جامعة نيويورك أبوظبي
NYU | ABU DHABI

Queen Mary
University of London

Univerza v Ljubljani

SAVANT

SYSTEMATICITY AND VARIATION IN WORD STRUCTURE PROCESSING ACROSS LANGUAGES: A NEURO-TYPOLOGY APPROACH

LINNAEA STOCKALL
DAVE KENNETH CAYADO
DUSTIN ALFONSO CHACÓN
KARIN KAVČIČ
SWARNENDU MOITRA
CHRSITINA MANOUILLIDOU
ALEC MARANTZ
SUHAIL MATAR
BOJANA RISTIĆ
SAMANTHA WRAY
ROK ŽAUCER

X-PPL 2002 September 12-13, 2022 University of Zurich

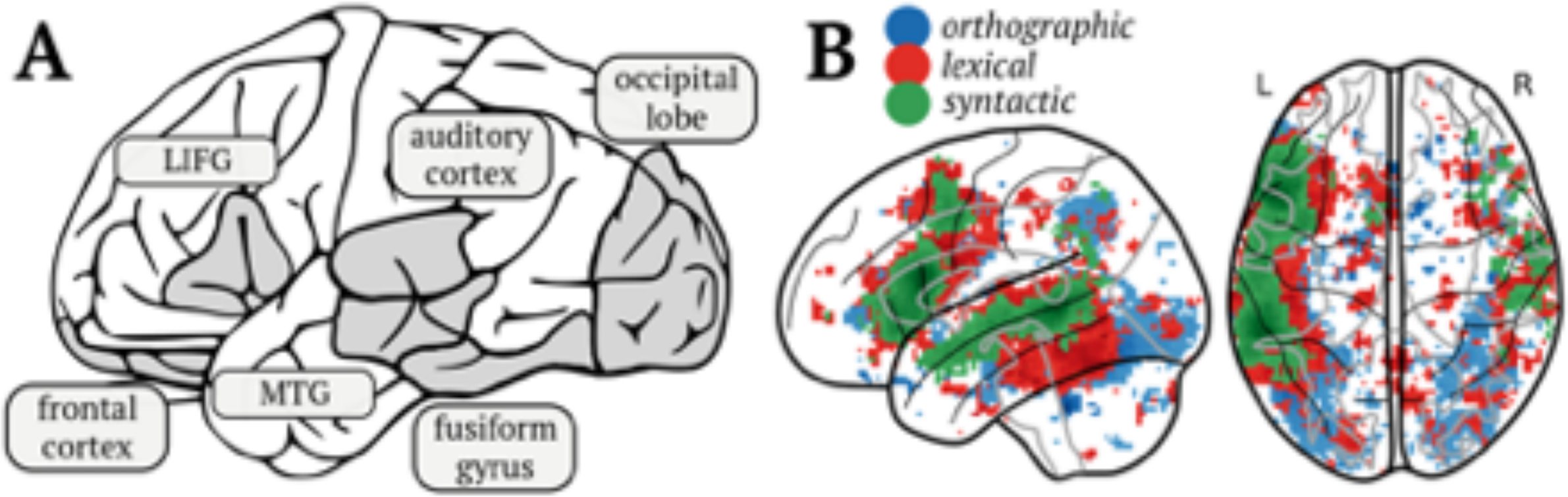
UKRI Economic and Social Research Council

THE TEAM/PROJECT STRUCTURE

Core Managment Team	PI: Linnaea Stockall, QMUL + Co-I: Alec Marantz, NYU/NYUAD and Co-I: Christina Manouilidou, U v Ljubljana

Languages	**Slovenian**	**Bosnian/ Croatian/ Serbian**	**Arabic**	**Tagalog**	**Bangla**
Language Team Leaders	Rok Žaucer, U v Nova Gorica		Samantha Wray, Dartmouth		Dustin Chacón, NYUAD/ Georgia
Research Staff	Karin Kavčič	Bojana Ristić	Suhail Matar	Dave Cayado	Swarnendu Moitra

BACKGROUND: COGNITIVE NEUROSCIENCE OF LANGUAGE



- well established, left-lateralised language network
- validated across: multiple modalities, multiple methods, the lifespan, atypical populations, and a range of languages (Poeppel et al 2012)

BUT LIMITED RANGE OF LANGUAGES STUDIED

Table 2. *Activation in the anterior temporal lobe*

Study	Coordinates	Location	
Mazoyer et al. (1993)	[Talairach]		French
syn (sent > WL, allowing local structure) sem (meaningful > meaningless sent)	no coordinates no coordinates	L temp pole L MTG/STG	
Stowe et al. (1998) visual	[Talairach]		Dutch
syn (sent > WL, with local structure) no sem condition	−40, 2, −20	L ant MTG/STG	
Friederici et al. (2000)			German
syn (sent > WL) (real words, only nouns) syn (sent > WL) (pseudowords) sem (real > pseudowords)	no coordinates no coordinates no effect	L ant STG L ant STG	
Humphries et al. (2005)	[MNI]		English
syn (sent > WL) no sem condition	−54, −1,- 24	L ant STS/MTG	
Humphries et al. (2006)	[MNI]		English
syn (sent > scrambled WL) interaction (sem/syn) sem (real > pseudowords)	−47, 14, −28 −54, 10, −8 no effect	L ant STS L ant STG	
Snijders et al. (2009)	[MNI]		Dutch
syn (sent > WL, allowing local structure) no sem condition	−54, 18, −30	L temp pole	
Friederici et al. (2003)	[Talairach]		German
syn violation (phrase structure) sem violation (select. restriction)	−53, −1, 0 no effect	L ant STG	
Vandenberghe et al. (2002) visual	[Talairach]		English
syn viol. (incorrect word order) sem viol. (sem random sent.) interaction (sem/syn)	−44, −6, −24 −48, −4, −20 −40, 8, −28	L ant temp pole L ant temp pole L ant temp pole	

Presence/absence of syntactic (syn) (sentences > word list) or semantic (sem) information (real words > pseudowords), violation of syntactic, or semantic information. L, left; ant, anterior; temp, temporal; STG, superior temporal gyrus; MTG, middle temporal gyrus; STS, superior temporal sulcus.

- Friederici, 2011. Physiol Rev 91: 1357–1392

WHY IS THIS A PROBLEM?

1. neurolinguistic theory building and testing largely ignores the vast linguistic diversity that plays such a critical role in all other language sciences. These well-studied languages share **an alphabet**, a basic **subject–verb–object or subject–object –verb word order**, have similar **subject–verb agreement patterns**, and make a binary **singular/ nonsingular** distinction, among many other grammatical features that vary considerably across the world's languages (Dryer & Haspelmath 2013)

WHY IS THIS A PROBLEM?

2. very narrow set of languages, and language speakers, have been essentially reified as 'worthy' of scientific inquiry (the WEIRD Science problem; Henrich et al. 2010)

Most people are not WEIRD

To understand human psychology, behavioural scientists must stop doing most of their experiments on Westerners, argue **Joseph Henrich**, **Steven J. Heine** and **Ara Norenzayan**.

People from Western, educated, industrialized, rich and democratic (WEIRD) societies – and particularly American undergraduates – are some of the most psychologically unusual people on Earth.

