UFPB

PROLING

LAPROL
Laboratório de Processamento Linguístico

TLB
Teorias Linguísticas de Base

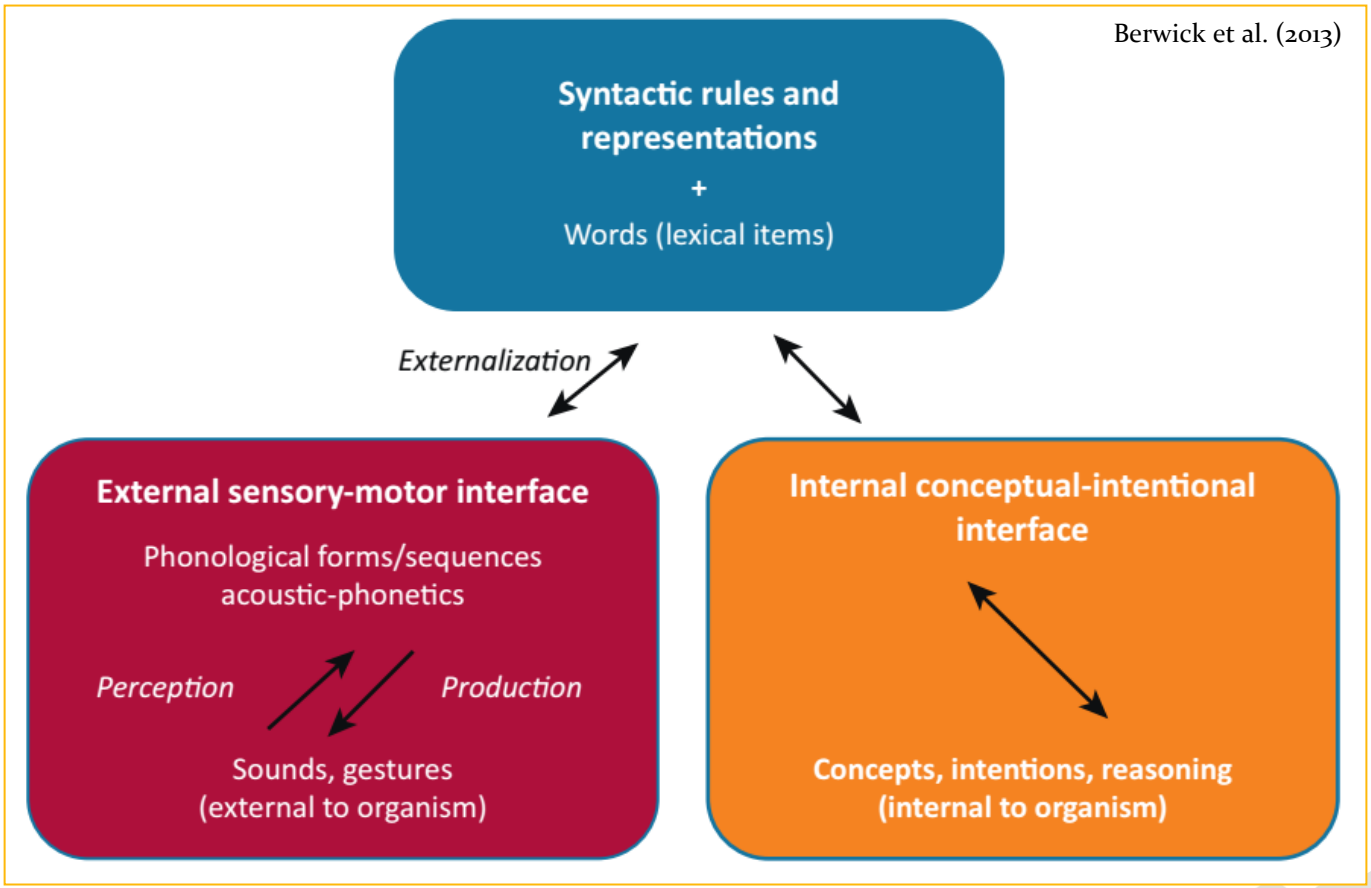
Processamento cognitivo e interface morfossinática

Sábado, 27 de julho de 2019

Gustavo L. Estivalet

Avant-Propos

How do we recognize words?
How are words stored in the lexicon?
How morphology is processed?



The Mental Lexicon

Word List

falo
falas
falamos
...

Morphemes

fal mos

Σ

a

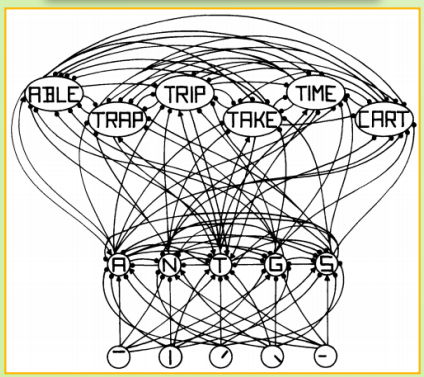
Empiricism

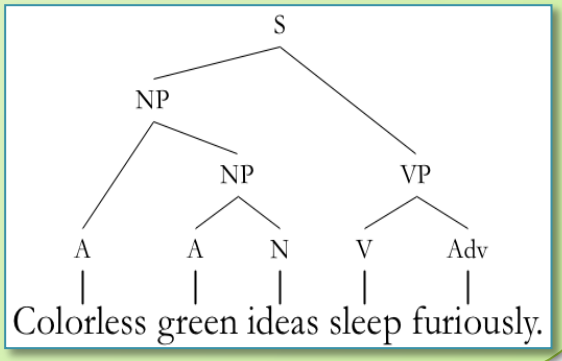
Vs.

