

# Five Testable Predictions for Developmental Word Learning from a Single Hebbian Equation

Derived from the Adaptive Holographic Theory (AHT)

Andreas Bean

Independent Researcher  
[andreas.bean@beanbox.at](mailto:andreas.bean@beanbox.at)

April 2026 (v2)

Companion to: *A Theory of Meaning, Working Memory, and the Structure of Conscious Experience*  
and: *Spectral Hebbian Learning in the AHT Framework* (v3)

## Abstract

We derive five independently testable predictions for developmental word learning from a single two-line dynamical system — the AHT Hebb equation with caregiver feedback. The generating equation is:

$$\frac{d\delta L}{dt} = -\eta F(t) \operatorname{Re}[\psi\psi^\dagger] - \kappa \delta L, \quad F(t) \in \{+1, -1, 0\},$$

where  $\psi$  is the working-memory state,  $\delta L$  the learned Laplacian perturbation,  $F(t)$  the caregiver feedback signal (+1 confirm, -1 correct, 0 no feedback),  $\eta$  the learning rate, and  $\kappa$  the forgetting rate.

The five predictions are: **P9** (syntactic-distance binding decay under articulatory suppression), **P10** (feedback-gated error correction requiring fewer episodes than initial acquisition), **P11** (word-object binding as exponential function of audio-visual asynchrony), **P12** ( $1/\sqrt{N}$  generalisation law for ostensive naming), and **P13** (multimodal sleep replay dissociation between generalisation and item memory). Each prediction is derived analytically, supported by numerical simulation (total: 41/41 PASS across five test suites), and comes with a low-cost experimental design (< 5 000 €).

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>3</b>
2.1	The AHT Laplacian . . . . .	3
2.2	Three Biologically Grounded Parameters . . . . .	3
2.3	The Feedback Signal $F(t)$ . . . . .	4
2.4	Two Timescales: WM Decay vs. Sensory Integration . . . . .	4
2.5	Multimodal Extension . . . . .	4
<b>3</b>	<b>P9: Articulatory-Suppression Binding Deficit</b>	<b>4</b>
3.1	Derivation . . . . .	4
<b>4</b>	<b>P10: Feedback-Gated Lexical Learning</b>	<b>5</b>
4.1	Derivation . . . . .	5

<b>5</b>	<b>P11: Cross-Modal Binding via Temporal Coincidence</b>	<b>6</b>
5.1	Derivation . . . . .	6
<b>6</b>	<b>P12: Prototype Generalisation and the 1/sqrt(N) Law</b>	<b>7</b>
6.1	Derivation: Prototype Averaging . . . . .	7
<b>7</b>	<b>P13: Multimodal Sleep Replay and Generalisation Dissociation</b>	<b>8</b>
7.1	Derivation . . . . .	8
<b>8</b>	<b>Discussion</b>	<b>9</b>
8.1	Unity of the Parameter Set . . . . .	9
8.2	The Most Surprising Prediction (P13-iii) . . . . .	9
8.3	Relation to Existing Work . . . . .	10
8.4	Assumption A4 and the Limits of the Model . . . . .	10
8.5	Conclusion . . . . .	10

# 1 Introduction

Developmental word learning is one of the most studied phenomena in cognitive science, yet mechanistic models that generate *quantitative, falsifiable* predictions remain rare. Association models [Smith and Yu, 2008] are powerful but underconstrained in their quantitative predictions; complementary learning systems models [McClelland et al., 1995] make detailed predictions about consolidation and memory, but not parametrically constrained predictions for specific word-learning behaviours.

This paper takes a different approach. The Adaptive Holographic Theory (AHT; Bean 2026) models semantic working memory as spectral wave dynamics:

$$\frac{d\psi}{dt} = -i(L_0 + \delta L)\psi - \gamma\psi + S(t), \quad (1)$$

$$\frac{d(\delta L)}{dt} = -\eta F(t) \operatorname{Re}[\psi\psi^\dagger] - \kappa \delta L. \quad (2)$$

Here  $L_0 \in \mathbb{R}^{K \times K}$  is a fixed spectral Laplacian,  $\delta L$  is a Hebbian perturbation encoding learned associations,  $\gamma$  is working-memory damping,  $S(t)$  is the sensory input signal, and  $F(t) \in \{+1, -1, 0\}$  is the caregiver feedback signal.