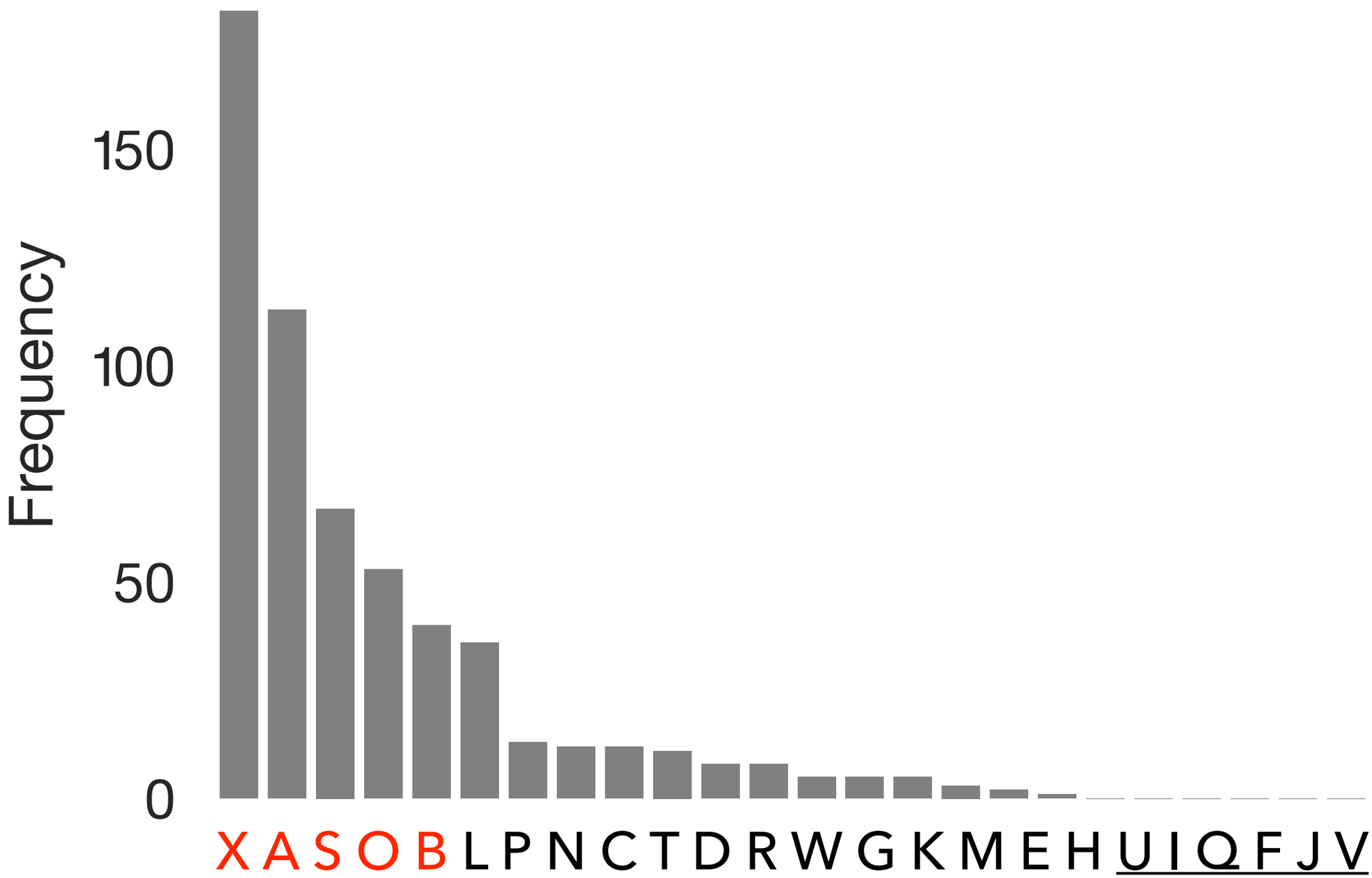
TYPOLOGICALLY-INFORMED THEORY BUILDING/TESTING

PARAMETER SETTING[a]	NUMBER SYSTEM	EXAMPLE
{ }	no number	Pirahã
{±atomic}	singular, plural	Svan
{±minimal}	minimal, augmented	Winnebago
{±minimal*}	minimal, unit augmented, augmented	Rembarrnga
{±minimal, ±atomic}	singular, dual, plural	Kiowa
{±additive, ±atomic}$_1$	singular, paucal, plural	Bayso
{±additive, ±atomic}$_2$	singular, plural, greater plural	Fula
{±additive, ±minimal}	minimal, paucal, plural	Mebengokre
{±minimal*, ±atomic}	singular, dual, trial, plural	Larike
{±additive, ±minimal*}	minimal, unit augmented, paucal, plural[b]	—?
{±additive*, ±atomic}	singular, plural, greater and greatest plurals	Banyun
{±additive*, ±minimal}	minimal, paucal, greater paucal, plural[b]	—?
{±additive*, ±minimal*}	minimal, unit aug., paucal, gr. paucal, plural[b]	—?
{±additive, ±minimal, ±atomic}$_1$	singular, dual, paucal, plural	Yimas
{±additive, ±minimal, ±atomic}$_2$	singular, dual, plural, greater plural	Mokilese
{±additive, ±minimal*, ±atomic}	singular, dual, trial, paucal, plural	Marshallese
{±additive*, ±minimal, ±atomic}$_1$	singular, dual, paucal, greater paucal, plural	Sursurunga
{±additive*, ±minimal, ±atomic}$_2$	singular, dual, paucal, plural, greater plural	Mele-Fila
{±additive*, ±minimal*, ±atomic}	singular, dual, trial, paucal, gr. paucal, plural[b]	—?

TABLE 3. A typology of number systems.

- Harbour, 2014. Language. Volume 90, Number 1

TYPOLOGICALLY-INFORMED THEORY BUILDING/TESTING:

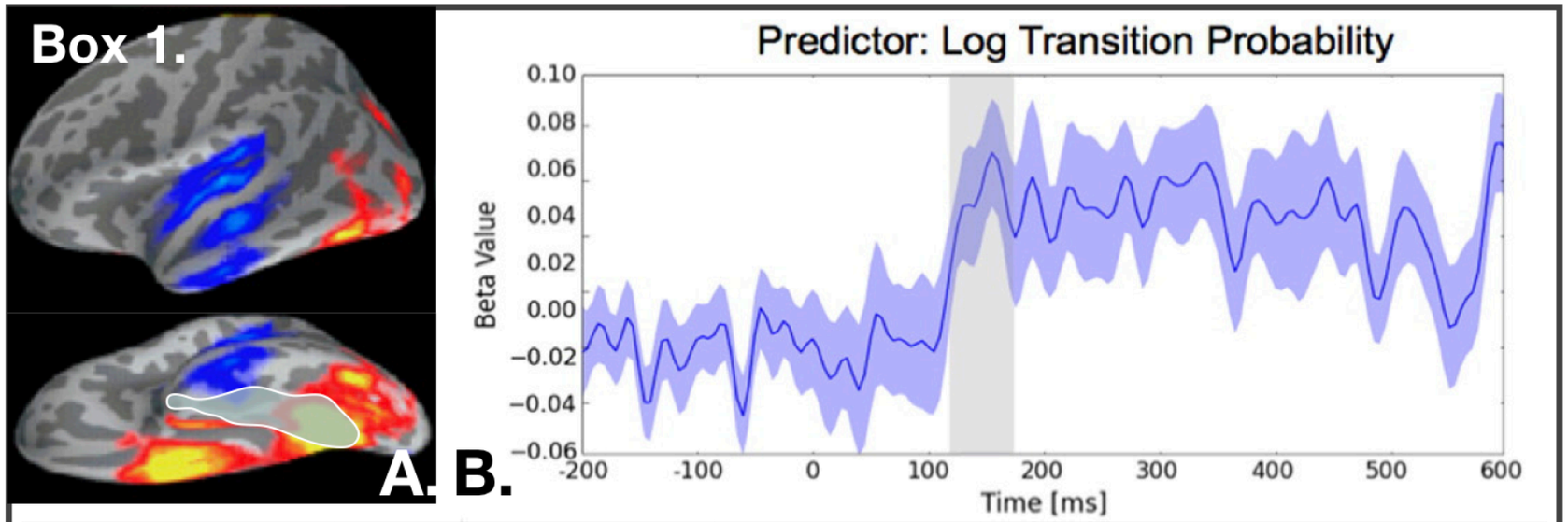


A: Dem-Num-A-N
B: Dem-Num-N-A
C: Dem-N-Num-A
D: N-Dem-Num-A
E: Num-Dem-A-N
F: Num-Dem-N-A
G: Num-N-Dem-A
H: N-Num-Dem-A
I: A-Dem-Num-N
J: A-Dem-N-Num
K: A-N-Dem-Num
L: N-A-Dem-Num
M: Dem-A-Num-N
N: Dem-A-N-Num
O: Dem-N-A-Num
P: N-Dem-A-Num
Q: Num-A-Dem-N
R: Num-A-N-Dem
S: Num-N-A-Dem
T: N-Num-A-Dem
U: A-Num-Dem-N
V: A-Num-N-Dem
W: A-N-Num-Dem
X: N-A-Num-Dem

WHY IS NEUROLINGUISTIC RESEARCH SO NARROW?