Rationalism

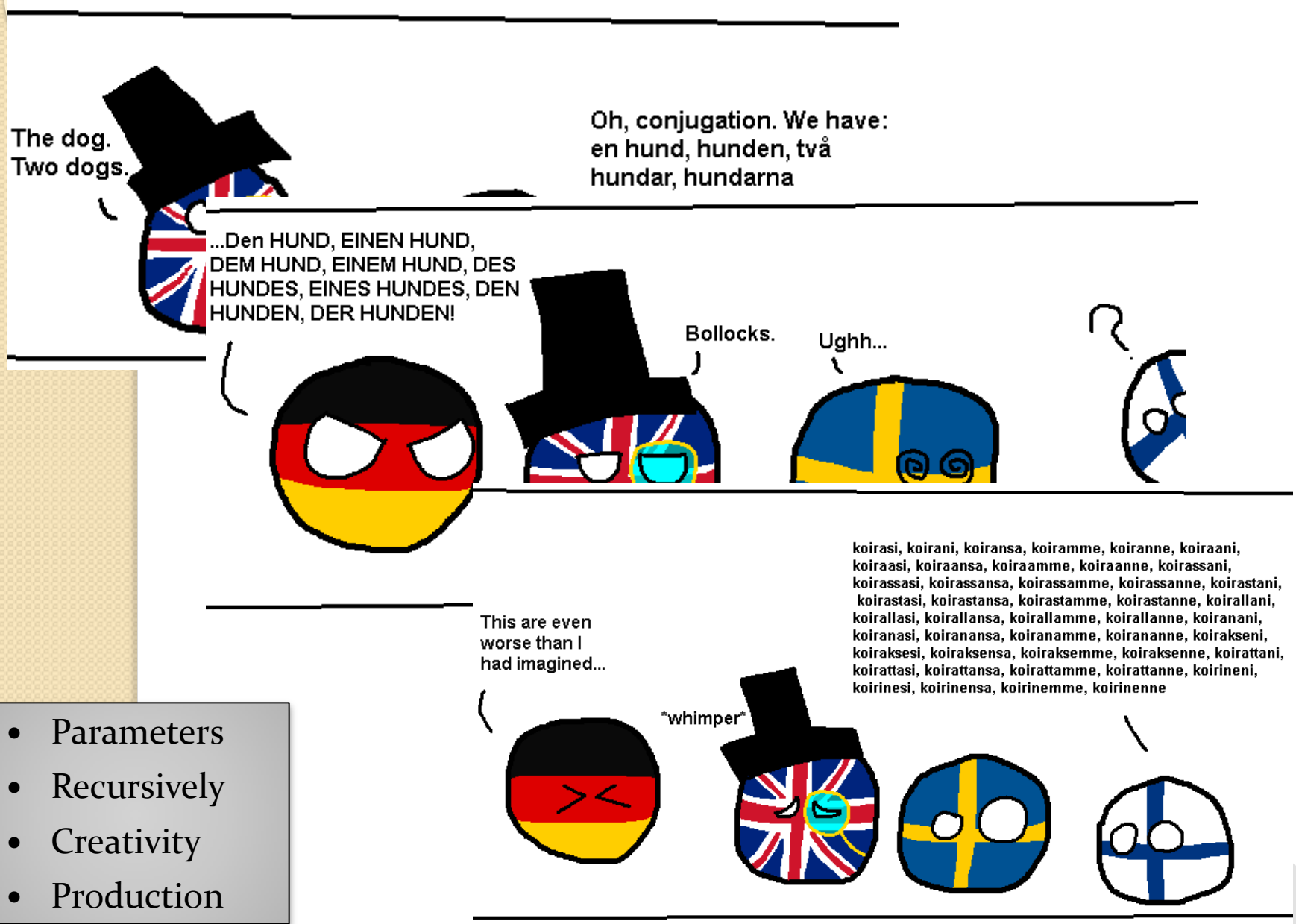
Associative
Interactive

Symbolic
Combination





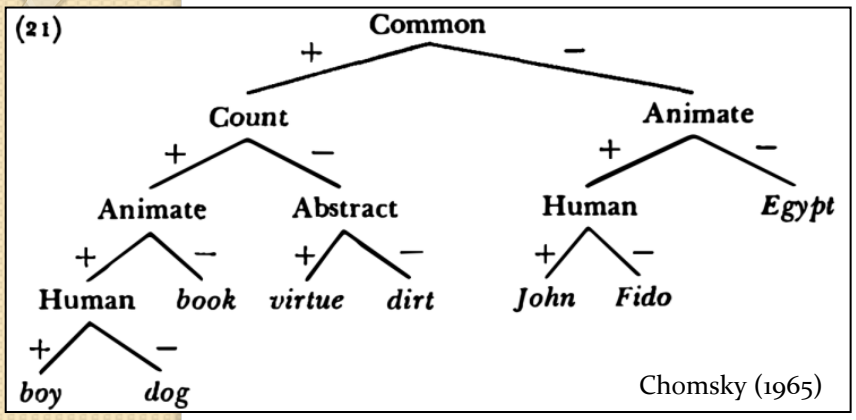
Morphology!



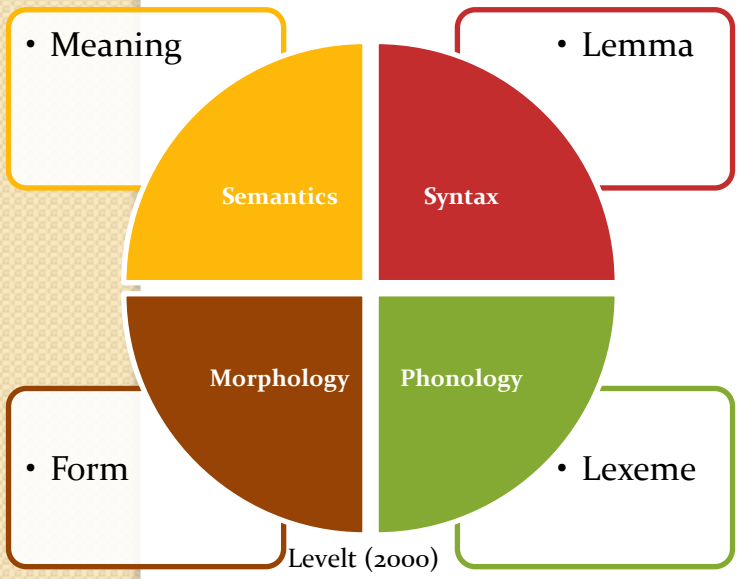
- Parameters
- Recursively
- Creativity
- Production

koirakaan, koirankaan, koiraakaan, koirassakaan, koirastakaan, koiraankaan, koirallakaan, koiraltakaan, koirallekaan,
koiranakaan, koiraksikaan, koirattakaan, koirineenkaan, koirinkaan, koirako, koiranko, koiraako, koirassako, koirastako,
koiraanko, koirallako, koiraltako, koiralleko, koiranako, koiraksiko, koirattako, koirineenko, koirinko, koirasikaan,
koiranikaan, koiransakaan, koirammekaan, koirannekaan, koiraanikaan, koiraasikaan, koiraansakaan, koiraammekaan,
koiraannekaan, koirassanikaan, koirassasikaan, koirassansakaan, koirassammekaan, koirassannekaan, koirastanikaan,
koirastasikaan, koirastansakaan, koirastammekaan, koirastannekaan, koirallanikaan, koirallasikaan, koirallansakaan,
koirallammekaan, koirallannekaan, koirananikaan, koiranasikaan, koiranansakaan, koiranammekaan, koirananneekaan,
koiraksenikaan, koiraksesikaan, koiraksensakaan, koiraksemmekaan, koiraksennekaan, koirattanikaan, koirattasikaan,
koirattansakaan, koirattammekaan, koirattannekaan, koirinenikaan, koirinesikaan, koirinensakaan, koirinemmekaan,
koirinennekaan, koirasiko, koiraniko, koiransak[illegible]iranneko, koiraaniko, koiraasiko, koiraansako,
koiraammeko, koiraanneko, koirassaniko, [illegible]irassansako, koirassammeko, koirassanneko,
koirastaniko, koirastasiko, koirastansa[illegible]irastanneko, koirallaniko, koirallasiko,
koirallansako, koirallammeko, koira[illegible]ansako, koiranammeko, koirananneko,
koirakseniko, koiraksesiko, koirak[illegible]rattaniko, koirattasiko, koirattansako,
koirattammeko, koirattanneko, k[illegible]iko, koirinensako, koirinemmeko,
koirinenneko, koirasikaanko, ko[illegible]nsakaanko, koirammekaanko,
koirannekaanko, koiraanikaank[illegible]o, koiraansakaanko, koiraammekaanko,
koiraannekaanko, koirassanika[illegible]asikaanko, koirassansakaanko,
koirassammekaanko, koirassan[illegible]oirastanikaanko, koirastasikaanko,
koirastansakaanko, koirastamm[illegible]astannekaanko, koirallanikaanko,
koirallasikaanko, koirallansaka[illegible]llammekaanko, koirallannekaanko,
koirananikaanko, koiranasikaan[illegible]ansakaanko, koiranammekaanko,
koiranannekaanko, koiraksenika[illegible]sesikaanko, koiraksensakaanko,
koiraksemmekaanko, koiraksenne[illegible]nikaanko, koirattasikaanko, koirattansakaa
koirattammekaanko, koirattanneka[illegible]nikaanko, koirinesikaanko, koirinensakaanl
koirinemmekaanko, koirinennekaan[illegible]okaan, koiranikokaan, koiransakokaan,
koirammekokaan, koirannekokaan, ko[illegible]an, koiraansakokaan, koiraammekokaan,
koiraannekokaan, koirassanikokaan, koiras[illegible]akokaan, koirassammekokaan, koirassannekokaa
koirastanikokaan, koirastasikokaan, koirastansakok[illegible]nekokaan, koirastannekokaan, koirallanikokaan,
koirallasikokaan, koirallansakokaan, koirallammekokaan, koirallannekokaan, koirananikokaan, koiranasikokaan,
koiranansakokaan, koiranammekokaan, koiranannekokaan, koiraksenikokaan, koiraksesikokaan, koiraksensakokaan,
koiraksemmekokaan, koiraksennekokaan, koirattanikokaan, koirattasikokaan, koirattansakokaan, koirattammekokaan,
koirattannekokaan, koirinenikokaan, koirinesikokaan, koirinensakokaan, koirinemmekokaan...