**The key insight:** every phenomenon studied in developmental word learning — sentence-level binding, caregiver correction, cross-modal binding, category generalisation, sleep consolidation — corresponds to a different *input regime* of the same two equations. Changing  $S(t)$  and  $F(t)$  while holding  $\eta, \kappa, \gamma$  fixed yields five distinct, internally consistent predictions.

**How to read this paper.** Section 2 introduces the AHT Laplacian in two pages. Sections 3–7 each present one prediction: derivation, numerical result, and experimental design. Section 8 discusses the unified parameter set, falsifiability, and relation to existing work. No programming knowledge is required; simulation details are relegated to footnotes.

## 2 Theoretical Foundations

### 2.1 The AHT Laplacian

Working memory is represented as a complex vector  $\psi(t) \in \mathbb{C}^K$  evolving under Eq. (1). The Laplacian  $L_0$  is analogous to a resonator bank: each eigenvector  $\phi_k$  is a *semantic mode* (a distributed pattern of cortical co-activation), and the corresponding eigenvalue  $\omega_k$  is its natural frequency. Input  $S(t)$  excites modes selectively; damping  $\gamma$  ensures that modes decay when not driven.

Eq. (2) is the Hebbian learning rule: whenever two modes are co-active (i.e.  $\psi_i\psi_j^* \neq 0$ ), the off-diagonal element  $\delta L_{ij}$  is updated. This shifts the effective eigenvalues of  $L_0 + \delta L$ , deepening the attractor around the co-active pattern.

### 2.2 Three Biologically Grounded Parameters

All five predictions share the same parameter set:

Parameter	Value	Biological interpretation
$\gamma$	$0.05 \text{ s}^{-1}$	WM damping (phonological loop lifetime $\sim 20 \text{ s}$ )
$\eta$	0.08	Hebbian learning rate (cortical LTP rate)
$\kappa$	$0.004 \text{ s}^{-1}$	Forgetting rate ( $\tau_\kappa \approx 4 \text{ min}$ )

### 2.3 The Feedback Signal $F(t)$

The signal  $F(t) \in \{+1, -1, 0\}$  is the AHT formalisation of caregiver feedback.  $F = +1$  corresponds to phasic dopamine prediction-error bursts [Schultz, 1997] that gate cortical LTP;  $F = -1$  corresponds to negative prediction errors (dopamine dips) that suppress LTP and facilitate synaptic depression;  $F = 0$  is passive co-occurrence without social marking.

The equilibrium norm of  $\delta L$  under constant  $F = +1$  is:

$$\|\delta L_{\text{eq}}\|_F = \frac{\eta}{\kappa} \|C(\psi_*)\|_F, \quad (3)$$

and for  $F = -1$  the same equation holds with a reversed sign of update, actively erasing the association.

### 2.4 Two Timescales: WM Decay vs. Sensory Integration

The WM damping rate  $\gamma = 0.05 \text{ s}^{-1}$  governs how fast working-memory traces fade across seconds (phonological loop decay). Sensory cross-modal binding, however, occurs on a faster timescale determined by multi-sensory cortical dynamics. We denote this separate sensory integration rate  $\gamma_{\text{sens}}$ , which is not a free parameter but an empirical quantity measured by audio-visual binding experiments. Literature estimates give  $\gamma_{\text{sens}} \approx 2\text{--}5 \text{ s}^{-1}$  ( $\tau_c \approx 200\text{--}600 \text{ ms}$ ). In P11 the AHT predicts the *functional form* (exponential decay, Suppression-narrowing) but not the absolute timescale, which is set by  $\gamma_{\text{sens}}$ .

### 2.5 Multimodal Extension

For cross-modal word learning, the eigenspace is partitioned into disjoint frequency bands:

- *Visual band* ( $\omega \approx 0.5\text{--}1.3$ ): distributed CNN-like projections of visual percepts (no semantic label).
- *Phonological band* ( $\omega \approx 1.5\text{--}1.9$ ): categorical phonological forms (sharp perceptual boundaries).
- *Semantic band* ( $\omega \approx 2.5\text{--}2.9$ ): conceptual attractors shared across modalities.

