On the temporal and structural organization of phonological Information

François PELLEGRINO
Dynamics of Language Lab & ASLAN Laboratory of Excellence
University of Lyon & CNRS
Francois.Pellegrino@univ-lyon2.fr

Munich – February 5 2020

OUTLINE

- General Background
- Structural organization of phonological information
- Temporal regulation of phonological information
- Conclusion

ASLAN UNIVERSITE DE LYON DDL

General Background

- Languages
  - ✓ Are not immune to influences from this environment
  - ✓ Are prone to evolve under pressures from this environment

Languages are embedded (and embodied) in their environment

Social

Biological

Physical/Ecological

This Language/Environment interaction constitutes an **ecological niche**
(Sinha, 2009; Laland et al., 2010)

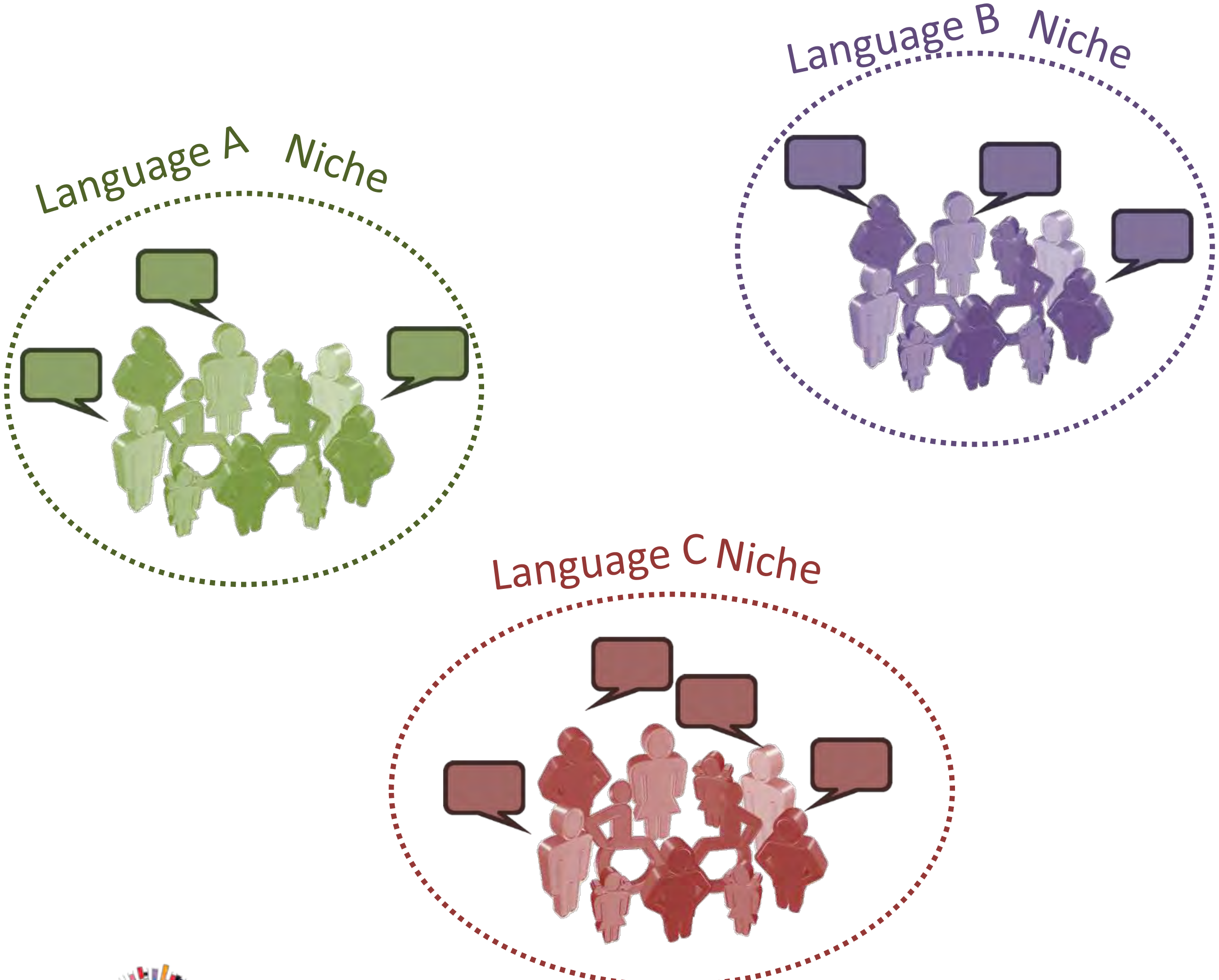
ASLAN UNIVERSITE DE LYON DDL

THE LANGUAGE NICHE IN A NUTSHELL

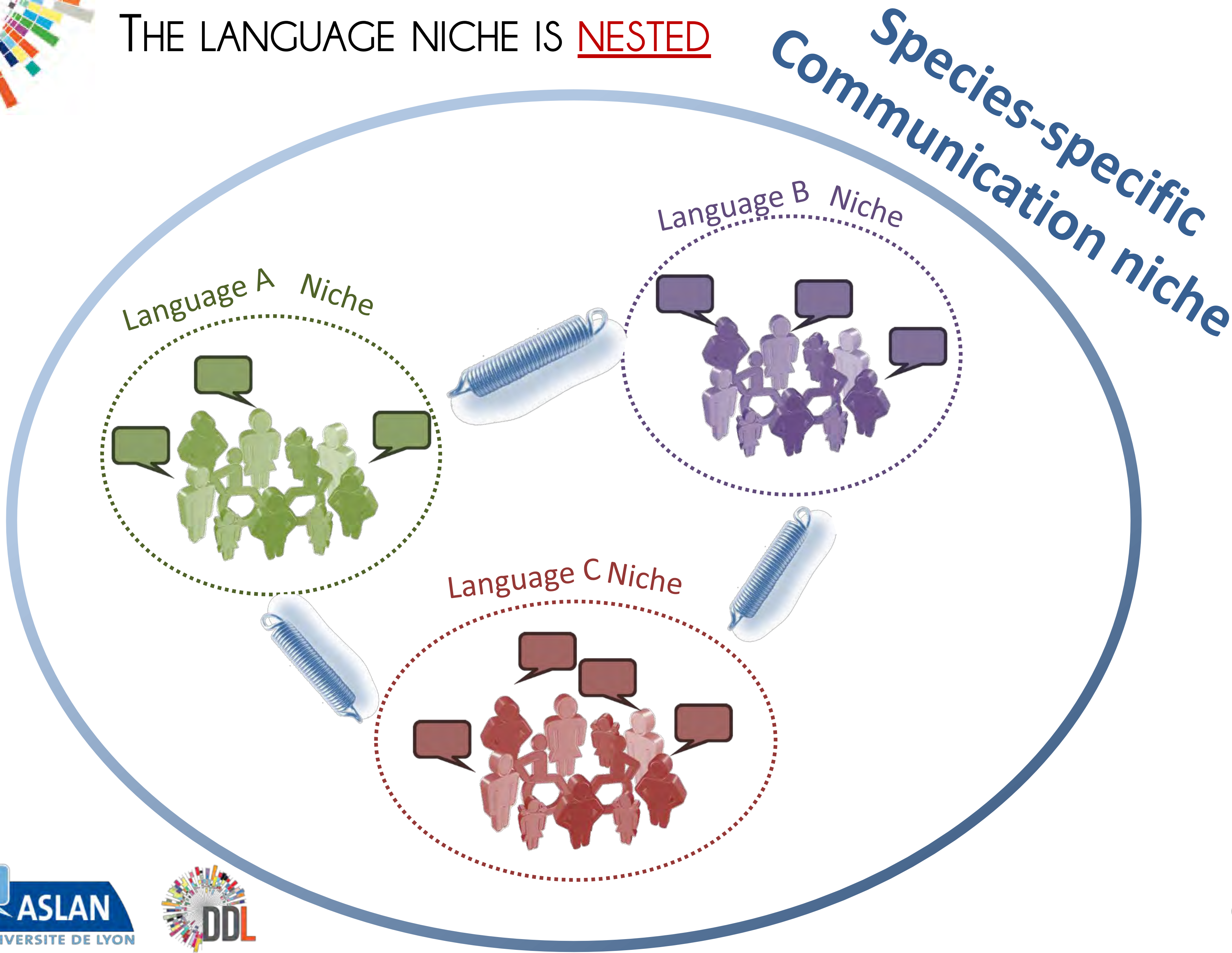
- Co-evolution of languages and their environment
- Environment (in the broad sense) provides numerous factors that may affect languages
  - ✓ Some of them are weak (but present other numerous generations of speakers)
  - ✓ Others are strong but localized (e.g. deafness)
- Potential causes of linguistic differentiation and diversity among **community-specific** language niches
- … but this is only a part of the story