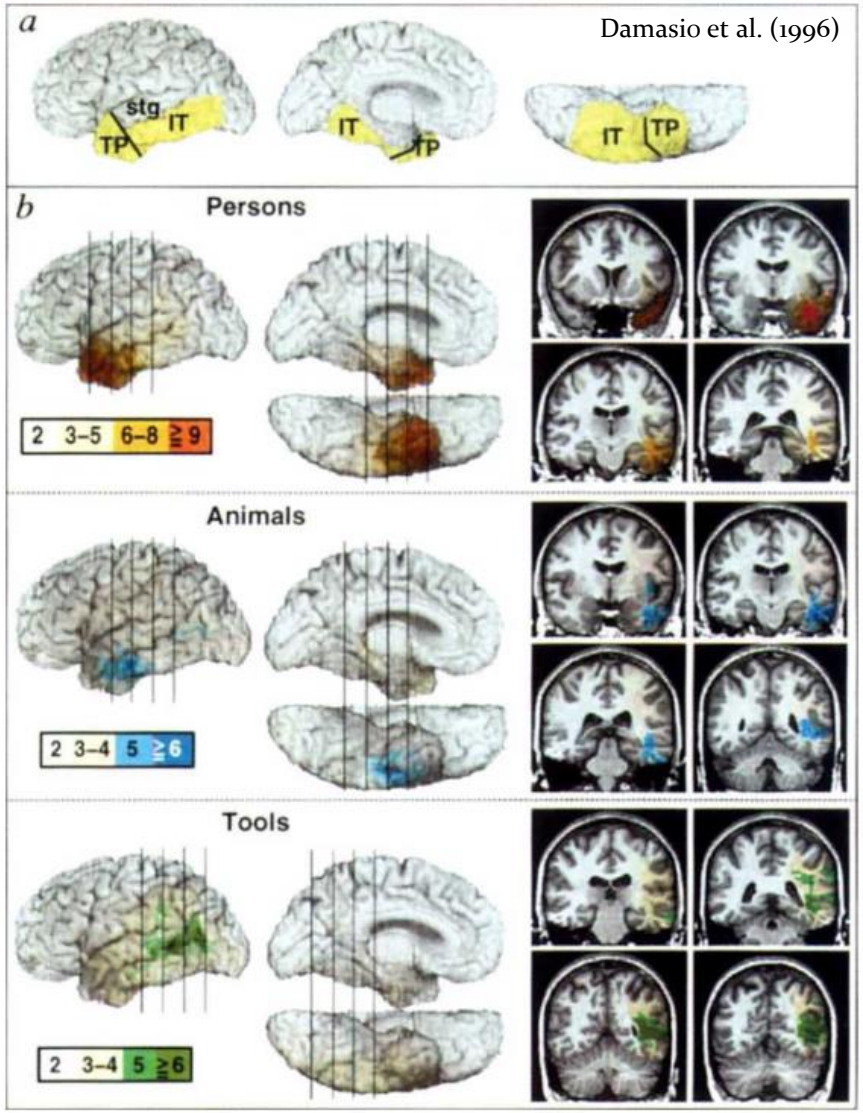
Word Representation



Chomsky (1965)



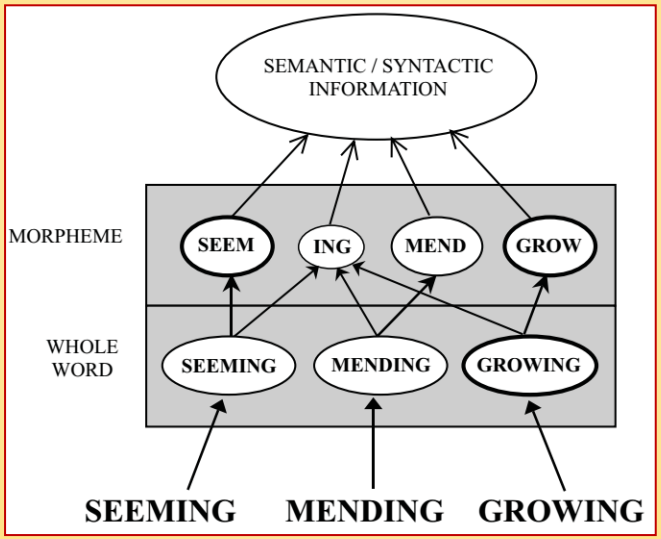
Levelt (2000)



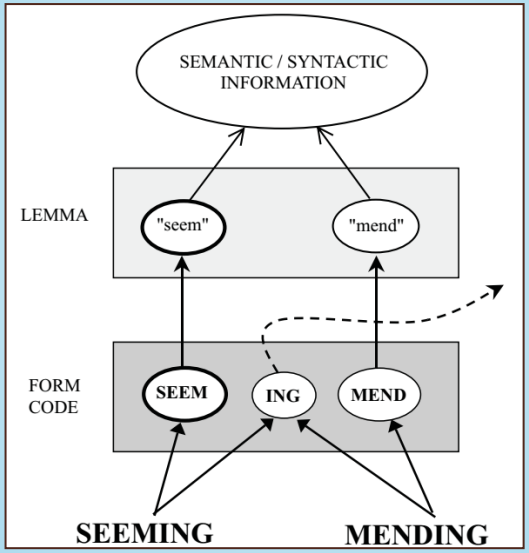
Damasio et al. (1996)

Morphological Processing

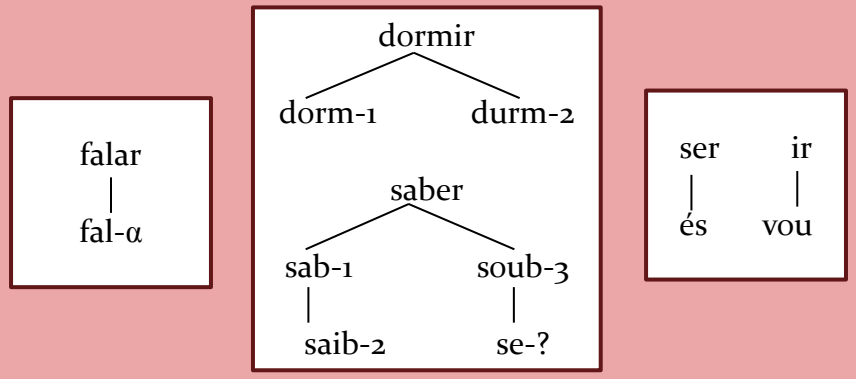
Full-Entry: early whole-word and late morphological activation (Jackendoff, 1975; Grainger & Giraudo, 2001)



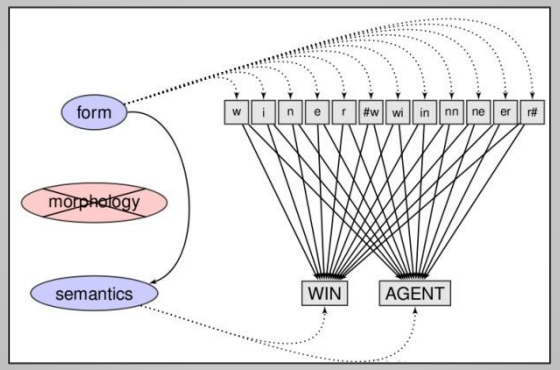
Obligatory Decomposition: morphological and lexical representations (Taft, 1979).



Minimalist Morphology (MM): default rule, sublexical structures, and suppletion (Wunderlich, 1996)

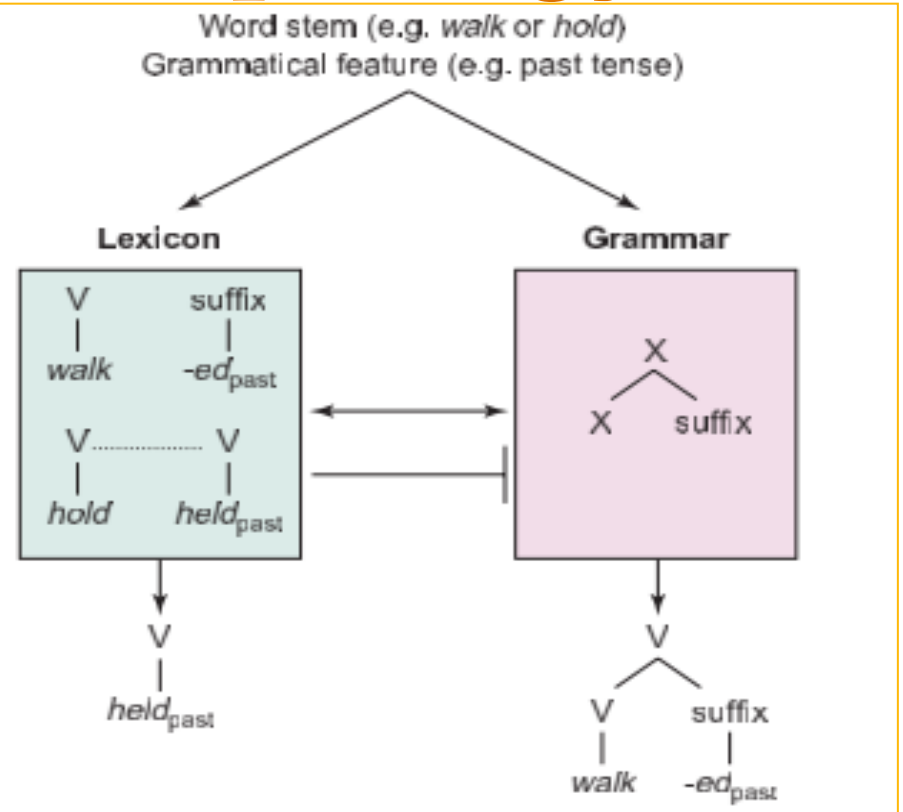


Naïve Discriminative Learning: statistical mapping from form onto meaning (Baayen et al., 2011)