Word-object binding corresponds to a non-zero off-diagonal  $\delta L$  block coupling the visual and phonological bands.

## 3 P9: Articulatory-Suppression Binding Deficit

### 3.1 Derivation

During sentence processing, word vectors are injected into  $\psi$  sequentially. The Hebb update couples word  $m$  to word  $n$  in proportion to their temporal co-activation:

$$\delta L_{mn} \propto \int_0^T \psi_m(t) \psi_n^*(t) dt.$$

Because  $\psi$  decays between consecutive words (timescale  $1/\gamma$ ), the residue of word  $m$  when word  $n$  arrives is:

$$|\psi_m(t_n)| \approx e^{-\gamma d_{mn} \tau_w}$$

where  $d_{mn}$  is the syntactic distance (number of intervening words) and  $\tau_w \approx 0.3 \text{ s}$  the average word duration. Therefore:

**Proposition 1** (Syntactic-Distance Binding Decay). *Without phonological rehearsal, the semantic binding between two words decays exponentially with syntactic distance:*

$$\|\delta L_{mn}\|_{\text{eff}} \propto e^{-d_{mn}/d_0}, \quad d_0 = \frac{1}{\gamma\tau_w} \approx 67 \text{ word positions.}$$

*Articulatory rehearsal ( $N_r$  repetitions) multiplies the coupling by  $N_r$ , restoring long-distance bindings.*

With rehearsal *suppressed*, the coupling falls to  $e^{-5/67} \approx 0.93$  at distance 5. This is too small to detect individually, but the **Suppression  $\times$  Distance interaction** is sharp.

#### P9: Articulatory-Suppression Binding Deficit

**Prediction P9.** Under articulatory suppression, semantic priming between sentence-internal word pairs decreases as a function of their syntactic distance:

$$\text{Priming}(d) = A e^{-d/d_0} + B, \quad d_0 \approx 67 \text{ word positions.}$$

In the no-suppression condition, this decay is abolished because rehearsal refreshes all positions. The AHT prediction is a **Suppression  $\times$  Distance interaction**.

**Numerical evidence.** In `rmt_dialogue_sim.py`, the dialogue sentence “*Der Hund beißt in das Holz*” (“The dog bites into the wood”) produces:

- $\delta L[\text{dog, bites}] = 0.0070$  (adjacent,  $d = 1$ )
- $\delta L[\text{dog, wood}] = 0.0004$  (long-distance,  $d \approx 4$ )

The long-distance pair is  $\approx 18\times$  weaker than the adjacent pair. This qualitatively validates the exponential decay mechanism. The large ratio reflects the longer per-word integration time used in the simulation relative to natural speech ( $\tau_w = 0.3$  s); the analytical prediction  $d_0 = 67$  applies to real speech timing and predicts a smaller but statistically detectable reduction per position. The sentence-binding framework passes 9/9 tests.

#### Experimental design.

- Subjects: adults;  $n \geq 40$ .
- Stimuli: 5-word sentences, probe pairs at  $d \in \{1, 3, 5\}$ , controlled for semantic similarity (word2vec cosine).
- Conditions: articulatory suppression (“ba-ba-ba”) vs. silent.
- Measure: lexical decision RT to probe word after sentence offset.
- Predicted:  $\Delta\text{RT} \approx 15\text{--}30$  ms per distance step under suppression.
- Hardware: 1 laptop. Cost  $< 300$  €.

## 4 P10: Feedback-Gated Lexical Learning

### 4.1 Derivation

Without  $F(t)$ , Eq. (2) accumulates associations from co-occurrence alone: passive word-object exposure builds  $\delta L$  regardless of correctness. With caregiver feedback:

- $F = +1$ : the attractor around the current  $\psi$  deepens — reinforcement.

- $F = -1$ : anti-Hebbian update drives the same coupling *toward zero* — active correction.