- Such research requires:
- (a) the use of expensive, mostly nonportable technology, typically only located in large cities in the developed world
- (b) researchers with expertise in the tools and techniques of neurolinguistics and native or near-native speaker competence in the languages under investigation in order to design appropriate materials, and
- (c) appropriately annotated linguistic corpora to ensure these materials are suitable for experimental designs, which don't exist for most languages.

OTHER NEURO-TYPOLOGY PROGRAMS: CUES FOR MORPHOLOGICAL DECOMPOSITION



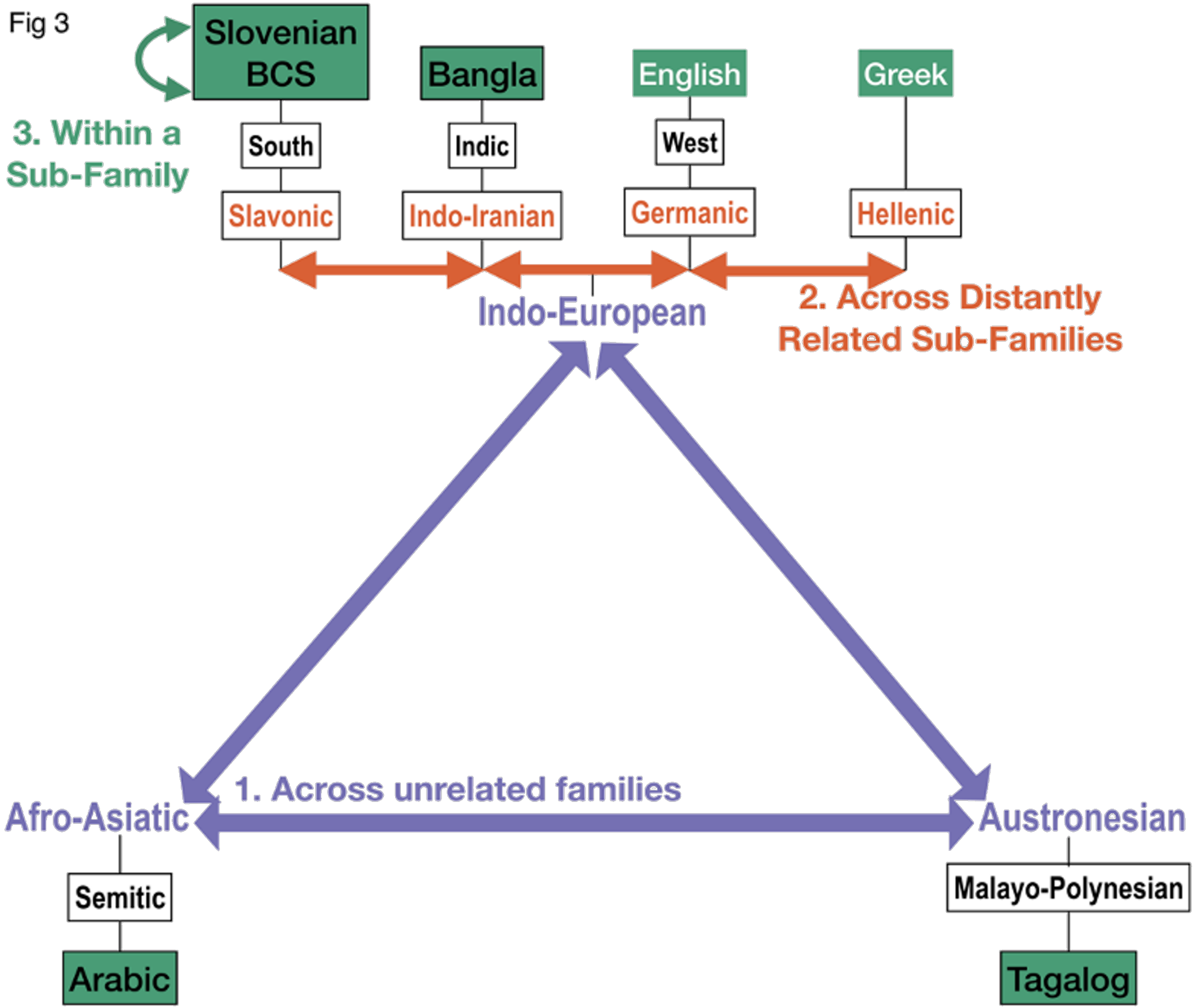
Language	Affix Type	Example	M170/TP Corr
English	suffix	teach<u>er</u>	LH, positive
English	pseudosuffix	broth<u>er</u>	LH, positive
English	prefix	<u>re</u>fill	**RH**, positive
Greek	suffix	φαγο<u>σιμοσ</u>	LH, positive
Tagalog	reduplicant prefix	<u>halo</u>halo	LH, positive
Tagalog	infix	t<u>in</u>awag	LH, **negative**
Japanese	suffix	染ま<u>る</u>	LH, positive

Cross-linguistic investigations of morpho-orthographic parsing using MEG consistently find correlations between activity originating in the left Fusiform Gyrus [A.] ~170ms after the onset of a visually presented complex word, and the transition probability from stem to whole word (TP), with increased M170 activity the more likely the whole word is given the stem [B.] (English: Lewis et al., 2011, Gwilliams et al., 2016, Greek: Neopytou et al., 2018, Japanese: Ohta et al., 2019). However, this M170/TP correlation goes in the opposite direction for Tagalog infixed words (Wray et al., submitted), and is right hemisphere lateralised for English prefixed words (Stockall et al., 2019), suggesting differences in the way form frequency cues to morphological structure are used in early parsing. Plot A from (Lewis et al., 2011); B from (Gwilliams et al., 2016)

THE SAVANT PROJECT:

- A program of Neuro-Typology, in which we combine the extremely high temporal resolution of MEG and EEG with the simplicity and practicality of a morphological manipulation paradigm implemented systematically across a typologically diverse and largely understudied set of languages: **Bosnian/Serbian/Croatian, Slovenian**, **Bangla**, **Tagalog**, **Arabic, Greek & English**.
- This paradigm allows the simultaneous investigation of distinct stages of processing that have not previously been investigated within a single experiment.

Fig 3


Slovenian BCS
Bangla
English
Greek
3. Within a Sub-Family
South
Indic
West
Slavonic
Indo-Iranian
Germanic
Hellenic
Indo-European
2. Across Distantly Related Sub-Families
1. Across unrelated families
Afro-Asiatic
Austronesian
Semitic
Malayo-Polynesian
Arabic
Tagalog

THE PARADIGM: DETECTING SENSITIVITY TO WORD-INTERNAL GRAMMAR

Brain and Language 183 (2018) 11–20

Contents lists available at ScienceDirect

Brain and Language

journal homepage: www.elsevier.com/locate/b&l

Syntactic and semantic restrictions on morphological recomposition: MEG evidence from Greek

K. Neophytou[a,*], C. Manouilidou[b], L. Stockall[c,a], A. Marantz[a,d,e]

[a] NYUAD Institute, New York University Abu Dhabi, United Arab Emirates
[b] Department of Comparative and General Linguistics, University of Ljubljana, Slovenia
[c] Department of Linguistics, Queen Mary University of London, United Kingdom
[d] Department of Psychology, New York University, USA
[e] Department of Linguistics, New York University, USA

ORIGINAL RESEARCH ARTICLE

Front. Psychol., 06 September 2019 | https://doi.org/10.3389/fpsyg.2019.01964

Prefix Stripping Re-Re-Revisited: MEG Investigations of Morphological Decomposition and Recomposition

Linnaea Stockall[1*], Christina Manouilidou[2], Laura Gwilliams[3,4], Kyriaki Neophytou[4,5] and Alec Marantz[3,4,6]