Dual-Route Morphology

Pinker & Ullman (1999)



Used for:	roots, idioms, irregulars, some regulars	phrases, sentences, any regular form
Form of computation:	lookup, association	combination, unification
Subdivision of:	declarative memory	procedural system
Associated with:	words, facts	rules, skills
Principal substrate:	temporo-parietal cortex	frontal cortex, basal ganglia

Behavioral Investigation

- Behavioral data
- Reaction Time, Accuracy
- Lexical Decision
- Masked Priming
- Online/Offline Tasks

- Behavioral Data
- Eye-Tracker Data
- Reading time, Nb fixations, Fixation time, Returns
- Reading profile, Reading deficits

Study 1: Stem Formation

languages MDPI

Article

Stem Formation in French Verbs: Structure, Rules, and Allomorphy

Gustavo L. Estivalet [1,2,*] and Fanny Meunier [1,2,3]

[1] Université Claude Bernard Lyon 1 (UCBL)
[2] Laboratoire sur le Langage, le Cerveau et la Cognition (L2C2, CNRS UMR5304)
[3] Laboratoire Bases, Corpus, Langage (BCL, CNRS UMR7320)

Questions

Which morphemes are represented in the mental lexicon? Theme vowel? Root? Stem?

How stems from specific micro-class are processed?

Objective

Investigate the Theme vowel representation and the verbal structure processing in the French mental lexicon

Verb Type	Control	Test	Identity	*Target*
a) 1st [-er]	aimer	parler	parlons	***PARLONS***
b) 3rd [-ir]	ouvrir	dormir	dormons	***DORMONS***
c) 3rd [-ire]	construire	écrire	écrivons	***ECRIVONS***
d) 3rd [-indre]	plaindre	joindre	joignons	***JOIGNONS***

Method

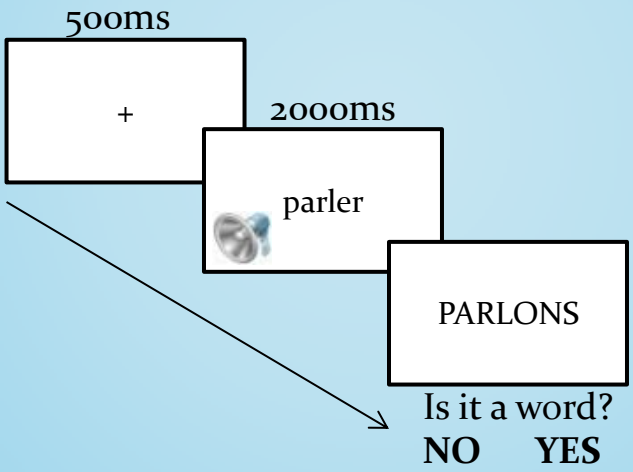
Target: 1st plural present inflected form [-ons]

Prime predictions:
-**Identity** = same target (full priming)
-**Control** = different infinitive (no priming)
-**Test** = target infinitive (?)

Stimuli:
-4 verb types, 3 conditions
-Experimental: 84 pair of verbs
(21 per verb type)
-Fillers: 84 pairs of pseudowords

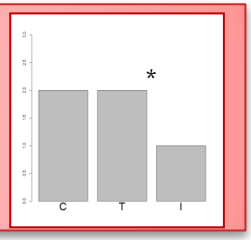
Experiment1: cross-modal priming

Subjects: N=54, 27 women, mean age 21.82, French as L1

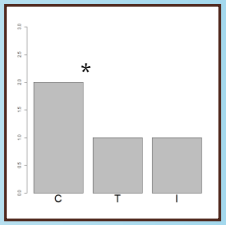


Hypothesis

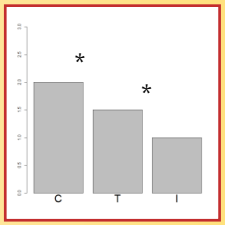
H0: No Priming = different representation
a) whole-word representation: [word]

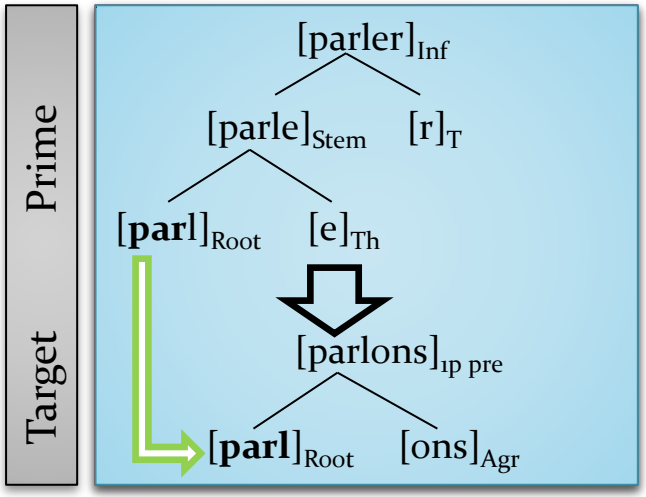


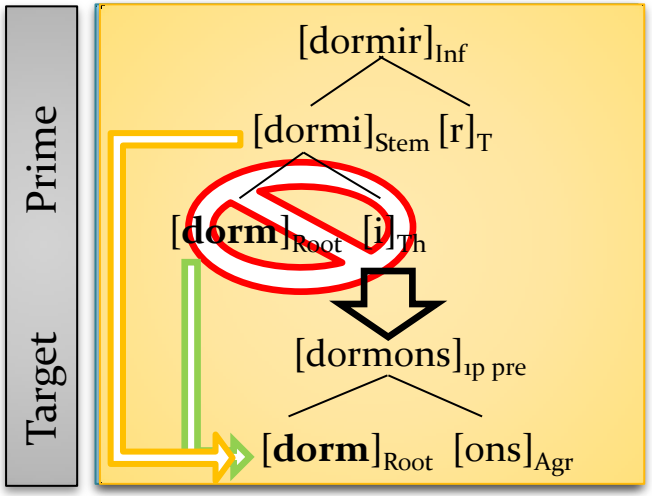
H1: Full Priming = same representation
a) full decomposition: [[[√][Th]][[T][Agr]]]
b) root-based stem formation



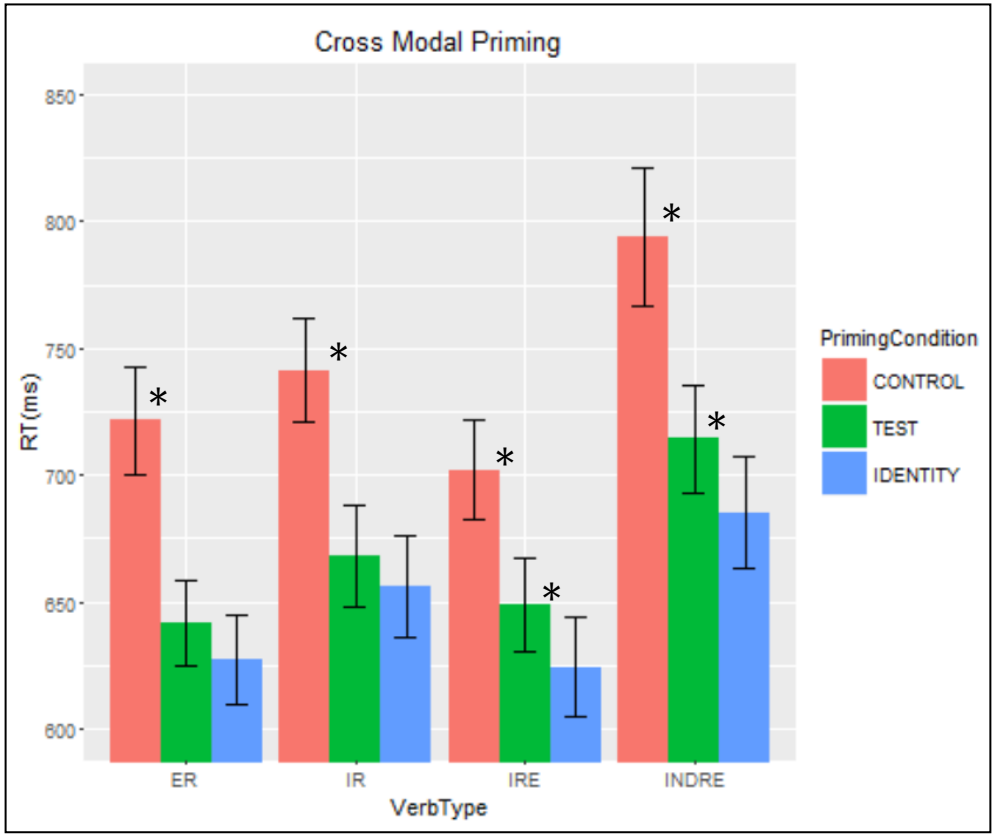
H2: Partial Priming= linked representation
a) partial decomposition [[Stem][[T][Agr]]]
b) stem-based allomorphic storage