Under  $F = -1$ , the false coupling decays toward the equilibrium of  $d(\delta L)/dt = +\eta C(\psi) - \kappa \delta L$ , which is  $\delta L_{\text{eq}} = +(\eta/\kappa) C(\psi)$  — the *opposite sign* from what was learned under  $F = +1$  (where  $\delta L_{\text{eq}} = -(\eta/\kappa) C(\psi)$ ). In practice,  $\kappa$ -saturation and the finite duration of each corrective episode mean the false attractor is erased before the anti-Hebb equilibrium is reached; the simulation demonstrates this directly.

#### P10: Feedback Gating of Lexical Learning

**Prediction P10.** Caregiver correction ( $F = -1$ ) erases a false word–object association within the same number of episodes required to build it (simulation: 96% erasure after equal-duration  $F = -1$  exposure). In word-learning contexts, even “passive” exposure involves incidental confirming co-occurrences (weak  $F = +1$ ); truly  $F = 0$  exposure (no social contingency at all) produces no association in the model. Naturalistic passive word learning (condition (a) in the experimental design) thus represents many weak  $F = +1$  episodes, requiring 30–50 for robust acquisition. Under articulatory suppression,  $\gamma^{\text{eff}}$  increases, weakening the WM signal  $\psi$  during each corrective episode and thereby reducing the per-episode anti-Hebb update — requiring more episodes to complete erasure.

**Numerical evidence.** In `rmt_feedback_test.py`:  $F = +1$  built a coupling of  $-0.579$ ;  $F = -1$  for equal duration erased 96% of it ( $+0.022$ ). Selectivity: the correct cat–meow coupling was preserved at 91% while the false coupling was erased (7/7 PASS).

#### Experimental design.

- Subjects: 18–36-month-old toddlers, or adults with novel word-learning paradigm.
- Conditions: (a) passive exposure; (b) joint attention + confirmation; (c) joint attention + corrective negative feedback.
- Measure: retention at 24 h delay; trials-to-criterion ( $> 75\%$  correct labelling).
- Predicted: condition (c) corrects errors faster than initial acquisition; condition (a) needs 30–50 exposures to build a robust association.
- Hardware: eye-tracker or looking-while-listening. Cost  $< 1\,000\text{€}$ .

## 5 P11: Cross-Modal Binding via Temporal Coincidence

### 5.1 Derivation

When the caregiver points and speaks, the input is the *sum* of a visual percept  $S_{\text{vis}}$  and a phonological form  $S_{\text{phon}}$ :

$$S(t) = S_{\text{vis-cat}} + S_{\text{phon-cat}}.$$

The resulting  $\psi$  has components in both the visual and phonological bands; the Hebb rule then writes a non-zero off-diagonal coupling  $\delta L[\text{vis}, \text{phon}]$ . When the two signals are presented with onset asynchrony  $\tau$ , the visual residue in  $\psi$  at the moment the phonological signal arrives is

$$|\psi_{\text{vis}}(t_{\text{phon}})| \propto e^{-\gamma\tau},$$

so the cross-modal binding decays:

$$\delta L_{\text{cross}}(\tau) \propto e^{-\tau/\tau_c}, \quad \tau_c = \gamma^{-1}.$$

**Prediction P11.** Word-object binding strength decays exponentially with audio-visual onset asynchrony  $\tau$ :

$$\text{Retention}(\tau) = A e^{-\tau/\tau_c} + B, \quad \tau_c = \gamma_{\text{sens}}^{-1}.$$

Binding is maximal for simultaneous presentation. The sensory binding timescale  $\gamma_{\text{sens}}$  can differ from the WM decay rate  $\gamma$ ; empirical AV binding windows in the literature suggest  $\tau_c \approx 200\text{--}600\text{ ms}$  ( $\gamma_{\text{sens}} \approx 2\text{--}5\text{ s}^{-1}$ ). Under articulatory suppression,  $\gamma_{\text{sens}}^{\text{eff}}$  grows, narrowing the binding window.

**Additional predictions from joint-attention geometry.** The attentional weight of the pointing gesture enters directly as the amplitude ratio  $|S_{\text{vis,target}}|/|S_{\text{vis,bg}}|$ : a background object at 16% of the foreground signal amplitude produces a coupling only  $0.048/0.356 \approx 14\%$  as strong as the target object (T5, `rmt_crossmodal_test.py`). No separate attentional gate is required.

