Grammatical Theory with Gradient Symbol Structures

The GSC Research Group

Paul Smolensky	Géraldine Legendre
Matt Goldrick	Colin Wilson
Kyle Rawlins	Ben Van Durme
Akira Omaki	Paul Tupper
Don Mathis	Pyeong-Whan Cho
Laurel Brehm	Nick Becker
Drew Reisinger	Emily Atkinson
Matthias Lalisse	Eric Rosen

Belinda Adam

Research Institute for Linguistics *12 January 2016* *Hungarian Academy of Science*

Context of the work

Problem: crisis of cognitive architecture. Unify symbolic & neural-network (NN) computation

Proposal: Gradient Symbolic Computation (**GSC**), a cognitive architecture

- Representation: symbol structures as vectors—Tensor Product Representations (TPRs)
- Knowledge: weighted constraints–-probabilistic Harmonic Grammars (HGs)
- Processing:
    - (1) (Multi-)linear feed-forward NNs
    - (2) Stochastic feed-back (higher-order) NNs

Smolensky, Goldrick & Mathis 2014 *Cognitive Science*
Smolensky & Legendre 2006 *The Harmonic Mind* MIT Press

Context of the work

Problem: crisis of cognitive architecture. Unify symbolic & neural-network (NN) computation

Proposal: Gradient Symbolic Computation (**GSC**), a cognitive architecture

- Representation: symbol structures as vectors—Tensor Product Representations (TPRs)
- Knowledge: weighted constraints–-probabilistic Harmonic Grammars (HGs)
- Processing:
  - (1) (Multi-)linear feed-forward NNs
  - (2) Stochastic feed-back (higher-order) NNs

Tests:

- symbolic side
  - ➤ computation
    - ✦ (1) can compute: ("primitive") recursive functions, β-reduction, tree adjoining, inference
    - ✦ (2) can specify/asymptotically compute: formal languages (type 0)
  - ➤ linguistic theory: HG/OT work in phonology, …, pragmatics
- NN side
  - ➤ computation
    - ✦ theory: stochastic convergence to global optima of Harmony
    - ✦ NLP applications (MS): question answering, semantic parsing (related: vector semantics etc.)
  - ➤ cognitive neuroscience: stay tuned (limited extant evidence)
- Together: (currently) psycholinguistics of sentence production & comprehension

Prediction: blended, gradient symbol structures play an important role in cognition

- NNs: phonetics, psycholinguistics: interaction of gradience & structure-sensitivity
- symbolic level, phonology: gradience in lexical representations & **French liaison**

Context of the work

Problem: crisis of cognitive architecture. Unify symbolic & neural-network (NN) computation

Proposal: Gradient Symbolic Computation (**GSC**), a cognitive architecture

Why go beyond classical symbol structures in grammatical theory?

Fundamental issue: Symbolic analyses in linguistics often offer tremendous insight, but typically *they don't **quite** work.*

Hypothesis: Blended, gradient symbol structures can help resolve long-standing impasses in linguistic theory.

Problem: Competing analyses posit structures A and B to account for X

Proposal: X actually arises from a gradient blend of structures A and B

Today: X = French liaison (& elision); Cs (& Vs) that ~ Ø; e.g., *peti t ami* ~ *peti copain*

A = underlyingly, *petit* is /pøti𝕋/ with deficient final *t*; *ami* is /ami/

B = underlyingly, *petit* is /pøti/; *ami* is {/tami/ (~ /zami/, /nami/, /ami/}

- cognitive neuroscience: stay tuned (limited extant evidence)

- Together: (currently) psycholinguistics of sentence production & comprehension

Prediction: blended, gradient symbol structures play an important role in cogn…

- NNs: phonetics, psycholinguistics: interaction of gradience & structure-sensit…
- symbolic level, phonology: gradience in lexical representations & **French liaison**

Thanks to Jennifer Culbertson

Context of the work

Problem: crisis of cognitive architecture. Unify symbolic & neural-network (NN) computation

Proposal: Gradient Symbolic Computation (**GSC**), a cognitive architecture

Why go beyond classical symbol structures in grammatical theory?

Fundamental issue: Symbolic analyses in linguistics often offer tremendous insight, but typically *they don't **quite** work.*

Hypothesis: Blended, gradient symbol structures can help resolve long-standing impasses in linguistic theory.

Problem: Competing analyses posit structures A and B to account for X

Proposal: X actually arises from a gradient blend of structures A and B

Today: X = French liaison (& elision); Cs (& Vs) that ~ Ø; e.g., *peti t ami* ~ *peti copain*

A = underlyingly, *petit* is /pøti$\mathbb{T}$/ with deficient final *t*; *ami* is /ami/

B = underlyingly, *petit* is /pøti/; *ami* is {/tami/ (~ /zami/, /nami/, /ami/}

See also Hankamer, Jorge. 1977. Multiple Analyses. In Charles Li (ed.) *Mechanisms of Syntactic Change*, pp. 583–607. University of Texas Press.

"we must give up the assumption that two or more conflicting analyses cannot be simultaneously correct for a given phenomenon" (pp. 583–4)

"such constructions have both analyses at once (in the **conjunctive** sense)" (p. 592)

Goals of the work

Show how Gradient Symbolic Representations (GSRs)

- enable enlightening accounts of many of the phenomena that have been claimed to occur in the rich scope of liaison
- putting aside the many divergent views on the actual empirical status of these alleged phenomena

The theoretical divergences in this field illustrate well how symbolic representations don't *quite* work.

➤ Can GSC help resolve these disputes?

Talk goal: show what GSRs *can do* in the analysis of liaison.

A *theoretical exploration* — **not** an *empirical argument*!

- The facts are much too murky for me to even attempt a definitive empirical argument (but stay tuned).
- Also, it takes considerable theoretical exploration of a new framework before it's appropriate to seek empirical validation.

Inspiration

Dowty sketch re: structural ambivalence (*PP complement vs. adjunct*)

Dowty, David. 2003. The Dual Analysis of Adjuncts/Complements in Categorial Grammar. In Ewald Lang, Claudia Maienborn, Cathrine Fabricius-Hansen, eds., *Modifying Adjuncts*. pp. 33–66. Mouton de Gruyter.

Inspiration

Dowty sketch re: structural ambivalence (*PP complement vs. adjunct*)

- children form an initial simple, maximally general, analysis
  - ➤ *adjuncts:* compositional semantics
- adults end up with a more complex, specialized analysis
  - ➤ *complements:* idiosyncratic semantics

but:

- ➤ general analysis persists in adulthood
- ➤ co-exists with more complex analysis
- ➤ the two blend and function jointly

"in some subtle psychological way, in on-line processing—though in a way that only connectionism or some other other future theories of the psychology of language can explain." [antepenultimate paragraph, yellow added]

Inspiration

Dowty sketch re: structural ambivalence (*PP complement vs. adjunct*)

- children form an initial simple, maximally general, analysis
  - ➤ *adjuncts:* compositional semantics
- adults end up with a more complex, specialized analysis
  - ➤ *complements:* idiosyncratic semantics

but:

- ➤ general analysis persists in adulthood
- ➤ co-exists with more complex analysis
- ➤ the two blend and function jointly

Here, formalize the adult blend, speculate about acquisition [*skip?*]

- liaison in French
  - ➤ ultimately involves prosody [*skip?*]

Outline

① Gradient Symbolic Computation in grammar: Nano-intro

② The adult blend: A gradient grammar of French liaison

Ⓐ The phonological phenomenon

Ⓑ GSC analysis: Idea

Ⓒ GSC analysis: Formal account

③ Acquisition: Speculations on formalizing Dowty's sketch [*skip (1)?*]

