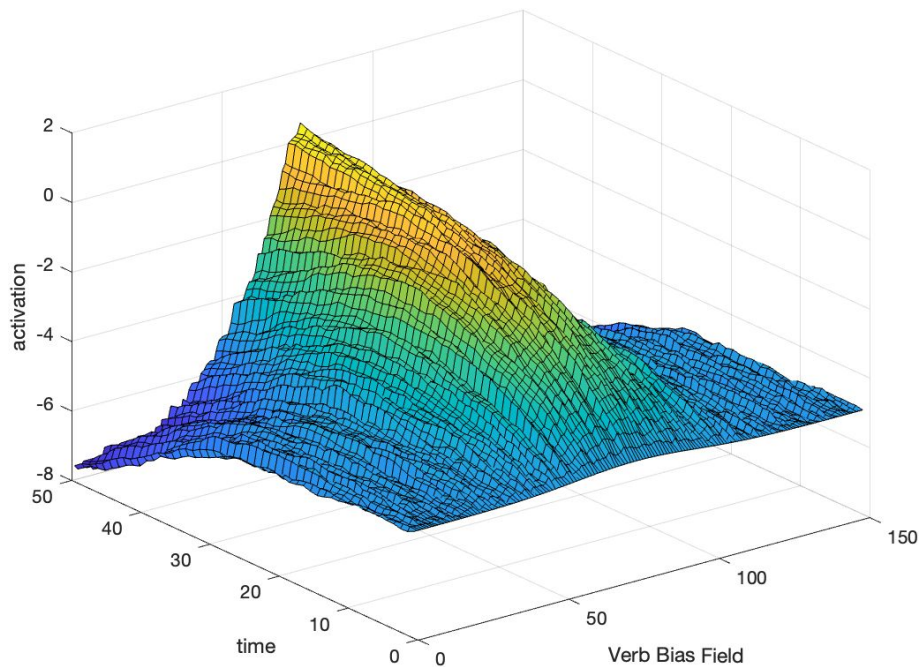
Capturing the Lexical Boost Effect
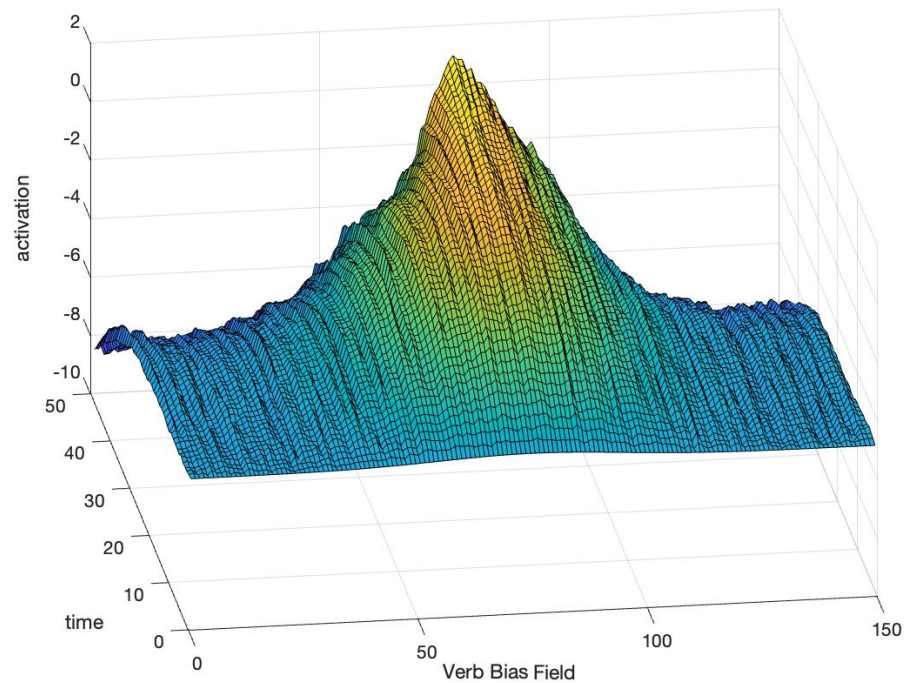