THE PARADIGM: DETECTING SENSITIVITY TO WORD-INTERNAL GRAMMAR

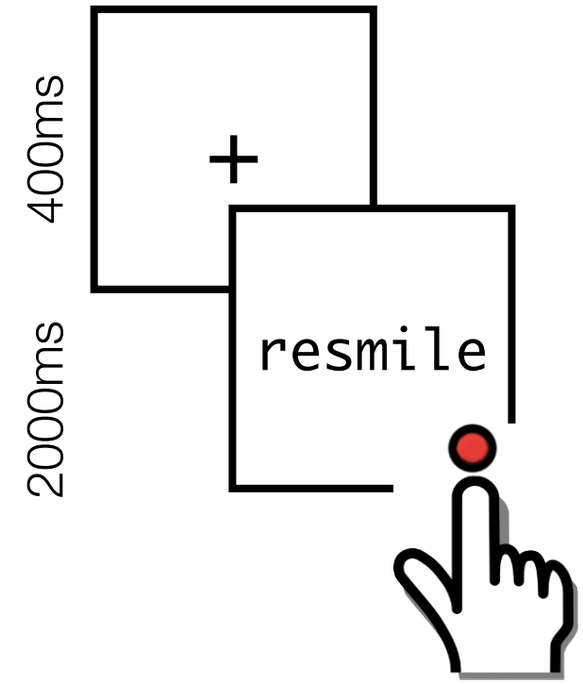
- Affixes put restrictions on the category of the stems they attach to:
  - ✓ re+verb = repaint, reclose, rewrap
  - ✗ re+noun = reidea, renapkin, resadness
  - ✗ re+adjective = rehappy, reloud, reblue

- and on the semantics:
  - Abi painted the wall again ~ Abi repainted the wall. ✓
  - Abi smiled at Kai again ~ Abi resmiled at Kai. ✗
  - Abi knew the answer again ~ Abi reknew the answer. ✗

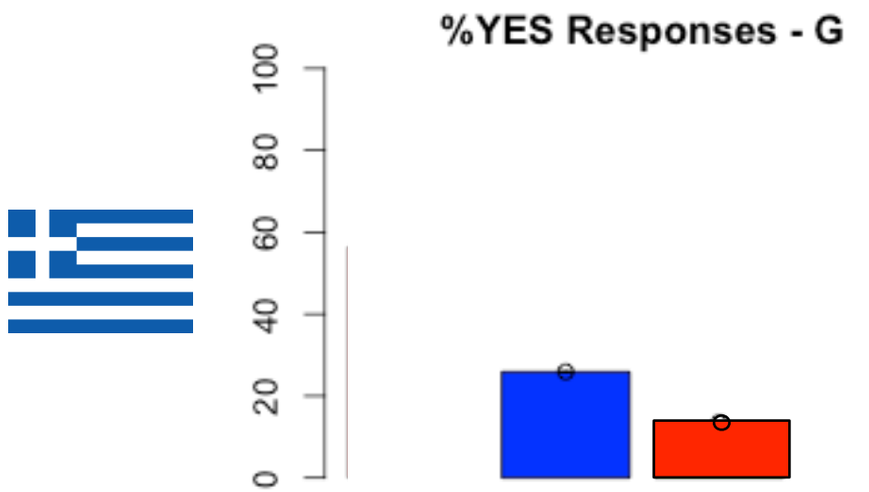
're-' can only attach to verbs

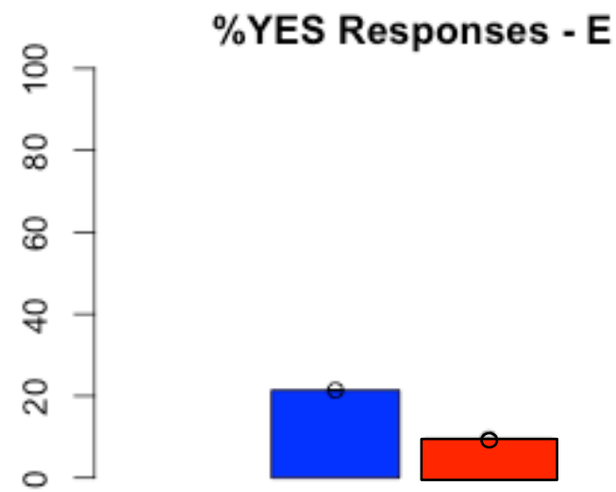
're-' can only attach to verbs that describe an action that causes a result

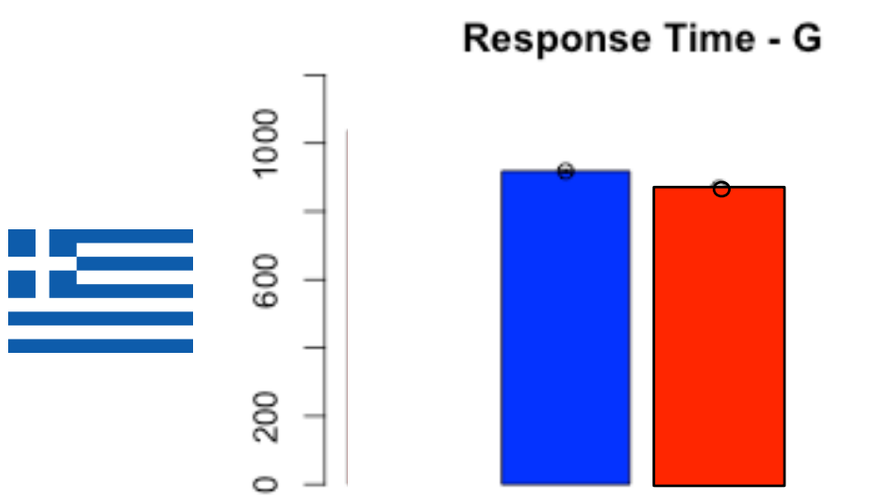
MANOUILIDOU & STOCKALL, 2014

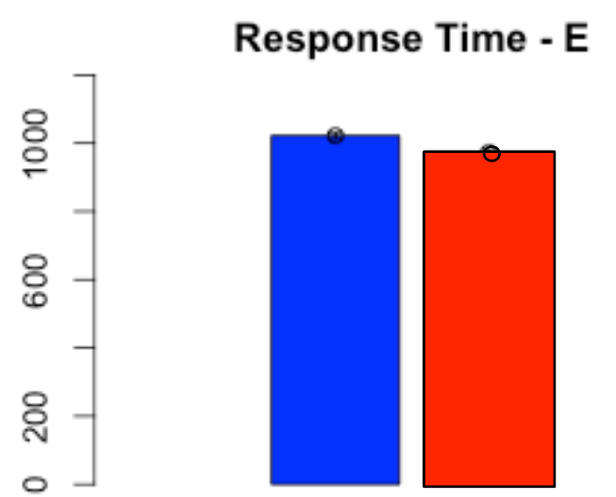


PSEUDO WORD TYPE	GREEK (N=46)	ENGLISH (N=30)
category violation	varelimos	reflat
argstruc violation	orimasismos	resmile