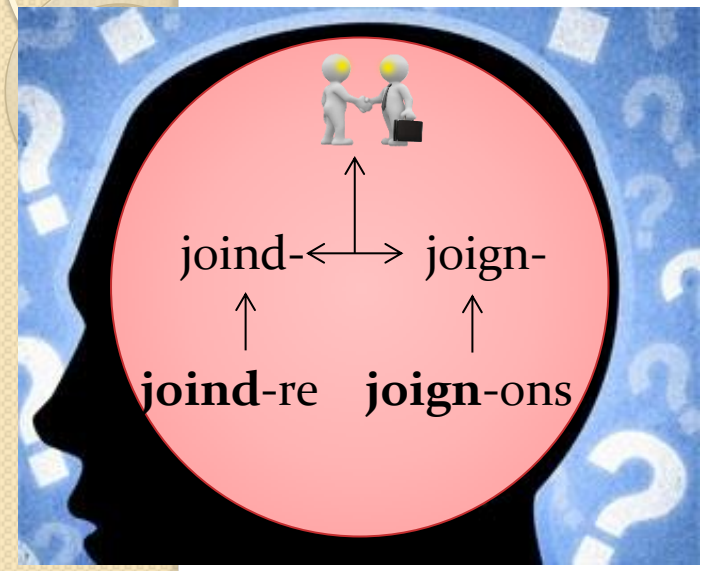
Results

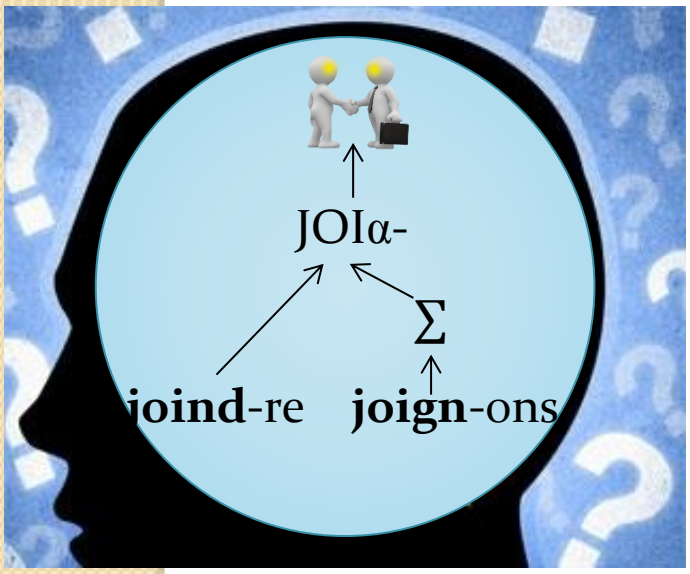


Full priming = [-er], [-ir]:
Th representation; same structure
Full decomposition
Micro-class productivity

Partial priming = [-ire], [-indre]
Different stem representations or stem rules

Word and Paradigm Vs. Item and Process





[-ir]
dor**m**ir → dor**m**ons
dor**m**ir → dort

$\sqrt{C} \rightarrow \emptyset / \_C_{[suffix]}$

[-ire]
dire → di**s**ons
écrire → écri**v**ons

$\sqrt{} \rightarrow \sqrt{<s>} / \_V_{[stress]}$
$\sqrt{} \rightarrow \sqrt{<v>} / \_V_{[stress]}$

[-indre]
joi**nd**re → joi**gn**ons

$\sqrt{<nd>} \rightarrow \sqrt{<gn>} / \_V_{[stress]}$

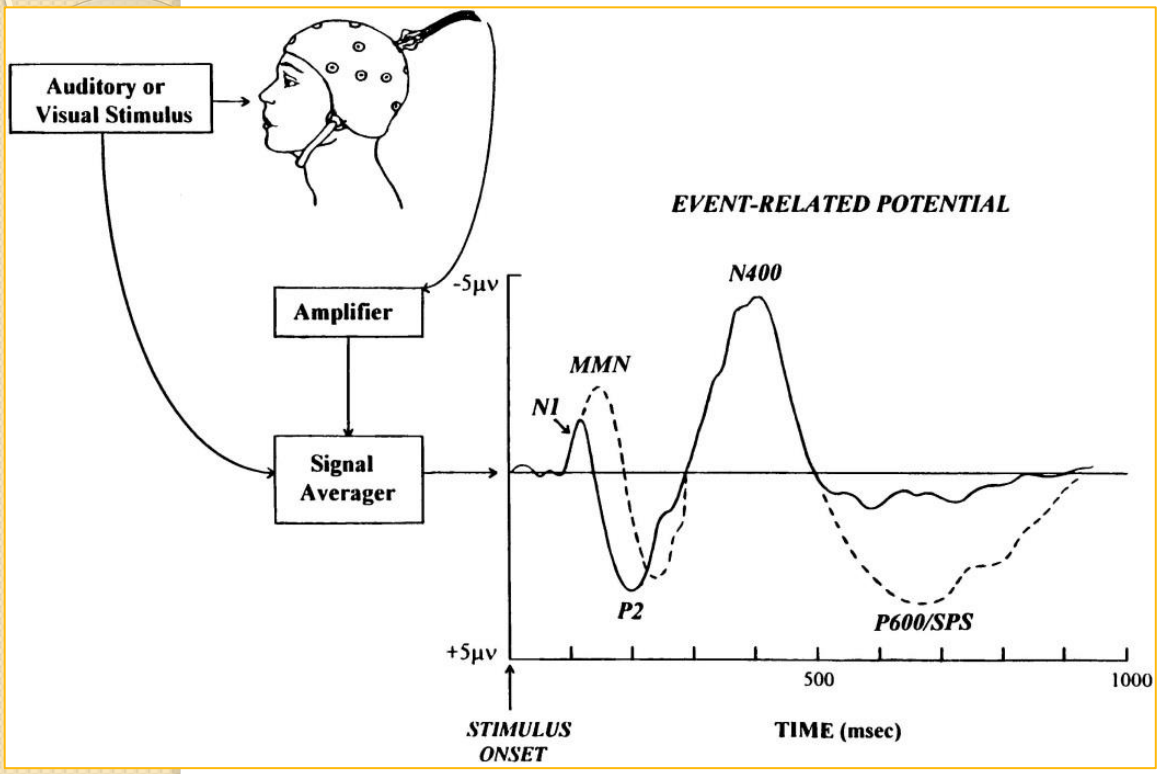
Discussion

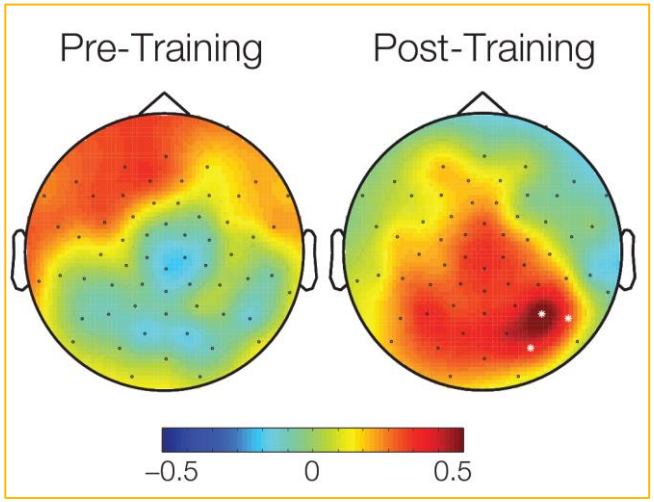
French = single combinatorial mechanism in 1st (Meunier & Marslen-Wilson, 2004), 2nd (Bonami et al., 2008), and 3rd classes (here!) ≠ other Romance languages

Theme vowel, root and structure representations in stem formation (Foley, 1979)

Allomorphic stem representations or allomorphic rules in stem formation (Embick & Halle, 2005)

Electroencephalography (EEG)





- Electric current in piramidal neurons
- High temporal resolution, spacial resolution accordingly to the number of electrodes
- Sensorial activity, cognitive activity

Study 2: Electrocircuiting Pseudoverbs


caligrama
revista de estudos românicos

Processing of the French verb by lexical decomposition

Gustavo L. Estivalet[1,2,*]; Fanny Meunier[1,2]

[1]Université Claude Bernard Lyon 1 (UCBL)
[2]Laboratoire sur le Langage, le Cerveau et la Cognition (L2C2, CNRS UMR5304)

Questions

Which is the cost in the processing the different morphemes ?