ASLAN UNIVERSITE DE LYON DDL

THE LANGUAGE NICHE

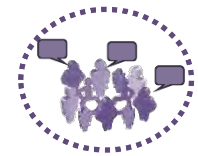


THE LANGUAGE NICHE IS NESTED

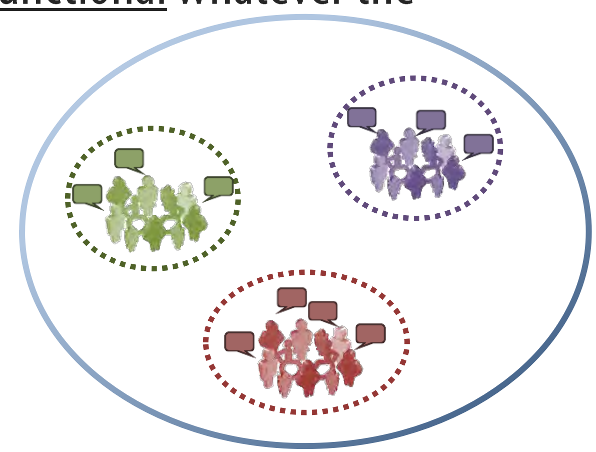


The language niche is <u>NESTED</u>

- Scale 1: Human groups (**Community-specific** language niche)
  - ✓ Languages adapt to specific environments (local fitness)
  - ✓ Language usage and structure impact each other



- Scale 2: Humankind (**Species-specific** communication niche)
  - ✓ Speech communication is the most pervasive mode of communication in our species
  - ✓ Human language is <u>ubiquitous</u> and <u>highly diverse</u>, and <u>functional</u> whatever the environment (global fitness)



- One of these functions is to <u>convey information</u>

The language niche is <u>nested</u>

- Illustration: Solomon Islands (data from Maddieson *et al.*, 2013)
  - ✓ Rotokas (Papuan, East; Bougainville Island), **6** consonants and **5** vowels
  - ✓ Yelî Dnye (Papuan, East; Rossel Island), **34** vowels and **58** consonants
    - ➔ Differences in information encoding (in a Shannonian framework)
    - ➔ Both languages fit the "human communication system" niche

- Main questions addressed in this presentation
  - ✓ How information distributes within a sound system (phonological repertoire)?
    - Study 1: Structural organization of phonological information
  - ✓ Can we define the human communication niche in terms of information transfer?
    - Study 2: Temporal regulation of phonological Information

ASLAN UNIVERSITE DE LYON
DDL

Study 1:
Structural organization of phonological information

‘The function of a phonemic system
is to keep the utterances of a language apart.