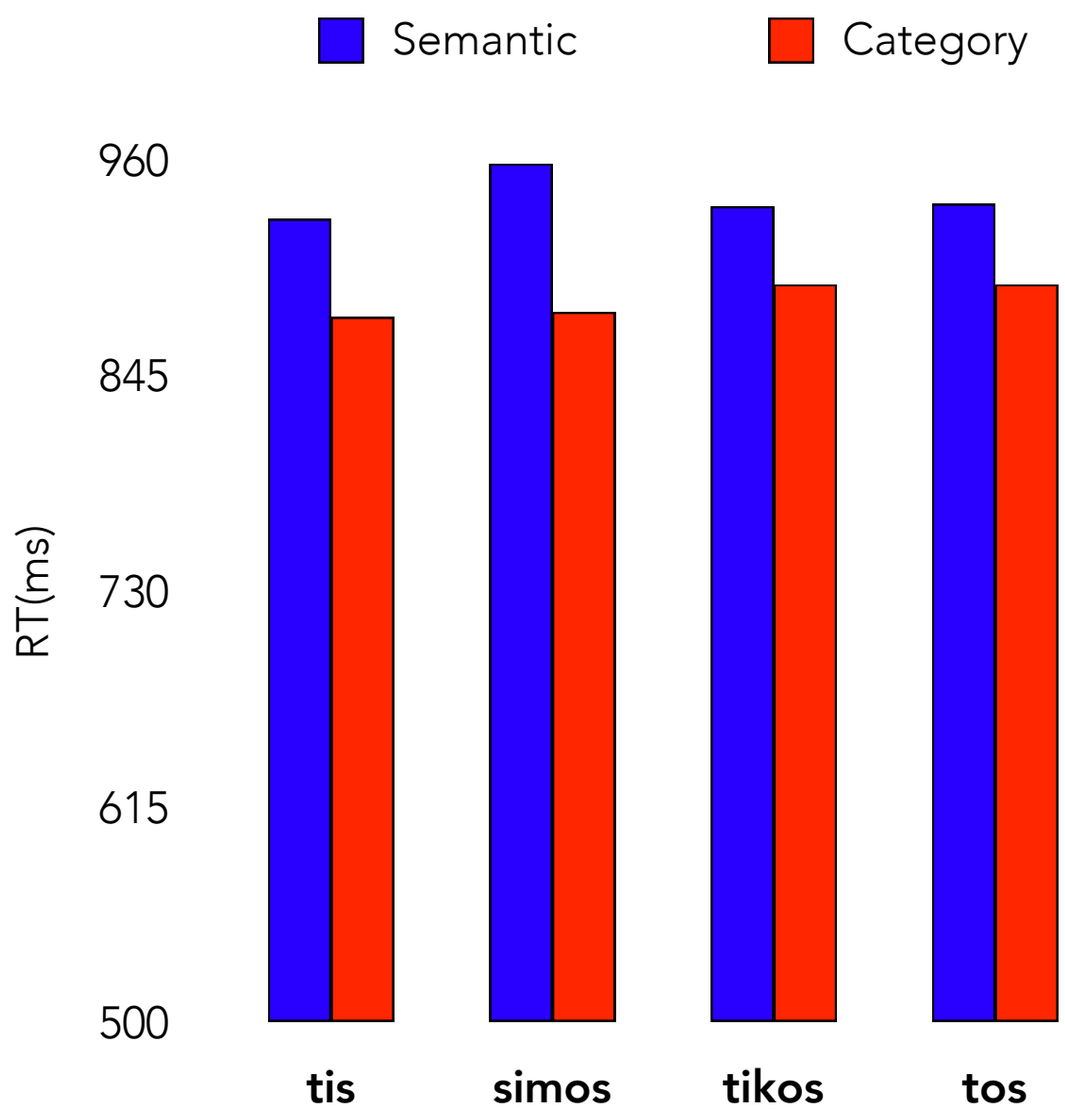




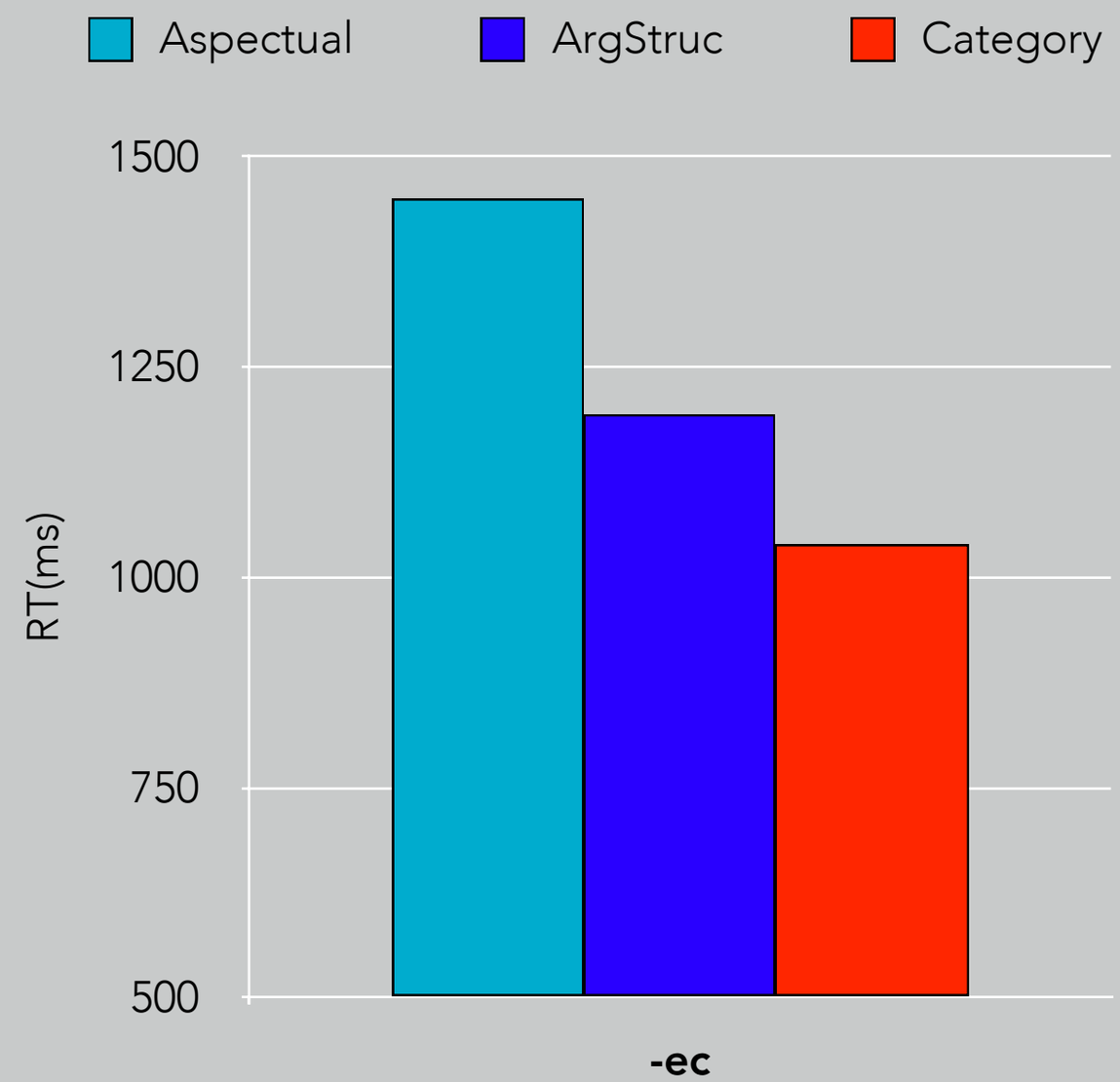


MORE BEHAVIOURAL EVIDENCE

- Manouilidou (2006): Greek deverbal adjectives and nominals



- Manouilidou, Dolenc, Marvin & Pirtošek (2016): Slovenian agentive -ec deverbal nominals



DETECTING SENSITIVITY TO WORD-INTERNAL GRAMMAR: GREEK & ENGLISH

PSEUDO WORD TYPE	GREEK	ENGLISH
syntactic category violation	varelimos (barrelable)	rehat
semantic violation	gelasimos (laughable)	relaugh

- 3 affixes for each language
- used **Magnetoencephalography** to reveal the time course of processing these distinct sources of information
- at NEL LAB جامعة نيويورك أبوظبي NYU | ABU DHABI