Which is the morphological processing hierarchy?

Which is the time-course of the morphological processing?

Objective

Investigate the time-course of the morphological processing in French inflected (pseudo)verbs

Verb Type	Roots	Agreement
a) Morphological Illegal (MI)	*barl	*ond
b) Only Base (OB)	parl	*ond
c) Only Suffix (OS)	*barl	ons
d) Existent Morphological Legal (EML)	parl	ons
e) Inexistent Morphological Legal (IML)	parl	ont*

Behavioral Results

Predictions

Whole-Word

EML < MI = OB = OS = IML

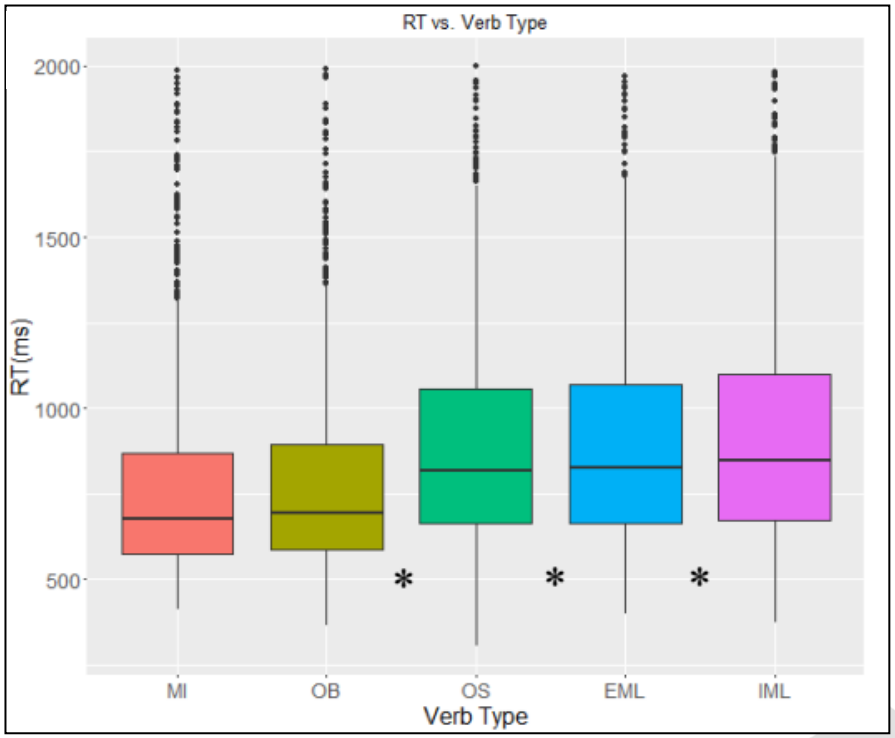
Obligatory Decomposition

MI = OB < OS < EML < IML

Dual-Mechanism

MI < EML < OB = OS < IML

Morphological Illegal (MI)
Only Base (OB)
Only Suffix (OS)
Existent Morphological Legal (EML)
Inexistent Morphological Legal (IML)

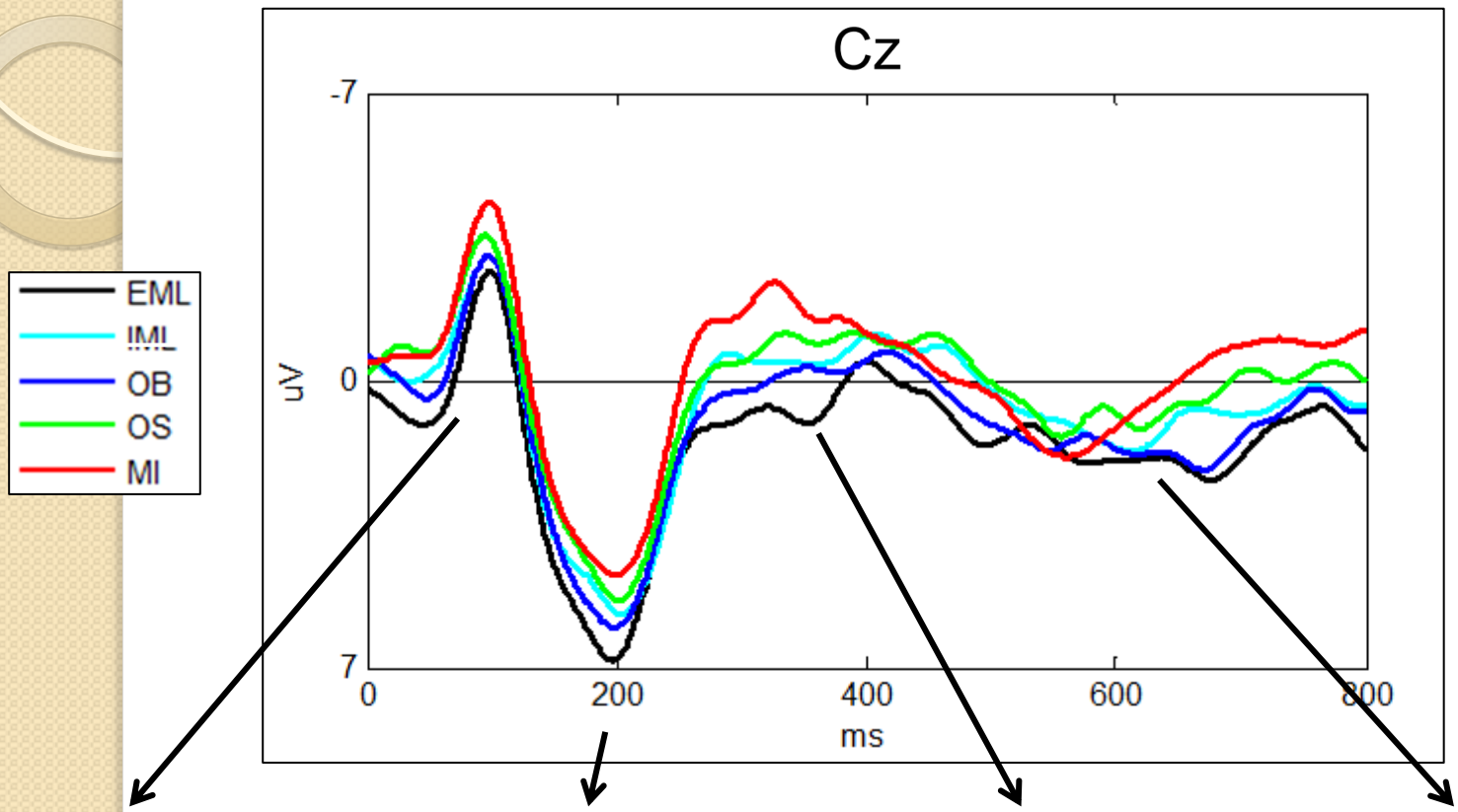


Discussion

MI = OB < OS < EML < IML

- ✓ MI do not decompose
- ✓ OB fast suffix rejection
- ✓ OS slow base rejection
- ✓ EML decompose and recombine
- ✓ IML recombination inhibition

Evoked-Related Potentials (ERP)