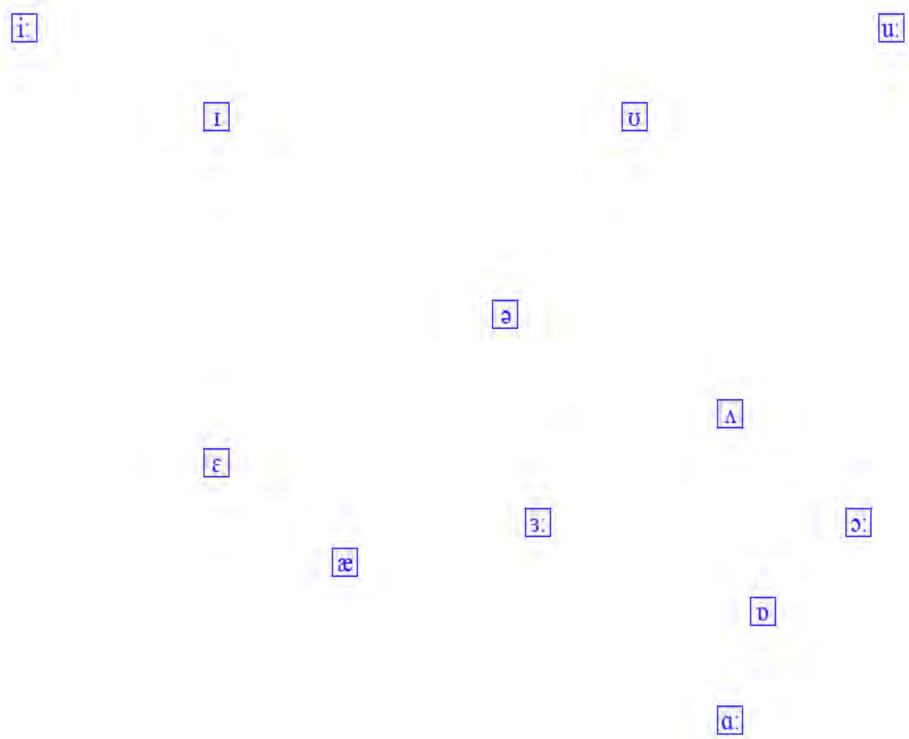
Some contrasts between the phonemes in a system apparently
do more of this job than others.’

Charles F. Hockett (1966)