④ Prosody: Tentative suggestions [*skip (6)?*]

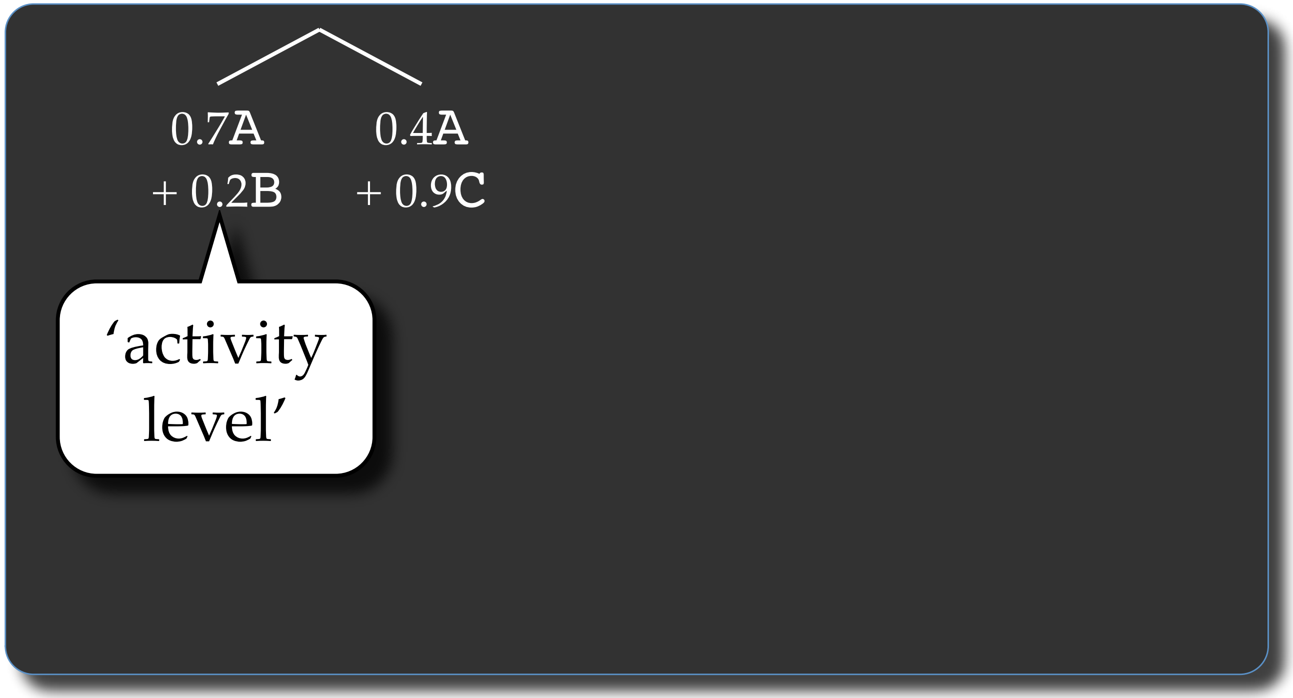
⑤ Summary

①
Gradient Symbolic Computation in grammar

Nano-intro

❶ Informal introduction to GSC

Examples of Gradient Symbolic Representations (GSRs)



❶ Informal introduction to GSC

Examples of Gradient Symbolic Representations (GSRs)

$$\boxed{\begin{array}{c}0.7\mathbf{A}\\ +\ 0.2\mathbf{B}\end{array}} \qquad \begin{array}{c}0.4\mathbf{A}\\ +\ 0.9\mathbf{C}\end{array}$$

Left child role filled
by blend of symbols

Phonology: Elements
change but stay in place

❶ Informal introduction to GSC

Examples of Gradient Symbolic Representations (GSRs)

0.7A 0.4A

+ 0.2B + 0.9C

A in *role* blend:

$0.7r_{\text{left}} + 0.4r_{\text{right}}$

Syntax etc.: Elements change their place (or occupy multiple roles)

❶ Informal introduction to GSC

Examples of Gradient Symbolic Representations (GSRs)

$0.7\text{A} + 0.2\text{B}$	$0.4\text{A} + 0.9\text{C}$

petit ami

$[^{\mathrm{M}}\text{pøti}(\lambda{\cdot}\text{t})]\ \ [^{\mathrm{M}}(\tau{\cdot}\text{t} + \zeta{\cdot}\text{z} + \nu{\cdot}\text{n})\text{ami}]$

A state in GSC is a probability distribution over GSRs

Computation with GS Representations

GSRs are implemented as distributed activity patterns/vectors

- this formalizes 'blend of symbols', 'blend of roles'

A in a *role* blend: $0.7r_{\text{left}} + 0.4r_{\text{right}}$

$0.7\mathrm{A} + 0.2\mathrm{B}$	$0.4\mathrm{A} + 0.9\mathrm{C}$

$$0.7\,\mathbf{A}\otimes\mathbf{r}_{\text{left}} + 0.4\,\mathbf{A}\otimes\mathbf{r}_{\text{right}} + 0.2\,\mathbf{B}\otimes\mathbf{r}_{\text{left}} + 0.9\,\mathbf{C}\otimes\mathbf{r}_{\text{right}}$$

r_{left} hosts a *filler* blend: $0.7\,\mathrm{A} + 0.2\,\mathrm{B}$

Computation with GS Representations

GSRs are implemented as distributed activity patterns/vectors

- this formalizes ‘blend of symbols’, ‘blend of roles’

Dynamics: stochastic optimization

Here **do *not* deal with dynamics**, but exploit the fact that the outcome of the dynamics is

(in the competence-theoretic approximation)

- a representation that **maximizes** well-formedness: ***‘Harmony’*** *H*
- $H(r)$ is the (weighted) sum of violations, by representation r, of constraints $\mathbb{C}_k$
- each $\mathbb{C}_k$ has a numerical weight (H is a *Harmonic Grammar*)

Computation with GS Representations

GSRs are implemented as distributed activity patterns/vectors

- this formalizes 'blend of symbols', 'blend of roles'

Dynamics: stochastic optimization

Here **do *not*** ... e

outcome o

(in the

$HT_{83/86} \rightarrow HG_{90} \rightarrow OT_{91/93} \rightarrow HG_{06}$

but gradient *representations* are new to GSC

☞ here, understanding the HG analysis

- a repres ... ***ny'*** *H*
- $H(r)$ is the (weighted) sum of violations, by repr sentation r, of constraints $\mathbb{C}_k$
- each $\mathbb{C}_k$ has a numerical weight (H is a *Harmonic Grammar*)
- the activity-vector implementation determines how $H(r)$ is computed when r is a GSR

②

The adult blend

Ⓐ The phonological phenomenon

Ⓑ GSC analysis: Idea

Ⓒ GSC analysis: Formal account

A gradient grammar of French liaison

Ⓐ The phonological phenomenon: Core

Latent consonants in French (liaison)

Core phenomena **Universal σ well-formedness: ONSET, NOCODA**

no coda, onset		coda, onset	no coda, no onset
.pø.ti.t̲a.mi.	.pø.ti.k̲o.pɛ̃.	.pø.tit̲.ko.pin.	.pø.ti.e̲.ʁo.
[t]	no [t]	[t]	no [t]
petit ami vs.	*petit copain* vs.	*petite copine* vs.	*petit héro*

[t]: only —V everywhere not —𝕍 (*h-aspiré*)

with *peti(t)*, final /t/ only surfaces 'when needed for syllable onset'
but before *héro*, no /t/ despite lacking onset (ʔ typically absent)
with *petite*, final /t/ always surfaces, even in coda

What is the (t) vs. t distinction in underlying (stored lexical) form?