RESULTS: DISSOCIATION BETWEEN SYNTACTIC & SEMANTIC PROCESSING

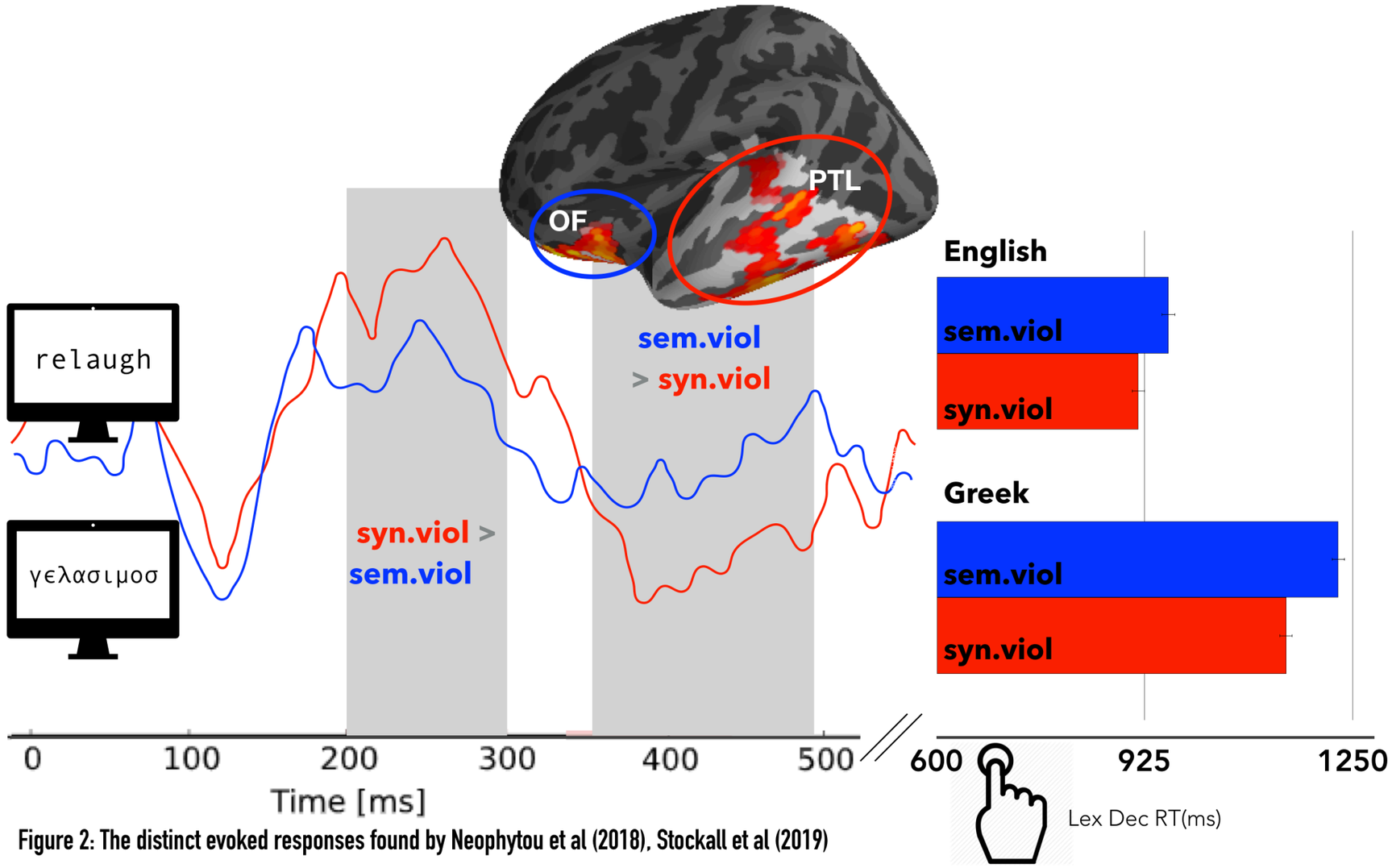


Figure 2: The distinct evoked responses found by Neophytou et al (2018), Stockall et al (2019)

- **SYNTACTIC WELL-FORMEDNESS:** from 200–300ms, we find that the syntactic-category-violating pseudowords evoke significantly larger responses than the semantics-violating pseudowords in left Posterior Temporal Lobe (PTL).
- **SEMANTIC WELLFORMEDNESS:** between 300–500ms, we find the reverse: the semantics-violating items evoke a larger response than the syntax-violating items from the OrbitoFrontal (OF) Cortex.

CONSISTENT WITH SENTENCE PROCESSING LITERATURE

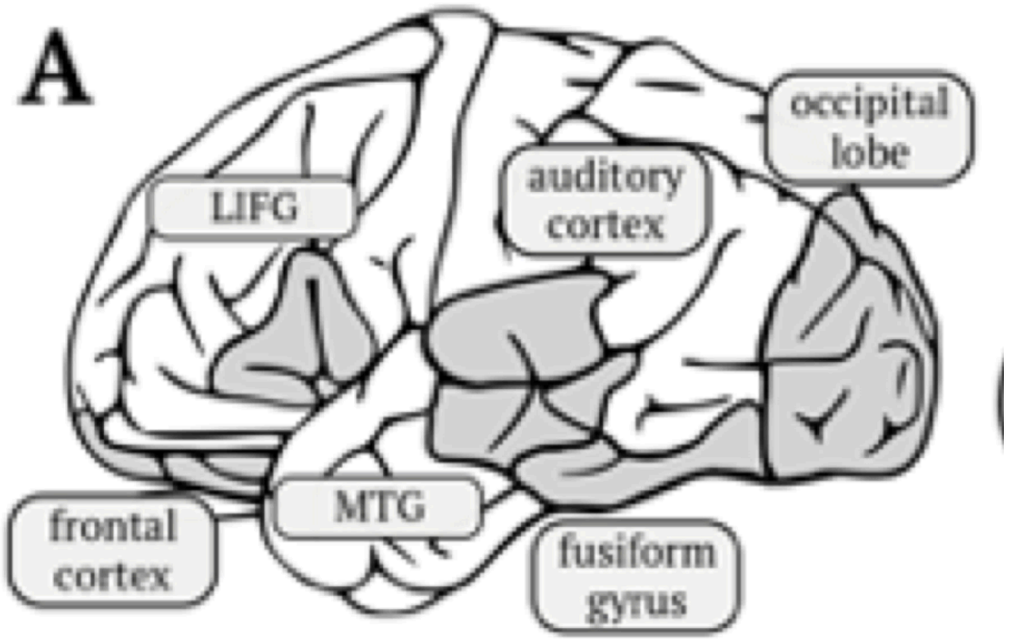
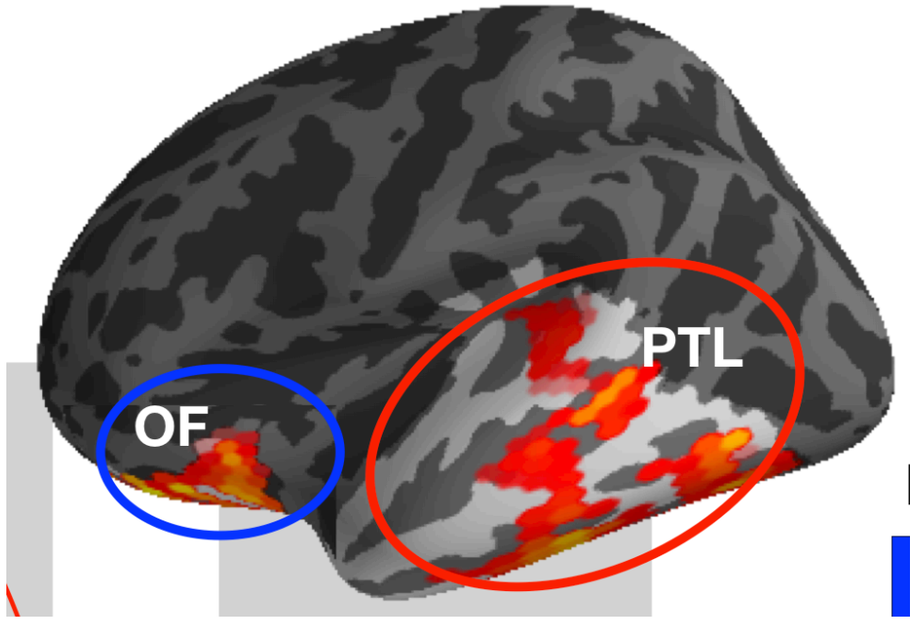


Figure 1: The left lateralised Language Network



- **PTL <-> syntactic processing:** Flick & Pylkkänen 2018, Brennan et al. 2017, Matchin, et al. 2017, 2020
- **OF <-> semantic interpretation and wellformedness:** Brennan & Pylkkänen 2008; Pylkkänen & McElree 2007; Pylkkänen, et al. 2009

WORD~SENTENCE CONTINUITY HYPOTHESIS

- The same mechanisms are involved in word-internal and sentence level grammatical processing (Halle & Marantz 1995; Borer, 1994, Marantz 2013).
- Word = sentence in many languages:

West Greenlandic: qimmersuasirusuppungalu

dog-large-buy-want-1.SG.IND-but

but I would like to buy a large dog!