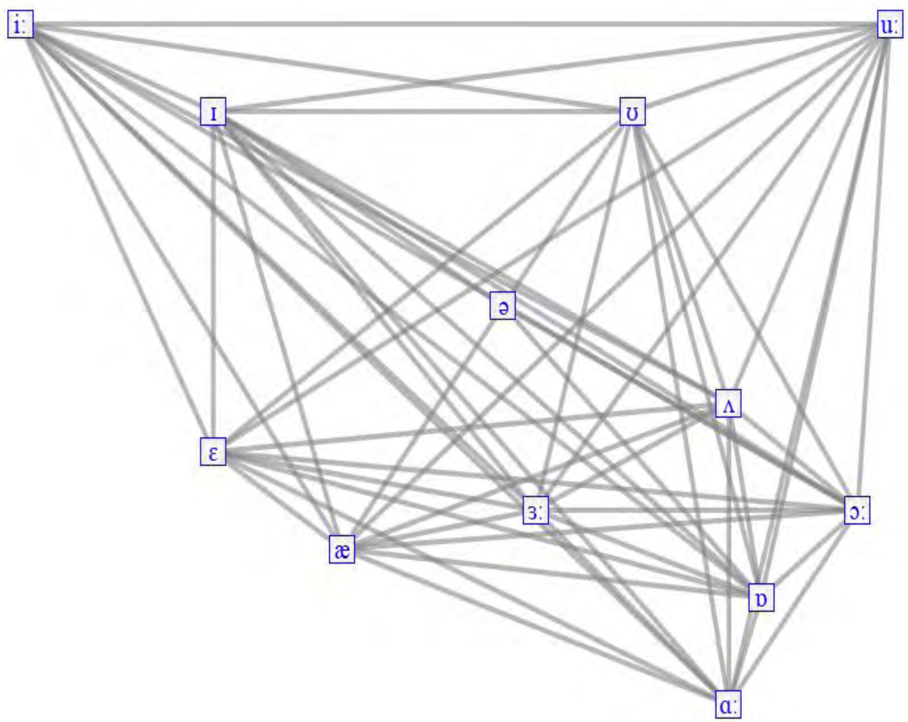
ASLAN
UNIVERSITE DE LYON

DDL

Vowel system as an (organized) set of vowel segments



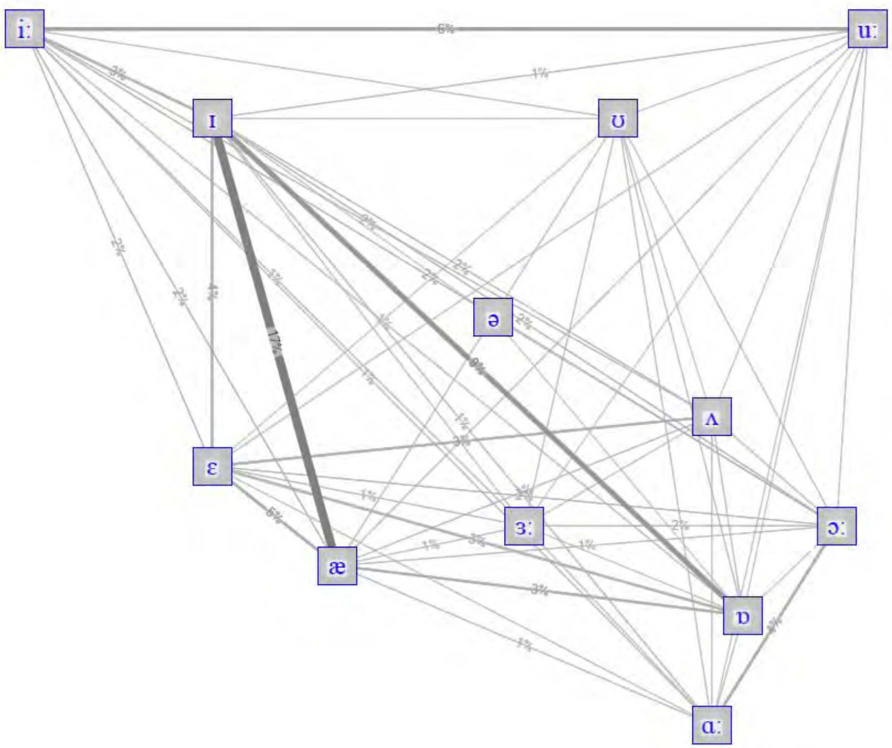
Vowel contrasts as a fully-connected graph