- 'liaison' ℒ [*petit*] vs. 'fixed' [*petite*] ℱ final consonants

Ⓐ The phonological phenomenon: Core

Latent consonants in French (liaison)

Core mappings

① $v\mathscr{L} + V \rightarrow v.\mathscr{L}V$	*peti(t) + ami*	$\rightarrow$	.pø.ti.t̲a.mi.
② $v\mathscr{L} + c \rightarrow v.c$	*peti(t) + copain*	$\rightarrow$	.pø.ti.ko.pɛ̃.
③ $v\mathscr{L} + \mathbb{V} \rightarrow v.\mathbb{V}$	*peti(t) + $\mathbb{H}$éro*	$\rightarrow$	.pø.ti.e.ʁo.
④ $v\mathscr{F} + c \rightarrow v\mathscr{F}.c$	*petite + copine*	$\rightarrow$	.pø.tit̲.ko.pin.

petit ami vs. *petit copain* vs. *petite copine* vs. *petit héro*

What is the (t) vs. t distinction in underlying (stored lexical) form?

- 'liaison' $\mathscr{L}$ vs. 'fixed' $\mathscr{F}$ final consonants

Ⓐ The phonological phenomenon: Core

Latent consonants in French (liaison)

What is the (t) vs. t distinction in underlying (stored lexical) form?

- ‘liaison’ $\mathcal{L}$ vs. ‘fixed’ $\mathcal{F}$ final consonants

Proposed GSC answer: ***activity level***

$\mathcal{F}$ is a fully active C, but $\mathcal{L}$ is *activity-deficient* — ‘weak’

$\mathcal{L}$ can surface only if it is provided with extra activity

$\mathcal{L}$ is exactly like $\mathcal{F}$ in *content* (a standard C) — but weaker in *activity*.

Ⓐ The phonological phenomenon

Latent consonants in French (liaison)

So far, following orthography, we've assumed a liaison C is *final* in the word it follows

- the **$\hat{W}_1\mathcal{L}$** (or **final-$\mathcal{L}$**) **Analysis**
  - ➤ also take to include *syllabification-driven alternation*

But a number of phonologists reject this theory.

some may find this inelegant

They favor an analysis in which a liaison C is *initial* in the word it precedes

  - ➤ consistent with syllabification
  - ➤ requires lexical entries *ami, tami, zami, nami, …*: allomorph selection is driven by the preceding word
- the **$\mathcal{L}\hat{W}_2$** (or **$\mathcal{L}$-initial**) **Analysis**

Ⓐ The phonological phenomenon

Latent consonants in French (liaison)

So far, following orthography, we've assumed a liaison C is
final in the word it follows

- the $\hat{W}_1\mathcal{L}$ (or **final-$\mathcal{L}$) Analysis**
  - ➤ also take to include *syllabification-driven alternation*

But a number of phonologists reject this theory.

Why? ['*external* evidence']

They favor an analysis in which a liaison C is
initial in the word it precedes

  - ➤ consistent with syllabification
  - ➤ requires lexical entries *ami, tami, zami, nami, …*:
    allomorph selection is driven by the preceding word

- the $\mathcal{L}\hat{W}_2$ (or **$\mathcal{L}$-initial) Analysis**

Ⓐ The phonological phenomenon: Complications

need at least a *3-way* contrast

Trouble for strictly syllabification-driven distribution of $\mathscr{L}$:

⑤ *Phrase-final* $\mathscr{L}$. In a few words: ***dix*** ‖ → dis‖ (but *deux* ‖ → dø‖)

⑥ *Coda* $\mathscr{L}$ (1). Can get **v$\mathscr{L}$.V** instead of v.$\mathscr{L}$V (but never *v$\mathscr{L}$.c)

⑦ *h-aspiré onset* $\mathscr{F}$ *(but not* $\mathscr{L}$*).* Can get **v.$\mathscr{F}\mathbb{V}$** (but not *v.$\mathscr{L}\mathbb{V}$)

⑧ *Post-pausal* $\mathscr{L}$. $\mathscr{L}$ can surface after a prosodic break: ‖$\mathscr{L}$

⑨ Frequency effect. Where optional, p($\mathscr{L}$ surfaces) ~ $p(W_2 | W_1)$

… as if $\mathscr{L}$ were part of the following word

Côté 2005, 2011
Tranel 1981 *et seq.*
⑥ Encrevé 1988
⑨ Ågren 1973, Bybee 2001

… so neither W_1 nor W_2 alone contains all lexically-specific relevant information

Ⓐ The phonological phenomenon: Complications

Errors that are *expected* under the **$\mathcal{L}\hat{W}_2$-** but *not* the **$\hat{W}_1\mathcal{L}$**-Analysis:

⑩ *Incorrect $\mathcal{L}$ selection.* When an incorrect C is substituted for $\mathcal{L}$, it is another liaison C: **v.$\mathcal{L}$′v** for v.$\mathcal{L}$v

⑪ *Exceptional $\mathcal{L}$ epenthesis.* When what should be V.V is illicitly repaired by C-insertion, it is a liaison C: **v.$\mathcal{L}$′v** for v.v

⑫ *Child $\mathcal{L}$-as-$\mathcal{F}$.* $\mathcal{L}\hat{W}_2$ treated as if word $\mathcal{F}\hat{W}_2$ — *joli '**n**ami'*

$\mathcal{L}\hat{W}_2$ Analysis: mis-selection of W_2 allomorph: **$\mathcal{L}'\hat{W}_2$ for $\mathcal{L}\hat{W}_2$**

ditto: /nami/ for /ami/

expected given **[$_{Wd}$ = [$_\sigma$** heuristic for word segmentation

$\mathcal{L}\hat{W}_2$ Analysis: mis-selection of W_2 allomorph: **$\mathcal{L}'\hat{W}_2$ for $\hat{W}_2$**

R. Shi 2011: ~20 mos., *ami, tami, zami, nami*

Ⓐ The phonological phenomenon: Complications

Challenges for the $\mathcal{L}\hat{W}_2$**-** but not the $\hat{W}_1\mathcal{L}$-Analysis:

⑬ W_2 *allomorph selection.* (None required in $\hat{W}_1\mathcal{L}$-Analysis)

⑭ *Coda* $\mathcal{L}$ (2). Can get **v$\mathcal{L}$.V** instead of v.$\mathcal{L}$V — but never *v$\mathcal{F}$.V

Another challenge for both analyses:

⑮ *Gender-bending* $\mathcal{L}$. *belle copine* and *belle amie;*
beau copain but *<u>beau a</u>mi*: instead ***bel ami***.

Proposed GSC theory appears to account for all ⓝs
(explanation? insight?)

②

The adult blend

Ⓐ The phonological phenomenon

Ⓑ GSC analysis: Idea

Ⓒ GSC analysis: Formal account

A gradient grammar of French liaison

Ⓑ A GSC analysis: Idea

Latent consonants in French (liaison)

So far, following orthography, we've assumed a liaison C is *final* in the word it follows:

- the **Ŵ$_1\mathcal{L}$ Analysis**
  - ➤ also take to include *syllabification-driven alternation*

But in children's ear After Dowty: propose that the adult state ...

ami, tami, zar

blends Ŵ$_1\mathcal{L}$-

— multiple al

… and $\mathcal{L}$Ŵ$_2$-Analyses

Presumably extracted from *joli.t ami, peti.t ami, le.s amis, u.n ami,*

… via a bias [$_{\text{morpheme}}$ = [$_{\text{syllable}}$

That is, a liaison C is ***initial*** in the word that it precedes:

- the **$\mathcal{L}$Ŵ$_2$ Analysis**

Ⓑ A GSC analysis: Idea

Underlying forms in $W_1 + W_2$

$/W_1/$

$= \hat{W}_1(\lambda\cdot\mathcal{L})$ *petit:* /pøti(λ·t)/

$= \hat{W}_1(1\cdot\mathcal{F})$ *juste:* /ʒys(1·t)/

blends $\hat{\mathbf{W}}_1\mathcal{L}$- [activity λ] ...