Lexical Boost

No Lexical Boost

Simulation Result: target$_{vb}$ = 75

Small IFE

Large IFE

Simulation Results

Prime Condition	Prime Verb Bias	Target Production Position	Priming Effect
Extremely DO-biased	55 (-20)	81	(+6)
Slightly DO-biased	65 (-10)	79	(+4)
No bias	75 (0)	78	(+3)
Slightly PD-biased	85 (+10)	77	(+2)
Extremely PD-baised	95 (+20)	76	(+1)

Implications & Conclusions

- **Generalizing the Contrastive Field based approach?:** the <u>error-driven learning mechanism</u> could happen at multiple linguistic levels, modeled as a field hosting the computation of differences between stimuli from field couplings;
- **Modeling abstract information in the DFT System?**: this study opens up the possibility of representing <u>frequency-based grammatical knowledge</u> within the DFT framework.
- **Unifying Two Mechanisms into One Model?:** past works have proposed two separate mechanisms for the different priming effects, one *activation*-based, and the other *error-driven learning*-based; I showed that they can be unified.
- **Future**: explicitly modeling the "*learning*" process by adding memory traces.

Thanks for Listening!
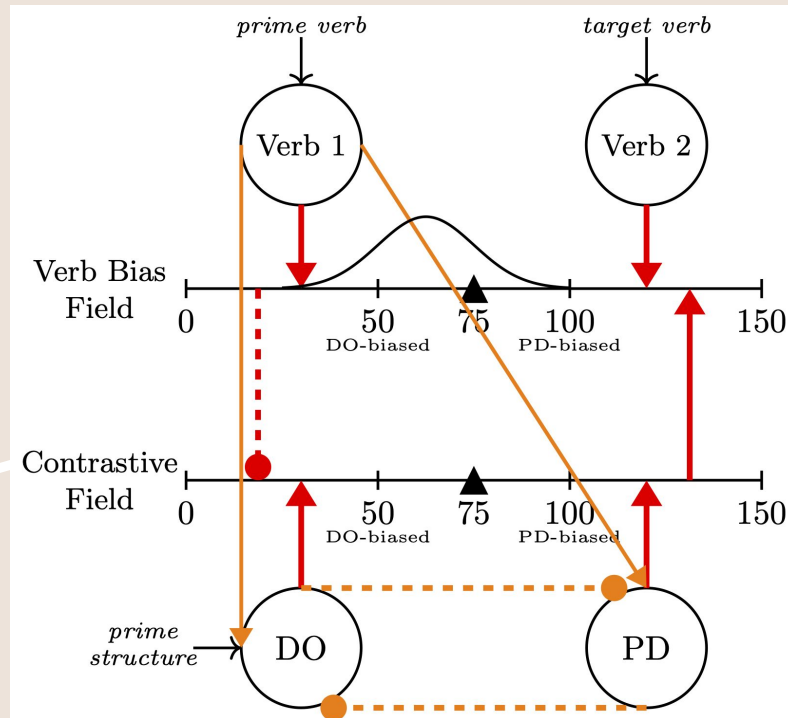
Appendix: Parameters

Parameter	Value
τ	20
h	-5
β	4
q	0.1
range	[-10, 10]
resolution	0.05
c_{11}	6
c_{21}	-3

(a) Parameter values for a node.

Parameter	Value
τ	20
h	-5
β	4
q	1
size	150

(b) Parameter values for a DNF.

Parameter	Value
c_{exc}	20
c_{inh}	5
c_{glob}	0.9
σ_{exc}	5
σ_{inh}	15

(a) Parameters for the selection kernel (used for the **[Verb Bias]** field.

Parameter	Value
c_{exc}	17.5
c_{inh}	15
c_{glob}	0
σ_{exc}	5
σ_{inh}	10

(b) Parameters for the selection kernel (used for the **[Contrastive]** field.