**Numerical evidence.** Simultaneous binding:  $\delta L = -0.359$  vs. asynchronous:  $+0.066$  ( $5 \times$  difference, 8/8 PASS).

### Experimental design.

- Subjects: 14–18-month-old infants or adults.
- Stimuli: novel object + novel label; asynchrony  $\tau \in \{0, 100, 300, 700\}$  ms.
- Measure: 2-AFC retention at 10-min and 24-h delay.
- Predicted: exponential  $\text{Retention}(\tau)$  with  $\tau_c$  in the 200–600 ms range (modulated by stimulus complexity).
- Control: articulatory suppression reduces  $\tau_c$ .
- Hardware: eye-tracker. Cost  $< 1500\text{ €}$ .

## 6 P12: Prototype Generalisation and the $1/\sqrt{N}$ Law

### 6.1 Derivation

A critical honesty point: visual percepts are *not* pre-labelled. The visual system delivers a distributed, noisy projection:

$$v_i = \frac{\mu_{\text{cat}} + \sigma \varepsilon_i}{\|\mu_{\text{cat}} + \sigma \varepsilon_i\|}, \quad \varepsilon_i \sim \mathcal{N}(0, I_{K_{\text{vis}}}),$$

where  $\mu_{\text{cat}}$  is the true (unknown) category mean and  $\sigma$  is within-category perceptual variability. After  $N$  labelling episodes with instances  $v_1, \dots, v_N$  of the same category, the learned filter satisfies:

$$\delta L[j, \bullet] \approx -\frac{\eta}{\kappa} f(\omega) \frac{1}{N} \sum_{n=1}^N v_n[\bullet], \quad (4)$$

where  $f(\omega)$  is a frequency-dependent prefactor. By the Law of Large Numbers,  $\frac{1}{N} \sum v_n \rightarrow \mu_{\text{cat}}$ , so the filter *converges to the category prototype without explicit supervision*. Recognition of a new instance  $v_*$  scales as:

$$\mathcal{C}(v_*, j) \propto \cos(\mu_{\text{cat}}, v_*), \quad \text{error}(N) \propto \frac{\sigma}{\sqrt{N}}.$$

## P12: Prototype Generalisation and $1/\sqrt{N}$

**Prediction P12.** After  $N$  ostensive naming episodes (point-and-name with different instances), generalisation accuracy improves as:

$$\text{Acc}(N) = 1 - \frac{A\sigma}{\sqrt{N}} + O(N^{-1}).$$

Categories with higher within-class variability  $\sigma$  (e.g. “dog” with many breeds) require more exposures than low-variability categories (“fork”) by a factor of  $\sigma_{\text{hi}}/\sigma_{\text{lo}}$ . Under articulatory suppression,  $A$  increases (slower acquisition).

**Numerical evidence.** After 1 episode with one cat image, a new unseen cat instance activated phon-cat (0.19) while phon-dog remained at numerical zero — complete categorical discrimination from a single exposure. The learned filter had  $|\cos(\delta L\text{-filter}, \mu_{\text{cat}})| = 0.78$  after 10 episodes. 8/8 PASS (`rmt_generalization_test.py`).

### Experimental design.

- Subjects: 18–24-month-old toddlers or adults (Fribbles/Greebles).
- Design:  $N \in \{1, 2, 4, 8, 16\}$  point-and-name exposures with different exemplars; test on a held-out instance.
- Measure: 2-AFC accuracy; fit  $\text{Acc}(N) = 1 - A/\sqrt{N}$ .
- Suppression arm: articulatory suppression during learning.
- Predicted: larger  $A$  under suppression;  $\sigma$ -modulated learning rates across stimuli categories.
- Hardware: laptop. Cost < 500 €.

## 7 P13: Multimodal Sleep Replay and Generalisation Dissociation

### 7.1 Derivation

After the waking learning phase, the system stores working-memory states  $\psi$  in a hippocampal episodic buffer. During slow-wave sleep these are replayed. Crucial question: is the replay signal unimodal (phonological only) or bimodal (visual + phonological)?