... and $\mathcal{L}\hat{W}_2$-Analyses [activity (τ, ζ, ν)]

$/W_2/$

$= C\hat{W}_2$ *copain:* /kopɛ̃/

$= \mathbb{V}\hat{W}_2$ *ℍéro:* /eʁo/ (*h-aspiré*)

$= \mathfrak{L}\hat{W}_2$ *ami:* /𝔏ami/ where $\mathfrak{L} \equiv (\tau\cdot t + \zeta\cdot z + \nu\cdot n)$

Ⓑ A GSC analysis: Idea

Underlying forms in $W_1 + W_2$

[for now $(\lambda, \tau, \zeta, \nu)$ are constants across the entire lexicon]

$\doteq (0.5, 0.3, 0.3, 0.3)$

$/W_1/$

$= \hat{W}_1(\lambda \cdot \mathcal{L})$ *petit:* /pøti(λ·t)/

$= \hat{W}_1(1 \cdot \mathcal{F})$ *juste:* /ʒys(1·t)/

petit ami:

/pøti(λ·t) (τ·t + ζ·z + ν·n)ami/

→ pø.ti.<u>**t**</u>a.mi

(τ in /W₂/ = $\mathfrak{L}\hat{W}_2$ gives /t/ the extra activity needed to bring λ up to the threshold level required to surface)

$/W_2/$

$= C\hat{W}_2$ *copain:* /kopɛ̃/

$= \mathbb{V}\hat{W}_2$ *ℍéro:* /eʁo/ (*h-aspiré*)

$= \mathfrak{L}\hat{W}_2$ *ami:* /𝔏ami/ where $\mathfrak{L} \equiv (\tau \cdot t + \zeta \cdot z + \nu \cdot n)$

Ⓑ A GSC analysis: Idea

Underlying forms in $W_1 + W_2$

[for now $(\lambda, \tau, \zeta, \nu)$ are constants across the entire lexicon]

$\doteq (0.5, 0.3, 0.3, 0.3)$

$/W_1/$

$= \hat{W}_1(\lambda\cdot\mathcal{L})$ *petit:* /pøti(λ·t)/

$= \hat{W}_1(1\cdot\mathcal{F})$ *juste:* /ʒys(1·t)/

petit copain:

/pøti(λ·t) kopɛ̃/

→ .pø.ti.ko.pɛ̃.

($/W_2/ = \mathrm{C}\hat{W}_2$ *lacks* the extra activity for /t/ needed to bring λ up to the threshold level required to surface)

$/W_2/$

$= \mathrm{C}\hat{W}_2$ *copain:* /kopɛ̃/

$= \mathbb{V}\hat{W}_2$ *ℍéro:* /eʁo/ (*h-aspiré*)

$= \mathfrak{L}\hat{W}_2$ *ami:* /𝔏ami/ where $\mathfrak{L} \equiv (\tau\cdot t + \zeta\cdot z + \nu\cdot n)$

Ⓑ A GSC analysis: Idea

Underlying forms in $W_1 + W_2$

[for now $(\lambda, \tau, \zeta, \nu)$ are constants across the entire lexicon]

$\doteq (0.5, 0.3, 0.3, 0.3)$

$/W_1/$

$= \hat{W}_1(\lambda \cdot \mathcal{L})$ *petit:* /pøti(λ·t)/

$= \hat{W}_1(1 \cdot \mathcal{F})$ *juste:* /ʒys(1·t)/

petit héro:

/pøti(λ·t̲) eʁo/

→ .pø.ti.e.ʁo.

(/W_2/ = $\mathbb{V}\hat{W}_2$ *lacks* the extra activity for /t/ needed to bring λ up to the threshold level required to surface)

$/W_2/$

$= \mathrm{C}\hat{W}_2$ *copain:* /kopɛ̃/

$= \mathbb{V}\hat{W}_2$ *ℍéro:* /eʁo/ (*h-aspiré*)

$= \mathfrak{L}\hat{W}_2$ *ami:* /𝔏ami/ where $\mathfrak{L} \equiv (\tau \cdot t + \zeta \cdot z + \nu \cdot n)$

②

The adult blend

Ⓐ The phonological phenomenon

Ⓑ GSC analysis: Idea

Ⓒ GSC analysis: Formal account

A gradient grammar of French liaison

Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v   *peti*(*t*) + *ami* → .pø.ti.ta.mi.

Environment: vCV; output: v.CV or v.V   [V ≡ $\mathcal{L}$v]

peti(*t*) *ami*   [^{M}pøti($\lambda\cdot t_1$)] [^{M}($\tau\cdot t_2+\zeta\cdot z_3+\nu\cdot n_4$)ami]

0.5   0.3   0.3   0.3

Harmonic Grammar (Legendre, Miyata & Smolensky 1990, Pater 2009 *et seq.*)

–10

[^{M}pøti($\lambda\cdot t_1$)] [^{M}($\tau\cdot t_2+\zeta\cdot z_3+\nu\cdot n_4$)ami]		DEP	UNIF
a	.pø.ti.a.mi.		
b	.pø.ti.t_{12}a.mi.	1–(λ+τ) 0.2	*
c	.pø.ti.t_1a.mi.	1–λ 0.5	

Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v   *peti*(*t*) + *ami* → .pø.ti.t̲a.mi.

Environment: vC̲V; output: v.CV or v.V   [V ≡ $\mathcal{L}$v]

peti(*t*) *ami*   [^{M}pøti($\lambda \cdot t_1$)] [^{M}($\tau \cdot t_2 + \zeta \cdot z_3 + \nu \cdot n_4$)ami]

0.5   0.3   0.3   0.3

Numbers are not derived *a priori*; they are fit to the data

	[^{M}pøti($\lambda \cdot t_1$)] [^{M}($\tau \cdot t_2 + \zeta \cdot z_3 + \nu \cdot n_4$)ami]	*–10* DEP	UNIF
a	.pø.ti.a.mi.		
b	.pø.ti.t_{12}a.mi.	1–(λ+τ) 0.2	*
c	.pø.ti.t_1a.mi.	1–λ 0.5	

Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v   *peti*(*t*) + *ami* → .pø.ti.ta.mi.

Environment: vCV; output: v.CV or v.V   [V ≡ $\mathfrak{L}$v]

peti(*t*) *ami*   [^{M}pøti($\lambda\cdot t_1$)] [^{M}($\tau\cdot t_2+\zeta\cdot z_3+\nu\cdot n_4$)ami]

0.5   0.3   0.3   0.3

All gradient versions of standard constraints from OT phonology

[^{M}pøti($\lambda\cdot t_1$)] [^{M}($\tau\cdot t_2+\zeta\cdot z_3+\nu\cdot n_4$)ami]		–10 DEP	UNIF
a	.pø.ti.a.mi.		
b	.pø.ti.t_{12}a.mi.	1–($\lambda+\tau$) 0.2	*
c	.pø.ti.t_1a.mi.	1–λ 0.5	

Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v   *peti*(*t*) + *ami* → .pø.ti.ta.mi.