- **Testing this hypothesis requires investigating theory critical languages**

COMPARISONS

Language	Writing System	Grammatical Domain	Semantic Dimension	Inflection/ Derivation	Affix Type	Fully productive
English	latin alphabet	verbs	argument structure	derivation	prefix	yes
Greek	greek alphabet	deverbal participles	argument structure	derivation	suffix	yes*
Arabic	**arabic abjad**	**roots**	argument structure/	**inflection**	prefix + CVC pattern	**semi**
Bangla	**bangla abugida**	**nominals**	**concrete/ abstract**	derivation	prefix	**semi**
Bosnian/ Croatian/ Serbian	latin alphabet	verbs	**state stability**	derivation	prefix	**variable**
Slovenian						
Tagalog	latin alphabet	verbs	voice	**inflection**	prefix	yes

✳ One of the three suffixes (-tos) is less productive when attaching to derived verbs than to roots

COMPARISONS

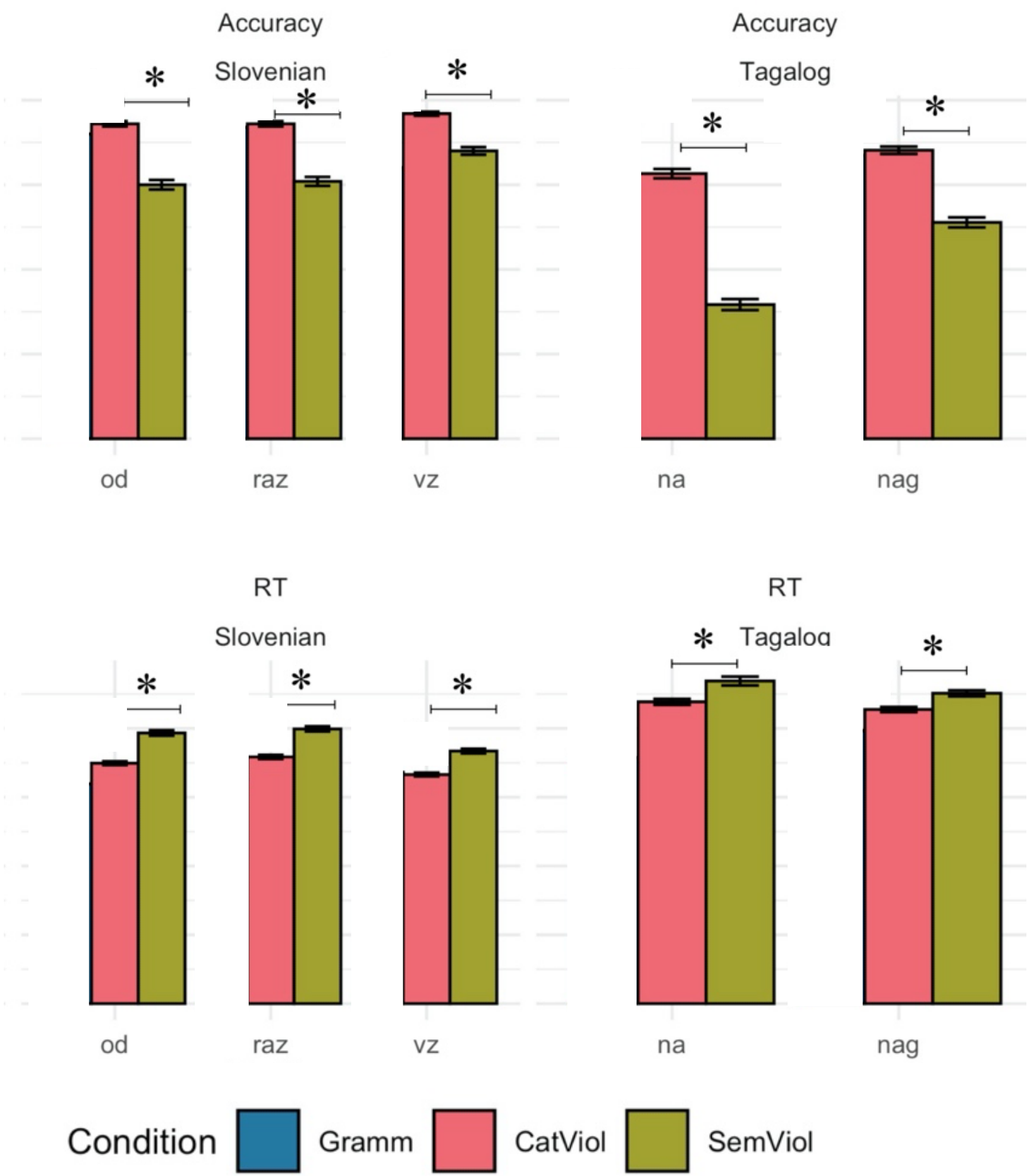
Language	Writing System	Grammatical Domain	Semantic Dimension	Inflection/ Derivation	Affix Type	Fully productive
Bangla	**bangla abugida**	**nominals**	**concrete/ abstract**	derivation	prefix	**semi**
Slovenian	latin alphabet	verbs	**state stability**	derivation	prefix	**variable**
Tagalog	latin alphabet	verbs	voice	**inflection**	prefix	yes

✳ One of the three suffixes (-tos) is less productive when attaching to derived verbs than to roots

PRELIMINARY RESULTS: ONLINE BEHAVIOURAL EXPERIMENTS

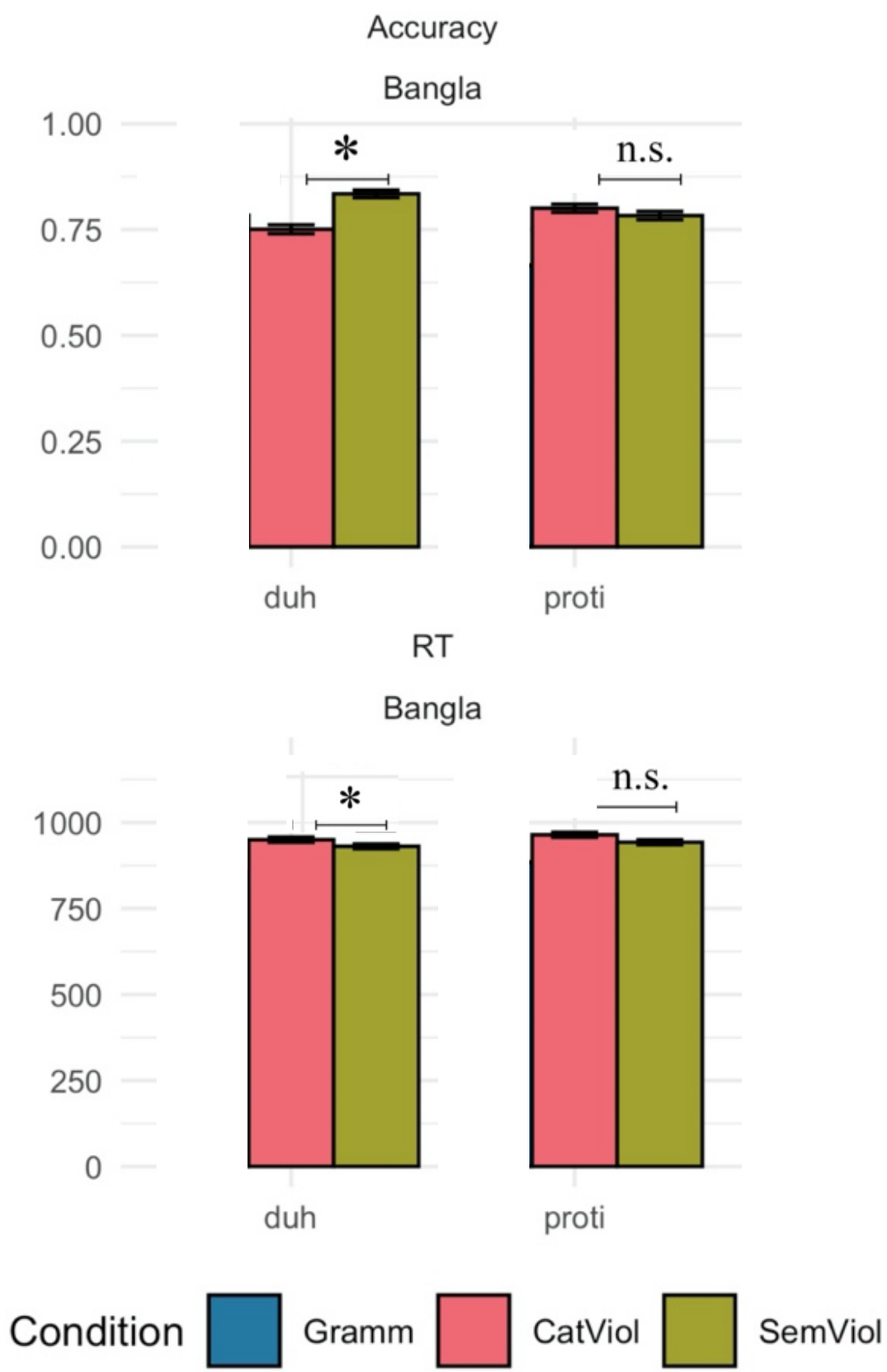
	Bangla	**Tagalog**	**Slovenian**
	2 prefixes	2 prefixes	3 prefixes
	abstract Ns	agentive Vs	‘unstable' Vs
Gramm	প্রতিহিংসা		
	prôti-hiṁsa	nag-ahit.	od-lepiti
	contra-violence	AG-shave	de-attach
	‘revenge’	‘shave'	'detach'
CatViol	প্রতিনীল		
	prôti-nil	nag-bintana	od-čebula
	contra-blue	AG-window	de-onion
SemViol	প্রতিরক্ত		
	prôti-rôktô	nag-bulok	od-čutiti
	contra-blood	AG-rot	de-feel