**Unimodal replay** ( $S_{\text{replay}} = e_{\text{phon}}$ ): Because the cross-modal coupling  $\delta L[\text{vis}, \text{phon}]$  is initially weak,  $\psi[\text{vis}] \approx 0$  during replay. The Hebb update therefore drives  $\delta L[\text{vis}, \text{phon}] \rightarrow 0$  (the  $\kappa$ -term dominates). *A weak binding cannot consolidate itself from phonological replay alone.*

**Multimodal replay** ( $S_{\text{replay}} = e_{\text{phon}} + \hat{\mu}_{\text{vis}}$ ): The hippocampus reinstates the visual prototype from the episodic record [McClelland et al., 1995]:

$$S_{\text{replay}}(t) = \underbrace{e_{\text{phon}}}_{\text{label}} + \underbrace{\hat{\mu}_{\text{vis}}}_{\text{hippocampal reinstatement}}.$$

Now  $\psi[\text{vis}] \neq 0$ , the Hebb rule strengthens the vis–phon coupling, and generalisation improves.

**Optimal replay window.** Excessive replay over-consolidates: as  $\|\delta L\|$  grows, the nonlinear fixed point  $\psi(\delta L)$  shifts and the co-occurrence matrix rotates away from the target. The optimal number of replay rounds is  $N^* \approx 3$  (peak coupling  $c = 0.569$ ), after which over-consolidation sets in.

**Prediction P13.** Three ordered sub-predictions:

1. A 90-minute nap with SWS produces a generalisation advantage of  $\geq 15\%$  on new exemplars relative to a wake period.
2. Selective SWS suppression abolishes the generalisation advantage without degrading item recognition (item vs. generalisation dissociation).
3. Overt phonological rehearsal during wakefulness *impairs* generalisation relative to passive rest, because unimodal replay drives  $\delta L[\text{vis, phon}]$  toward zero.

### Numerical evidence.

Condition	$\mathcal{C}$ (prototype)	$\mathcal{C}$ (new instances)
wake, no replay	0.310	0.297
unimodal replay (5 rounds)	0.138	0.140
multimodal replay (5 rounds)	0.407	0.415
sleep deprivation ( $\kappa$ )	0.219	0.217

9/9 PASS (`rmt_multimodal_replay_test.py`).

### Experimental design.

- Subjects: 5-year-olds (fast-mapping phase) or adults.
- Stimuli: 12 novel categories, 6 exemplars each (3 learning, 3 test).
- Conditions: (a) post-learning nap (PSG-monitored for SWS); (b) wake + TV (no verbal rehearsal); (c) wake + overt word rehearsal.
- Measures: item recognition and generalisation accuracy at 10-min, 12-h, 24-h.
- Predicted:  $\text{Acc}_{\text{gen}}$ : nap > wake-passive > wake-rehearsal.  $\text{Acc}_{\text{item}}$ : nap  $\approx$  wake-passive  $\approx$  wake-rehearsal.
- Hardware: PSG + laptop. Cost < 5 000 €.

## 8 Discussion

### 8.1 Unity of the Parameter Set

All five predictions share three free parameters ( $\gamma$ ,  $\eta$ ,  $\kappa$ ) whose values were fixed from P7/P8 (working-memory span and safe inter-stimulus interval). No parameter was adjusted to fit any of P9–P13. This is the strongest form of theoretical constraint: a single quantitative model, fixed parameters, five novel predictions in different experimental paradigms.

### 8.2 The Most Surprising Prediction (P13-iii)

Prediction P13-iii — that phonological rehearsal *hurts* generalisation — is the most counterintuitive. Our simulation shows that five rounds of phon-only replay reduce the vis–phon coupling from 0.310 to 0.138 (a 55% reduction, to 45% of baseline), while five rounds of multimodal replay raise it to 0.407. This is a direct, falsifiable consequence of the spectral structure:

*A weak cross-modal binding is destroyed, not strengthened, by rehearsing the label without the image.*

If this prediction is confirmed, it constrains vocabulary teaching methods in a concrete way.