Environment: vCV; output: v.CV or v.V   [V ≡ $\mathcal{L}$v]

peti(*t*) *ami*   [^{M}pøti($\lambda{\cdot}t_1$)] [^{M}($\tau{\cdot}t_2+\zeta{\cdot}z_3+\nu{\cdot}n_4$)ami]

0.5 0.3 0.3 0.3

ALIGN([m, [σ) [positive]

[^{M}pøti($\lambda{\cdot}t_1$)] [^{M}($\tau{\cdot}t_2+\zeta{\cdot}z_3+\nu{\cdot}n_4$)ami]	DEP *–10*	MAX *2*	ALIGN *1*	UNIF
a .pø.ti.a.mi.				
b .pø.ti.t_{12}a.mi.	1–(λ+τ) 0.2	λ+τ 0.8	1	*
c .pø.ti.t_1a.mi.	1–λ 0.5	λ 0.5		

Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v   *peti*(*t*) + *ami* → .pø.ti.ta.mi.

Environment: vCV; output: v.CV or v.V   [V ≡ $\mathcal{L}$v]

peti(*t*) *ami*   [^{M}pøti($\lambda \cdot t_1$)] [^{M}($\tau \cdot t_2 + \zeta \cdot z_3 + \nu \cdot n_4$)ami]

0.5 0.3 0.3 0.3

W₂ *allomorph selection.* It's automatic: only the matching $\mathcal{L}$ can coalesce ⇒ surface; next case shows coalescence is necessary

[^{M}pøti($\lambda \cdot t_1$)] [^{M}($\tau \cdot t_2 + \zeta \cdot z_3 + \nu \cdot n_4$)ami]	DEP *–10*	MAX *2*	ALIGN *1*	ONSET *–0.9*	UNIF *–0.7*	*H*
a .pø.ti.a.mi.				1		**–0.9**
b .pø.ti.t_{12}a.mi. ☞	1–(λ+τ) 0.2	λ+τ 0.8	1		‡	**–0.1**
c .pø.ti.t_1a.mi.	1–λ 0.5	λ 0.5				**–4**

© A GSC analysis: Formal account

$$H(b) - H(a) = [(1 - \lambda - \tau)D + (\lambda + \tau)M + U + A^{L}] - [O]$$
$$= (\lambda + \tau)[M - D] + D + U + A^{L} - O$$
$$> 0$$
$$\text{iff } (\lambda + \tau) > -[D + U + A^{L} - O]/[M - D] \equiv \theta(\mathrm{v}\underline{\mathrm{C}}\mathrm{V})$$
$$\doteq -[-10 - 0.7 + 1 - (-0.9)]/[2 - (-10)]$$
$$= 0.73$$

✓ since $\lambda + \tau \doteq 0.5 + 0.3 = 0.8$

same procedure for all elements & environments gives corresponding θ: activity of gradient segments must > θ to surface

			–10	*2*	*1*	*–0.9*	*–0.7*	
[ᴹpøti(λ·t_1)] [ᴹ(τ·t_2+ζ·z_3+ν·n_4)ami]			DEP	MAX	ALIGN	ONSET	UNIF	***H***
a	.pø.ti.a.mi.					1		**–0.9**
b	.pø.ti.t_{12}a.mi.	☞	1–(λ+τ) 0.2	λ+τ 0.8	1		1	**–0.1**
c	.pø.ti.t_1a.mi.		1–λ 0.5	λ 0.5				**–4**

Ⓒ A GSC analysis: Formal account

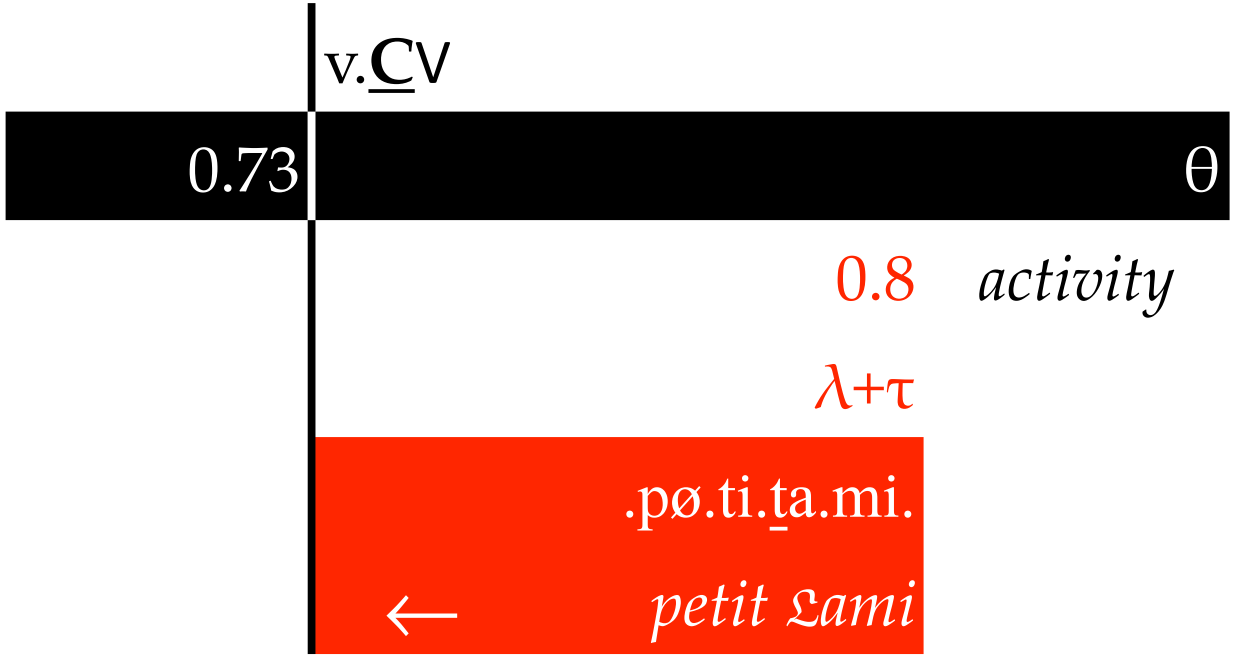
Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v *peti*(*t*) + *ami* → .pø.ti.t̲a.mi.

Environment: v<u>**C**</u>V; output: v.**C**V or v.V

pøtit ami [$^{\mathrm{M}}$pøti($\lambda \cdot t_1$)] [$^{\mathrm{M}}$($\tau \cdot t_2 + \zeta \cdot z_3 + \nu \cdot n_4$)ami]

$\theta(\mathrm{v\underline{C}V}) = -[D + U + A^{\mathrm{L}} - O]/[M - D] \doteq 0.73$