PRELIMINARY RESULTS: ONLINE BEHAVIOURAL EXPERIMENTS



- Replication of previous pattern (faster and more accurate rejections for CatViol than SemViol) for **Tagalog verbal Voice Inflection**, and **Slovenian event-subtype sensitive verbal derivation**

PRELIMINARY RESULTS: ONLINE BEHAVIOURAL EXPERIMENTS



- But not for **Bangla nominal derivation sensitive to the abstract/concrete distinction**
- Opposite direction pattern for duh-, ns for proti-
- Are nominals different?
- Is the abstract/concrete distinction qualitatively different?
- MEG data should help clarify

LEXI-VAULT

- Because the kinds of databases standardly used to select and norm lexical materials are not available for our languages, building our own databases is required.
- S. Wray leading work to develop **morphologically parsed databases from which critical variables such as word, morpheme, bigram and Part Of Speech (POS) frequencies can be extracted** (see github.com/s- wray/tag-lex and github.com/s-wray/jor-multimod)
- Developing a multi-language searchable lexical characteristics database Lexi-Vault to host our (and other searchable parsed corpora)

WRAPPING UP

- The SAVANT project is critical both to confirming the generalisability of the distinct neural profile for syntactic structure building vs. semantic composition (Neophytou et al. 2018, Stockall et al. 2019), and to uncovering and interpreting differences between languages. Because some of the languages we propose to study have, to the best of our knowledge, never been investigated using any neuroimaging tools before, the data sets we produce and discoveries we make will constitute a significant contribution to the general literature on the cognitive neuroscience of language, as well as to the specific questions addressed in this project.

THANK YOU

جامعة نيويورك أبوظبي
NYU | ABU DHABI

Queen Mary
University of London

Univerza v Ljubljani

SAVANT

SYSTEMATICITY AND VARIATION IN WORD STRUCTURE PROCESSING ACROSS LANGUAGES: A NEURO-TYPOLOGY APPROACH

DUSTIN ALFONSO CHACON (FOR … NAMES IN ORDER FROM X-PPL PROGRAM)

https://savant.qmul.ac.uk

UKRI Economic and Social Research Council

SCIENTIFIC ADVISORY BOARD

	Expertise	Language Expertise
Daniel Harbour, QMUL	Morphosemantics, Typology	Kiowa, Georgian, Hebrew
Hagit Borer, QMUL	Syntax, morphosyntax	Hebrew
Liina Pylkkänen, NYU	semantics, neurolinguistics (MEG)	Finnish
Artemis Alexiadou, Humboldt	syntax, heritage languages	Greek, German
Phaedra Royle, UdeM	morphology, neurolinguistics (EEG), acquisition, SLI	French, Spanish
Karsten Steinhauer, McGill	sentence processing, neurolinguistics (EEG), bilingualism	German
Isabelle Roy, Nantes	syntax, acquisition	French

DISCUSSION/FEEDBACK

PATHWAYS TO IMPACT

IMPACT PLAN

- multi-pronged approach, aimed at a lasting impact both within and outside of academia: (1) sharing results and tools with the existing research communities; (2) inspiring and training the next generation of native speaker researchers; and (3) embedding the knowledge, interest and resources required to perpetuate this training beyond the lifespan of the project in local schools and communities.

SHARING RESULTS AND TOOLS

- **An inter-disciplinary research and training workshop.**
- bring together our research team, our scientific advisory board, and researchers working on related topics from around the world.
- will include a Showcase evening where school-age researchers involved in our Schools collaboration will present the results of their research to the workshop attendees, and to their teachers, friends and families.
- will also include a 1 day Lexical Database-athon: an event in which we will guide interested researchers with only basic coding backgrounds through building and validating Lexical Databases for their languages

SCHOOLS PROJECT/NEXT GEN RESEARCHERS

- a programme of Research Training and Collaboration with local schools at both our QMUL and UvL sites
- training teachers and their students to develop the skills necessary to scientifically investigate the languages in their lives, inspiring some to pursue further study and careers in the cognitive neuroscience of language or in S.T.E.M. fields more broadly

SCHOOLS PROJECT

- Stage 1: Developing UG mentors
  - Advanced Experimental and Quantitative Tools Workshop (summer school) + additional Neuroling for UG students (in the summer between YR2&3)
  - Hire UG RAs (with grant budget)
  - QM Access funding to hire trained UG RAs as mentors/ ambassadors for…

SCHOOLS PROJECT

- Stage 2:
  - (1) Co-Is, with postdocs and UG student mentors will visit local secondary schools and colleges to give talks and demos about our research to students taking relevant courses
  - (2) we'll run an adapted version of the one week intensive summer school again, in the summer between YR3&4, this time targeted to the school/college students

SCHOOLS PROJECT

- Stage 2:
  - (1) Co-Is, with postdocs and UG student mentors will visit local secondary schools and colleges to give talks and demos about our research to students taking relevant courses
  - (2) we'll run an adapted version of the one week intensive summer school again, in the summer between YR3&4, this time targeted to the school/ college students (along the model of the successful Constructing a Language summer school)
  - (3) one-on-one practical training in designing, running and analysing an EEG experiment using our morphological wellformedness paradigm adapted to the language of the student's choice, offered by our postdocs and mentors, and (4) a showcase event for students to present their research to their peers, teachers, families and communities. The QMUL showcase event will run as part of the final project meeting in month 42. The UvL event will run shortly before that and it will also include a presentation in Hiša Experimentov (The House of Experiments), a venue in central Ljubljana dedicated to bringing science close to children and their families.

EMBEDDING KNOWLEDGE, INTEREST AND RESOURCES TO PERPETUATE TRAINING:

- We will work collaboratively with the teachers who welcome us into their classrooms to begin developing **practical materials for lessons**, following the established model for KE developed in the QMUL Linguistics Constructing a Language and Teach Real English projects. Crucially, these materials will reflect and celebrate linguistic diversity and highlight the important contributions to scientific knowledge that researchers from linguistically and socially diverse backgrounds are able to make. Work on these materials is expected to continue past the end of this project.

STILL TO DO

- recruit Impact/KE Advisory Board
- get letters of support from local schools

DISCUSSION/FEEDBACK