P1
MI, OS≠OB, IML, EML
Visual processing

N1 (50-150ms)
MI, OS≠OB, IML, EML
Orthography

P200 (150-250ms)
EML<OB<IML<OS<MI

MI ≠ all

Searching
Acivation

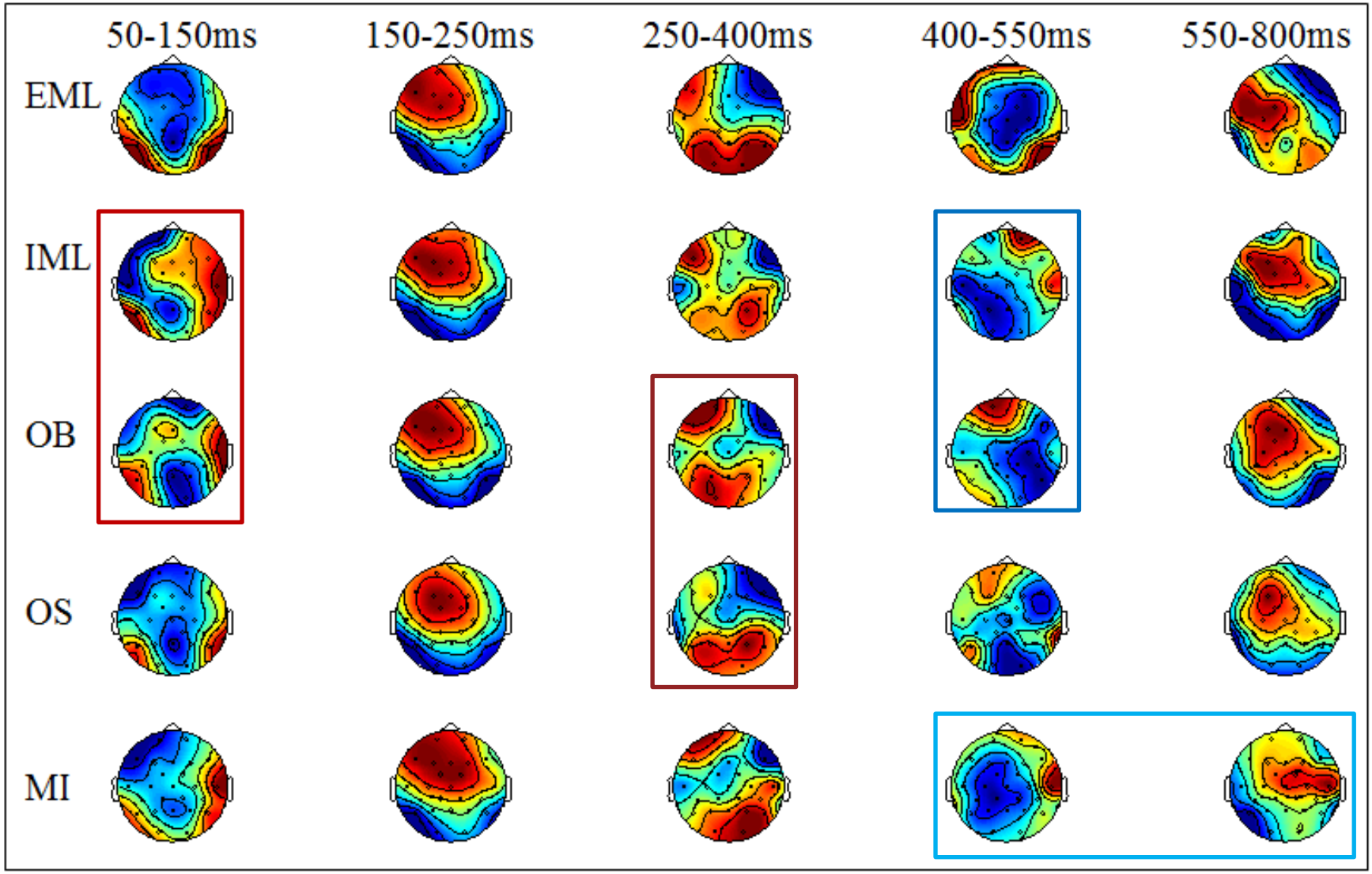
LAN (250-400ms)
OS ≠ MI, OB, EML
Morphosyntax

N400 (400-550ms)
MI, OS, IML≠OB, EML
Lexeme Activation

P600 (550-800ms)
OS ≠ IML ≠ OB, EML
Reanalysis
Repair

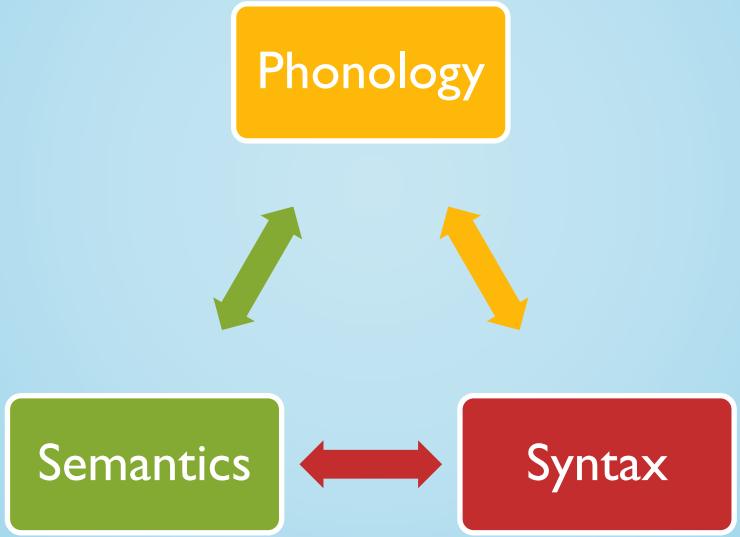
Obs.:
MI/OB/EML
2 cycles

ERP Topography



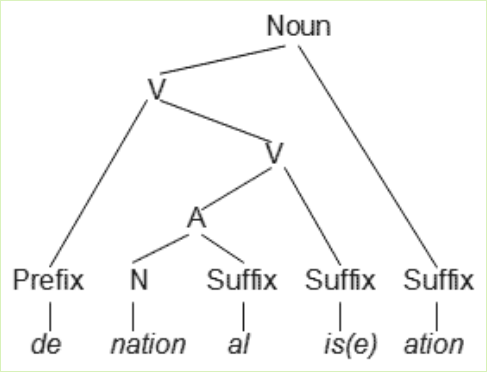
Concluding

- Morphology allows the the research in the different grammatical interfaces



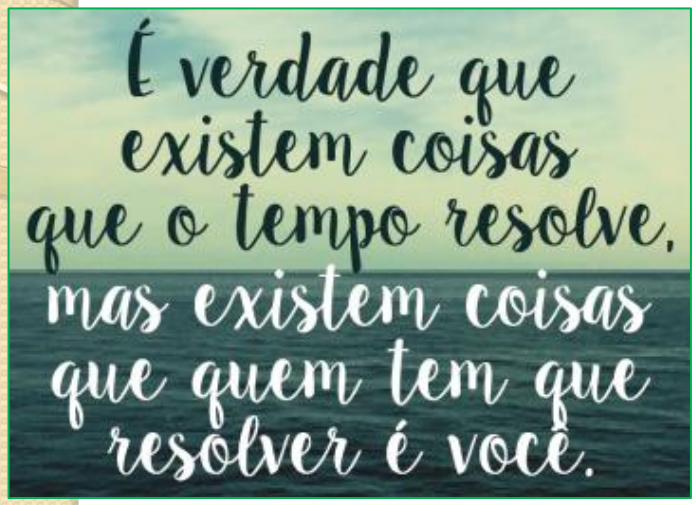
- Words seem to be stored on atomic units and computed on each realization

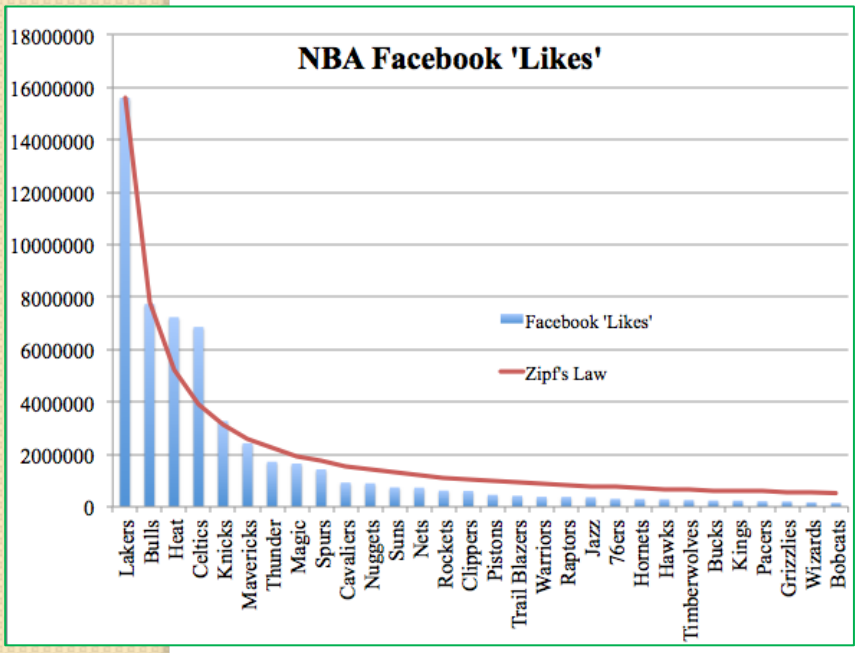
- How about the differences between inflection and derivation?