Ⓒ A GSC analysis: Formal account

Consider *joli ami* /joli ($\tau\cdot$t + $\zeta\cdot$z + $\nu\cdot$n)ami/

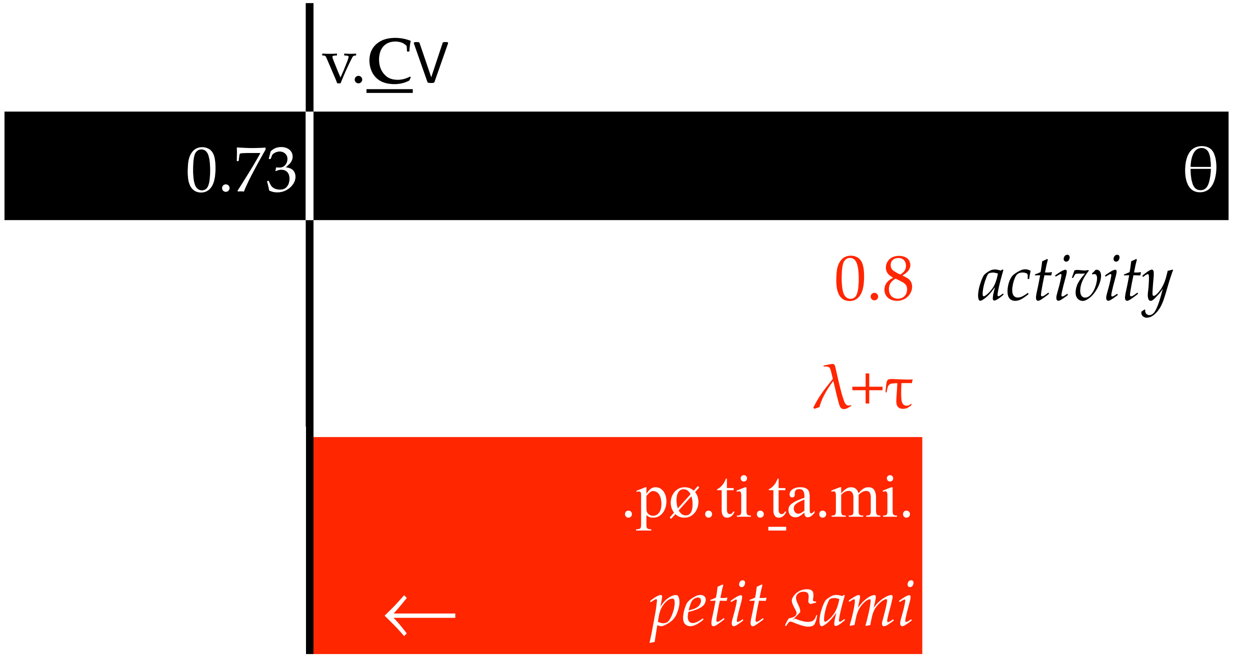
The pre-W_2 $\mathfrak{L}$ consonants are in the same environment as the post-W_1 consonant $\mathcal{L} = \lambda\cdot$t for *peti*(*t*) *ami*.

But now the only activity for any liaison C is

$\tau \doteq 0.3 \doteq \zeta = \nu$:

$< \theta(\text{v}\underline{\text{C}}\text{V}) \doteq 0.73$

$\Rightarrow$ no C surfaces ✓



Ⓒ A GSC analysis: Formal account

Core phenomena:

② v$\mathcal{L}$ + c → v.c   *peti*(*t*) + *copain* → .pø.ti.ko.pɛ̃.

Environment: v<u>**C**</u>c

petit copain [ᴹpøti(λ·t_1)] [ᴹkopɛ̃]

skip [3]

When does $\mathcal{L}$ = /t/ surface? I.e., when is $b > a$?

$H(b) - H(a) = [(1-\lambda)D + \lambda M + N + A^L + A^R] - [A^L]$

$= \lambda[M - D] + D + N + A^R > 0$

iff $\lambda > -[D + N + A^R]/[M - D] \equiv \theta(\text{v}\underline{\text{C}}\text{c})$

$\doteq -[-(10) - 0.2 + 0.1]/[2 - (-10)] = 0.84$

✗ ⇒ $\mathcal{L}$ does *not* surface
$\mathcal{F}$ *does* surface ✓

		–10 DEP	2 MAX	–0.2 NOCODA	1 ALIGN-L	0.1 ALIGN-R	*H*
a	.pø.ti.ko.pɛ̃.				1		**1**
b	.pø.tit$_1$.ko.pɛ̃.	1–λ 0.5	λ 0.5	1	1	1	**– 3.2**

Ⓒ A GSC analysis: Formal account

Core phenomena:

② v$\mathcal{L}$ + c → v.c   *peti*(*t*) + *copain* → .pø.ti.ko.pɛ̃.

Environment: v<u>C</u>c

petit copain [ᴹpøti(λ·t_1)] [ᴹkopɛ̃]

$$\theta(\mathrm{v\underline{C}c}) = -[D + N + A^{R}]/[M - D] \doteq \mathbf{0.84}$$

v<u>C</u>.c

0.84 | θ

0.5 | 1 | *activity*

λ | $\mathcal{F}$

.pø.ti.ko.pɛ̃. | .pø.tit̲.ko.pin.

petit̲ copain → | ← *petite copine*

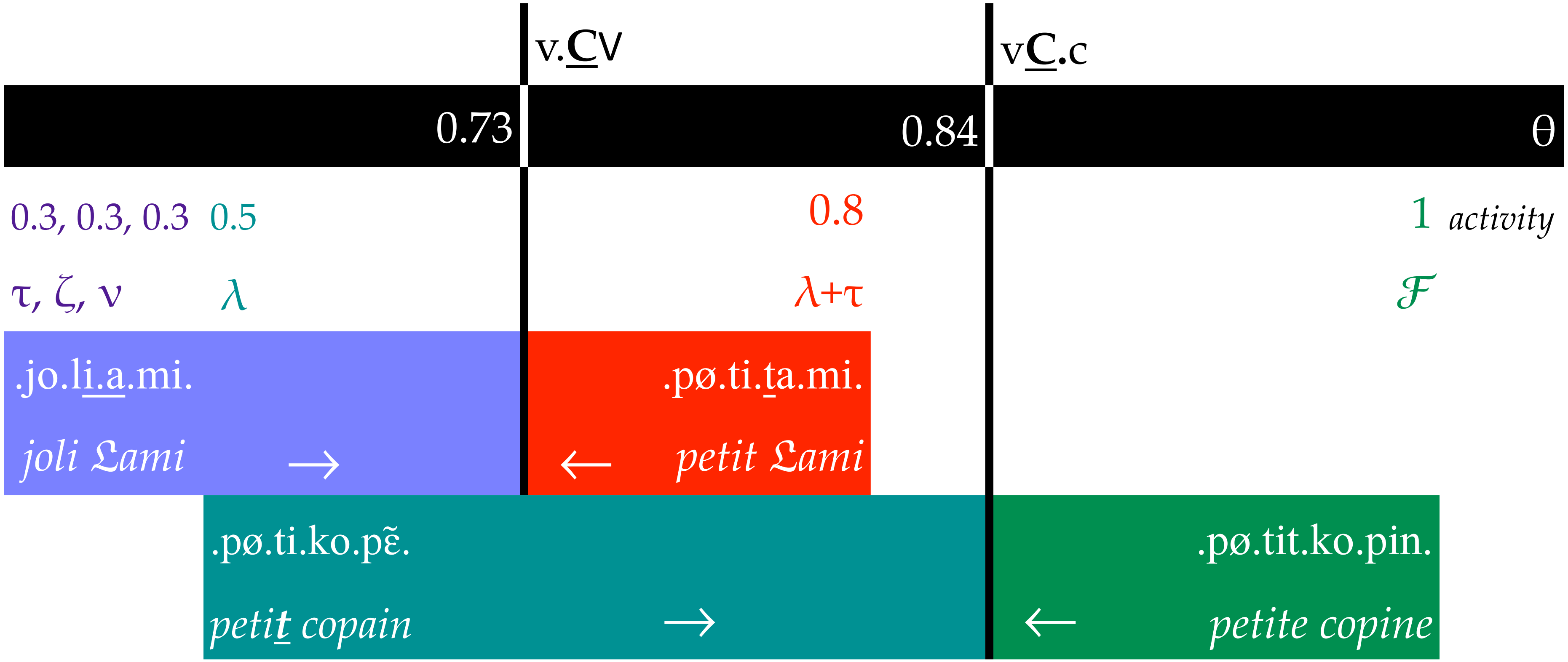
Ⓒ A GSC analysis: Formal account

Core phenomena:

① v$\mathcal{L}$ + v → v.$\mathcal{L}$v *peti*(*t*) + *ami* → .pø.ti.t̲a.mi.