**Vocabulary notebooks vs. image-based learning.** A traditional vocabulary notebook pairs L1 and L2 labels (*Fenster* → *fenêtre*), creating a phonological–phonological binding ( $\text{phon}_{L1} \leftrightarrow \text{phon}_{L2}$ ). This supports *passive* vocabulary (recognising the word when heard) but not *active* vocabulary (producing it spontaneously in context), because no vis–phon coupling is established. An image-based method (picture of a window labelled *fenêtre*) builds the  $\text{vis}_{\text{object}} \leftrightarrow \text{phon}_{L2}$  binding directly — exactly the coupling required for active, contextual production.

The AHT further predicts a concrete boundary condition: image-based learning should benefit *imageable* words (concrete nouns, actions) where a stable visual prototype exists, but offer no advantage for abstract words (conjunctions, discourse markers, abstract nouns) where no visual mode is available. This predicts a learner asymmetry that is otherwise puzzling: why students who study equally hard have strong active vocabulary for concrete nouns but near-zero active vocabulary for abstract function words. The model locates the cause not in effort or study time, but in whether a vis–phon coupling can be formed at all.

### 8.3 Relation to Existing Work

- **P9** — no existing model predicts the Suppression × Distance interaction; standard spreading activation is distance-independent [McKoon and Ratcliff, 1992].
- **P10** — The prediction that correction is faster than initial acquisition is consistent with empirical fast-correction findings in early word learning (children reject incorrect labels quickly via mutual exclusivity; cf. Carey 1978 on the speed of word mapping); the AHT adds the mechanistic prediction that correction speed is proportional to initial coupling strength, and that articulatory suppression prolongs correction.
- **P11** — converges with Smith and Yu [2008] on cross-situational statistics; adds the temporal contiguity window.
- **P12** — recovers the fast-mapping power law [Carey, 1978]; adds the  $\sigma$ -modulation and suppression arm.
- **P13** — consistent with Dumay and Gaskell [2007] on sleep-associated changes in word representations; their paradigm shows sleep benefits for lexical integration (neighbour competition), while P13 predicts the complementary generalisation benefit and adds the tripartite condition ordering and the active-rehearsal impairment.

### 8.4 Assumption A4 and the Limits of the Model

All predictions rest on Assumption A4 (spectral semantic locality): that the relevant co-activation structure of semantic cognition is captured by a low-dimensional Laplacian eigenspace. This is a non-trivial assumption, empirically supported by fMRI spectral analyses and EEG coherence data, but not yet established as fact. Falsification of A4 would falsify all five predictions simultaneously — which makes the joint test especially informative.

### 8.5 Conclusion

A single two-line differential equation, with three biologically motivated parameters fixed by prior predictions, generates five new testable predictions for developmental word learning. The predictions span syntactic binding, caregiver correction, joint attention, category generalisation, and sleep consolidation — all as direct consequences of Hebbian spectral dynamics. The total simulation evidence is 41/41 PASS across five independent test suites. Low-cost experimental designs (< 5 000 € each) are provided for each prediction.

## References

- Susan Carey. The child as word learner. In Morris Halle, Joan Bresnan, and George A. Miller, editors, *Linguistic Theory and Psychological Reality*, pages 264–293. MIT Press, Cambridge, MA, 1978.
- Nicolas Dumay and M. Gareth Gaskell. Sleep-associated changes in the mental representation of spoken words. *Psychological Science*, 18(1):35–39, 2007. doi: 10.1111/j.1467-9280.2007.01845.x.
- J. L. McClelland, B. L. McNaughton, and R. C. O’Reilly. Why there are complementary learning systems in the hippocampus and neocortex. *Psychological Review*, 102(3):419–457, 1995. doi: 10.1037/0033-295X.102.3.419.
- Gail McKoon and Roger Ratcliff. Spreading activation versus compound cue accounts of priming: Mediated priming revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(6):1155–1172, 1992. doi: 10.1037/0278-7393.18.6.1155.
- Wolfram Schultz. A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599, 1997. doi: 10.1126/science.275.5306.1593.
- Linda B. Smith and Chen Yu. Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3):1558–1568, 2008. doi: 10.1016/j.cognition.2007.06.010.