	present	*imperfect*	*preterite*	*pluperfect*	*future*	*condit. present*
1s	falo	falava	falei	falara	falarei	falaria
2s	falas	falavas	falaste	falaras	falarás	falarias
3s	fala	falava	falou	falara	falará	falaria
1p	falamos	falávamos	falámos	faláramos	falaremos	falaríamos
2p	falais	faláveis	falastes	faláreis	falareis	falaríeis
3p	falam	falavam	falaram	falaram	falarão	falariam

Corpus baseado em palavras





Total de palavras: 19

Palavra	**Freq. Lexema**
QUE:	4
COISAS:	2
É:	2
EXISTEM:	2
VERDADE, O, TEMPO, RESOLVE, MAS, QUEM, TEM, RESOLVER, VOCÊ:	1

Lema	**Freq. Lema**
RESOLV-:	2

Total	**14**	**1.000.000**
QUE	4	285.714
É	2	142.857
TEMPO	1	71.428
CEDEIS	2.000.000 = 1	
	1.000.000 = 0,5	

Informações do LexPorBR

Nb	Column	Description
1	orthography	Orthographic representation
2	gram_cat	Grammatical category
3	gram_inf	Grammatical information
4	ortho_freq	Orthographic frequency
5	ortho_freq/M	Orthographic frequency per million
6	log10_ortho_freq	Log10 from ortho_freq
7	zipf_scale	Standardized frequency scale
8	zipf_rank	Zipf's rank-frequency distribution
9	nb_letters	Number of letters
10	nb_homogr	Number of homographs
11	homographs	Homograph grammatical categories
12	pu_ortho	Orthographic uniqueness point
13	ortho_neigh	Orthographic neighborhood
14	old20	Orthographic Levenshtein Distance 20 words
15	cvcv_ortho	Consonant/vowel CVCV structure
16	bigrams	Bigrams representation
17	bigram_freq	Bigram frequency
18	trigrams	Trigrams representation
19	trigram_freq	Trigram frequency
20	rev_ortho	Reverse orthography
21	rev_cvcv_ortho	Reverse CVCV structure
22	rev_bigrams	Reverse bigrams
23	rev_trigrams	Reverse trigrams
24	random	Random number between 0–1
25	id	Identity number (position)

Em breve…

MPLPB
1008 palavras
1008 pseudopalavras
RT/ACC

SUBTLEX-PB/BR
Frequências

Fonologia
IPA/SAMPA
N Fonemas
Homófonos
PU/viz./PLD20
Bifones/trifones

Sílabas
CVCV
Tônica

Tagger
Cat./Inf. gramaticais

LexPorBR


PLOS | ONE

RESEARCH ARTICLE

The Brazilian Portuguese Lexicon: An Instrument for Psycholinguistic Research

Gustavo L. Estivalet[1,2]*, Fanny Meunier[1,2]

1 CNRS UMR5304, Laboratoire sur le Langage, le Cerveau et la Cognition, Institut de Sciences Cognitives, Bron, France, 2 Université Claude Bernard Lyon 1, Université de Lyon, Lyon, France

SOLETRAS REVISTA

N. 33 – 2017.1 – GUSTAVO LOPEZ ESTIVALET FANNY MEUNIER

Corpus psicolinguístico Léxico do Português Brasileiro

Gustavo Lopez Estivalet[1]

Fanny Meunier[2]

LexPorBR - Léxico

Não seguro | www.lexicodoportugues.com/?textarea=palavras%0D

Léxico do Português Brasileiro - LexPorBR

English

Selecione o idioma

Links

Léxico
Pseudopalavras
Downloads
Ferramentas
Atualizações
Créditos
Linguistica Estatistica
Linguateca
NILC

Manual: LexPorBR

Pesquisa simples

palavras
em
linhas

Ordenar por:
ortografia | crescente | Procurar
Limpar

Pesquisa complexa

1 | sim
2 | sim
3 | sim
4 | sim

Ordenar por:
crescente | Procurar | Limpar
+ Critérios

Utilize

_ : substituir uma ou mais letras
% : subsituir uma cadeia de letras
< : menor que
> : maior que

Categorias gramaticais

adj, adv, gram, nom, num, ver

Resultados

Anterior | Posterior | Exportar .csv

Página 1 de 1
0 - 3 palavras de um total de 3 palavras encontradas

Estatísticas

categoria	freq_orto	log10_freq_orto	zipf_escala	nb_letras	viz_orto	old20
Média	1861.0000	2.29313333	3.80593333	7.3333	4.6667	1.633333
Minimo	1	0.0000	1.5128	6	2	1.40
Maximo	3254	3.5124	5.0252	8	10	1.75

ortografia	cat_gram	inf_gram	freq_orto	freq_orto/M	log10_freq_orto	zipf_escala	nb_letras	nb_homogr	homografas	pu_orto	viz_orto	old20	cvcv_orto	bigramas	trigramas
linhas	nom		2328	74.1929	3.3670	4.8798	6	1		6	10	1.40	CVCCVC	#l_li_in_nh_ha_as_s#	#li_lin_inh_nha_has_as#
palavras	nom		3254	103.7043	3.5124	5.0252	8	2	adv,	8	2	1.75	CVCVCCVC	#p_pa_al_la_av_vr_ra_as_s#	#pa_pal_ala_lav_avr_vra_ras_as#
palavras	adv		1	0.0319	0.0000	1.5128	8	2	nom,	8	2	1.75	CVCVCCVC	#p_pa_al_la_av_vr_ra_as_s#	#pa_pal_ala_lav_avr_vra_ras_as#

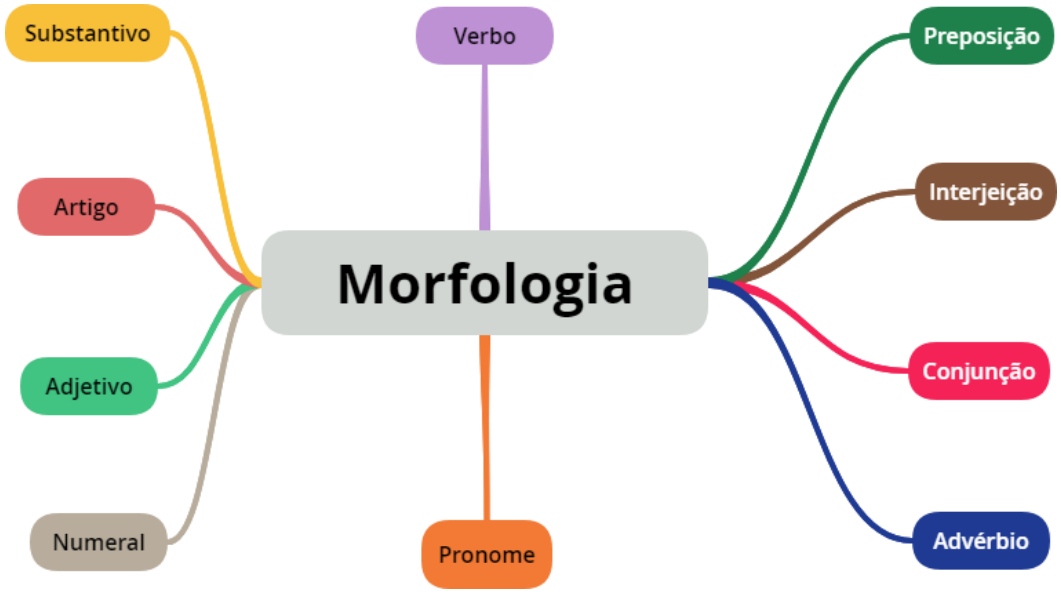
Ultima atualização: 12/10/2015.

Activity TLB

- Choose one phenomena related to morphology in Portuguese and perform the basic analysis.

- Morphology
- Morphophonology
- Morphosyntactics

- Inflection X Derivation X Compound



Obrigado pela atenção!