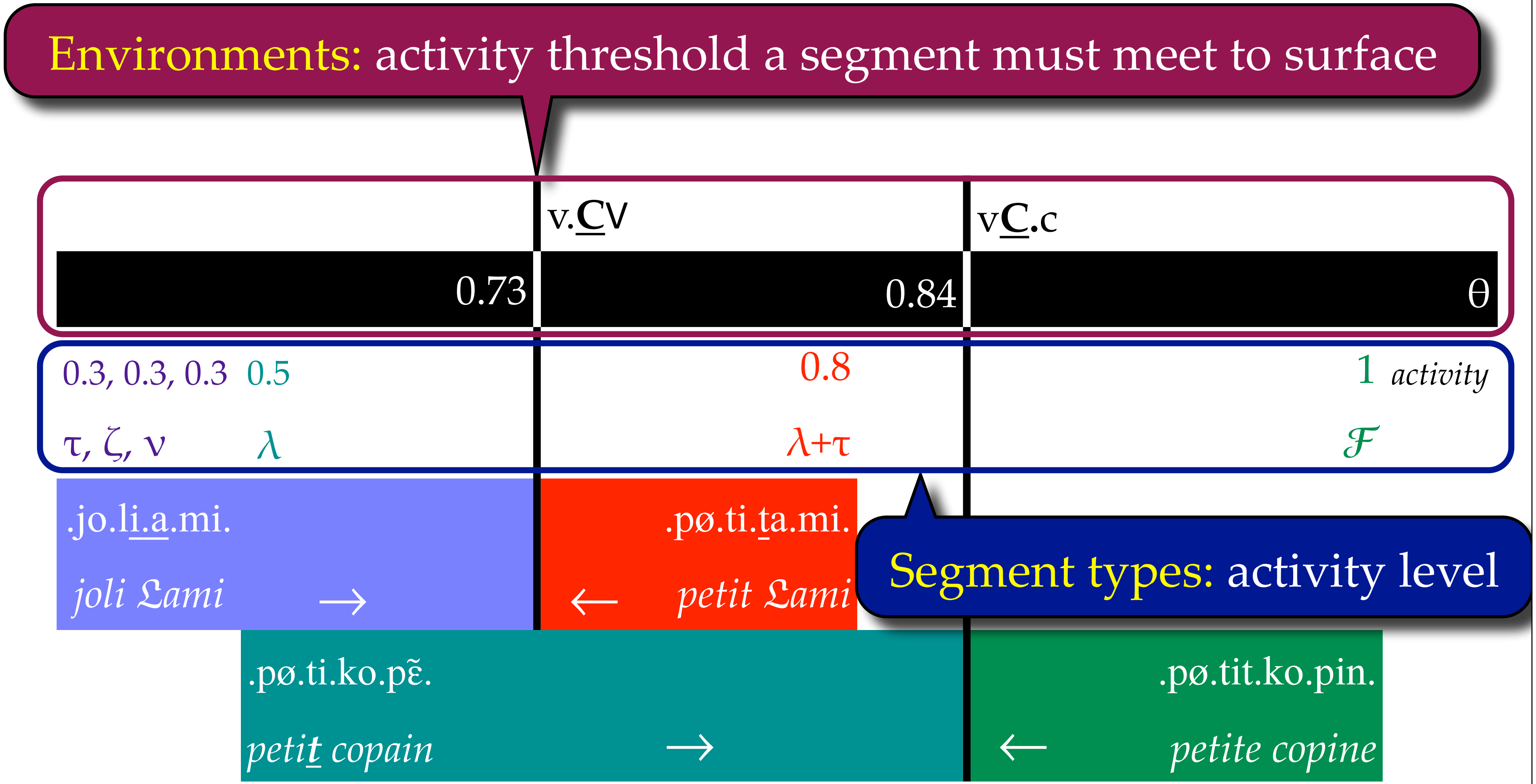
② v$\mathcal{L}$ + c → v.c *peti*(*t*) + *copain* → .pø.ti.ko.pɛ̃.

Environments: v**C̲**V, v**C̲**c



Ⓒ A GSC analysis: Formal account

The analysis consists of 2 crossed dimensions:



Ⓒ A GSC analysis: Formal account

Core mappings

①	$\mathrm{v}\mathcal{L} + \mathrm{V}$	$\rightarrow$	$\mathrm{v}.\mathcal{L}\mathrm{V}$	*peti*(*t*) + *ami*	$\rightarrow$	.pø.ti.t̠a.mi.
②	$\mathrm{v}\mathcal{L} + \mathrm{c}$	$\rightarrow$	$\mathrm{v.c}$	*peti*(*t*) + *copain*	$\rightarrow$	.pø.ti.ko.pɛ̃.
③	$\mathrm{v}\mathcal{L} + \mathbb{V}$	$\rightarrow$	$\mathrm{v}.\mathbb{V}$	*peti*(*t*) + ℍ*éro*	$\rightarrow$	.pø.ti.e.ʁo.
④	$\mathrm{v}\mathcal{F} + \mathrm{c}$	$\rightarrow$	$\mathrm{v}\mathcal{F}.\mathrm{c}$	*petite* + *copine*	$\rightarrow$	.pø.tit̠.ko.pin.

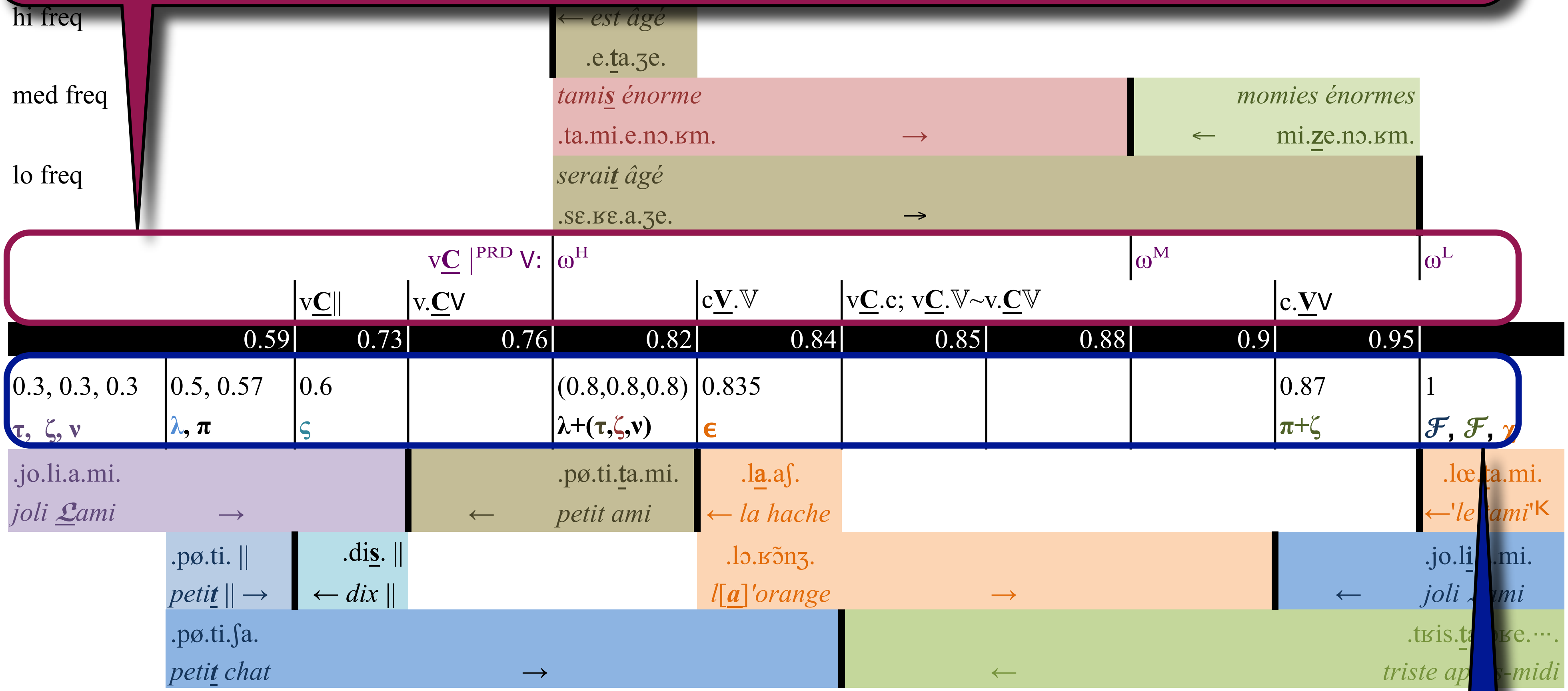
Analysis handles these 4 core patterns and nearly a dozen peripheral patterns: so far, handles all phenomena covered by both the $\hat{\mathrm{W}}_1\mathcal{L}$ and $\mathcal{L}\hat{\mathrm{W}}_2$ accounts

© An analysis of the GSC analysis

A less incomplete diagram of the analysis:

skip

Environments: activity threshold a segment must meet to surface



Segment types: activity level

Ⓒ An analysis of the GSC analysis

Restrictiveness

Environments: activity threshold a segment must meet to surface

- no matter the underlying activity of a segment x, if x surfaces in an environment with a threshold θ, then x must surface in any environment with a threshold $< \theta$

			vC $|^{PRD}$ V:	ω^H				ω^M		ω^L
		vC‖	v.CV		cV.𝕍	vC.c; vC.𝕍~v.C𝕍			c.VV	
	0.59	0.73	0.76	0.82	0.84	0.85	0.88	0.9	0.95	
0.3, 0.3, 0.3	0.5, 0.57	0.6		(0.8,0.8,0.8)	0.835				0.87	1
τ, ζ, ν	λ, π	ς		λ+(τ,ζ,ν)	ϵ				π+ζ	ℱ, ℱ, γ

- no matter the threshold of an environment E, if a segment x with activation $a \leq 1$ surfaces in E, then a segment x with any activation $> a$ (and ≤ 1) must also surface in E

Segment types: activity level

③

Acquisition

[1]

Speculations on formalizing Dowty's sketch

skip

❸ Notes: Acquisition

Comprehension-directed optimization &

- ALIGN-L(Morpheme, Syllable)

→ start in free variation *ami* ~ *tami* ~ *zami* ~ *nami*

  ➤ from: *joli. ami, peti.t ami, le.s amis, u.n ami*

Error signal *ʒoli t̲ami/ʒoli ami →

- weakens initial *t* of *tami*, say by 0.1;
  eventually, reduces to say $(0.7 \cdot t)$*ami*; [assume $\theta = 0.73$ as above];then
- to get *peti.tami* (when correctly choose /*tami*/)
  ➤ need "more *t* activity"
  ➤ increase activity of *t* on both sides, say by 0.05: *peti*$(0.05 \cdot t)$ $(0.75 \cdot t)$*ami*
- error *ʒoli t̲ami returns; reduce to (0.65 . *t*)*ami*
  ➤ to get *petit.ami* need to increase again: *peti*$(0.1 \cdot t)$ $(0.70 \cdot t)$*ami*
  ➤ ...

☞ gradual shift of *t* activity from *t̲ami* to *petit̲*

Adult blend analysis ⇒ the shift does not go all the way!

④

Prosody

[6]

Tentative suggestions

skip

❹ The role of prosody: Formalization

'$[W_1W_2]$' lexical entry (input to grammar):

$[_{m}\ \mathbb{W}_1\ (-\phi\cdot {}_{m}][_{m})\ \mathbb{W}_2\ {}_{m}]$

- ➤ $\mathbb{W}_1$ means this contributes only to inputs with a particular W_1; $\mathbb{W}_2$ means this contributes only to inputs with a particular W_2 *or* to inputs in which W_2 belongs to a particular syntactic category $\mathbb{X}$
  - ✦ e.g., $[_{m}\ \mathbb{quand}\ (-0.7\cdot {}_{m}][_{m})\ \mathbb{N}\ {}_{m}]$ 'when N'

Call this a *collocation schema*

Input for *quand on (va)* is the blend:

$$[_{m}\ \text{quand}\ {}_{m}]\ [_{m}\ \text{on}\ {}_{m}]\ +\ [_{m}\ \text{quand}\ (-0.7\cdot {}_{m}][_{m})\ \text{on}\ {}_{m}]$$

$$=[_{m}\ \text{quand}\ (0.3\cdot {}_{m}][_{m})\ \text{on}\ {}_{m}]$$

i.e. *quand* and *on* are separated by a morpheme boundary of activity 0.3 $\rightarrow$ *quand* [t] *on (va)*

➃ The role of prosody: Formalization

The outputs from the grammar (candidates):

- contain morphological structure = that of the input (*containment*)
- are evaluated by constraints:

  *CROSS(Morph, PCat): [Morph] and (PCat) constituents cannot cross

  I.e., can have neither

  [Morph (PCat $\mu \cdot$ Morph] PCat) nor

  (PCat $\mu \cdot$ [Morph PCat) Morph]

  Penalty: $\mu \cdot w_{\text{*CROSS(Morph, PCat)}}$

  which form a universal markedness hierarchy:

  if PCat′ is higher in the prosodic hierarchy than PCat, then

  $w_{\text{*CROSS(Morph, PCat')}} > w_{\text{*CROSS(Morph, PCat)}}$

Crucially: liaison violates *CROSS from coalescence:

(PCat [m1 peti PCat) (PCat [m2 t12 m1] ami m2] PCat) *peti.t ami*

❹ The role of prosody: Formalization

Penalty from liaison: $\mu \cdot w_{*\text{CROSS}(\text{Morph, PCat})}$

probability $\propto e^{-\text{Penalty}}$ greater Penalty $\Rightarrow$ lower probability

p(liaison) increases both from

- increasing collocation frequency (decreases μ) and
- decreasing prosodic-hierarchy-level of the boundary separating W_1 and W_2,

because if PCat is lower in the hierarchy than PCat′:

$$w_{*\text{CROSS}(\text{Morph, PCat})} < w_{*\text{CROSS}(\text{Morph, PCat}')}$$

⑤

Summary

Summary

Gradient Symbolic Representations crucial uses:

- adult blend: $0.5 \cdot [\hat{\mathbf{W}}_1\mathcal{L}$-analysis$]\ +\ 0.3 \cdot [\mathcal{L}\hat{\mathbf{W}}_2$ -analysis$]$

formalization of Dowty (2003)

Summary

Gradient Symbolic Representations crucial uses:

- adult blend: $0.5\cdot[\hat{\mathbf{W}}_{\mathbf{1}}\mathcal{L}$-analysis] + $0.3\cdot[\mathcal{L}\hat{\mathbf{W}}_{\mathbf{2}}$ -analysis]
- many crucially different gradient activity levels for different $\mathcal{L}$s
  - ➤ $\mathcal{L}$ of W_1
  - ➤ $\mathcal{L}$ of W_2
  - ➤ *z* of PLURAL
  - ➤ *z* of *dix*
  - ➤ pure floating activity of FEM
  - ➤ Vs that elide

crucial dependence on Harmonic Grammar to enable grammatical computation over Gradient Symbolic Representations

- acquisition process of gradually shifting activity of $\mathcal{L}$ from W_2 to W_1
- usage-based gradual increase of activity in lexicon of $[W_1W_2]$
  - ➤ implemented with *negative morpheme boundary activity*

That's all folks! — Thanks for your attention