Equal Thickness = equal amount of the job done by each contrast

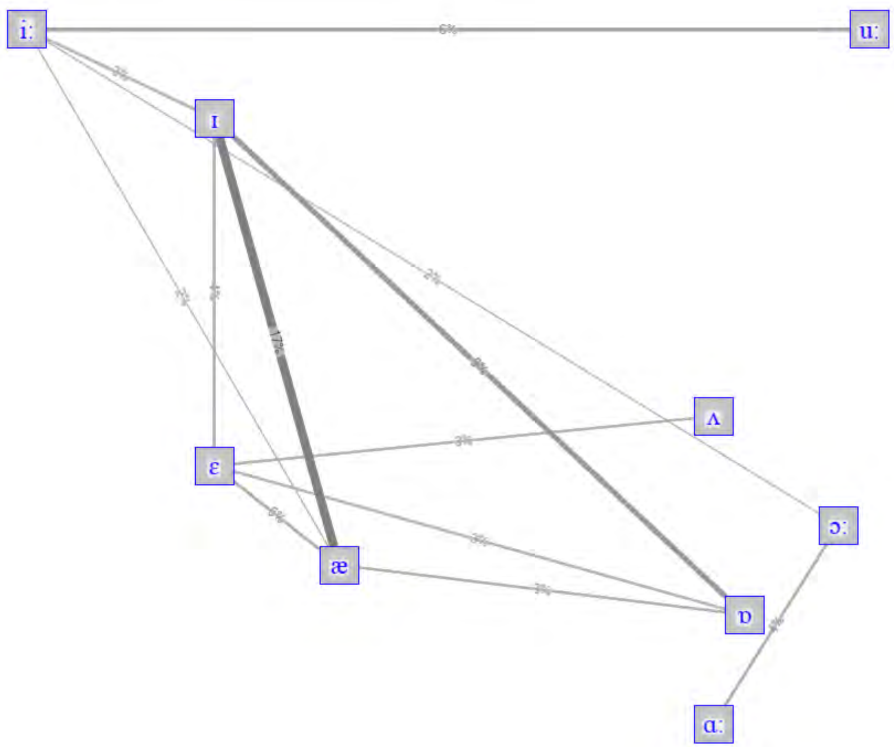
ASLAN UNIVERSITE DE LYON DDL

- But some contrasts actually do more of the job…



Edge thickness illustrates relative amount of the job done by each contrast

- Some segments are involved in none of the major contrasts… Are they nevertheless useful?



Less “important” contrasts and vowels erased

Research questions

- How is (lexical) information distributed in phonological systems?
- Could it change our view on phonological systems in a typological perspective ?
- This study:
  - ✓ Cross-linguistic and information-theory based perspective (Functional Load)
  - ✓ Multi-scale approach: Features, Segments, (Syllables), Phonological Subsystems

ASLAN UNIVERSITE DE LYON
DDL

Overview

- The notion of Functional Load
- Methodology & Data
- Results
- Conclusion

THE NOTION OF FUNCTIONAL LOAD (FL)

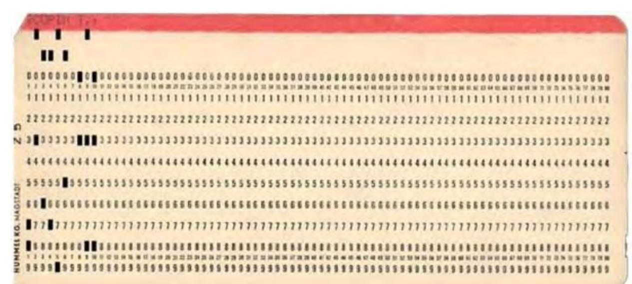
The origins

- Cercle Linguistique de Prague
  - ✓ "Rendement fonctionnel: Degré d'utilisation d'une opposition phonologique pour la différenciation des diverses significations des mots dans une langue donnée". (*TCLP 4, 1931*)
- Trubetzkoy (1939)
  - ✓ 'it is also possible to determine (. . .) the extent to which the individual phonological oppositions are utilized distinctively (their functional load) (. . .). It develops that there are "economical" and "wasteful" languages in this respect (. . .).' (Trubeztkoy, Principles, 1969:268).
- Martinet (since 1933)
  - ✓ Link between *Functional Load* and *Sound change*

ASLAN UNIVERSITE DE LYON DDL

The origins (cont'd)

- Comments on these seminal works
  - ✓ Strong intuitions, but lack of data and of mathematical concepts to test them
  - ✓ The diffusion of Shannon's Communication theory (aka Information theory) will provide conceptual tools to go beyond mere intuitions
- Hockett' methodological proposal (1955, 1961/1966)
  - ✓ FL = loss of entropy under the hypothesis of a phoneme coalescence (But no implementation and assessment with real data)
- A few computational and corpus-based attempts… and an eclipse
  - ✓ Kučera (1963), King (1967), Wang (1967)



ASLAN UNIVERSITE DE LYON

DDL

A new century, a new dawn for FL…

- Initiative: D. Surendran and colleagues
  - ✓ Measuring the Usefulness (Functional Load) of Phonological Contrasts. *Technical Report* TR-2003-12., Department of Computer Science, University of Chicago. 2003.
  - ✓ Articulatory complexity, ambient frequency and functional load as predictors of consonant development in children, *Journal of Speech and Hearing Research* 48(3). 2005.
- Van Severen, Gillis, Molemans, Van Den Berg, De Maeyer, & Gillis (2013)
  - ✓ The relation between order of acquisition, segmental frequency and function: the case of word-initial consonants in Dutch. *Journal of Child Language*, 40(4)
- Wedel, Kaplan, & Jackson (2013)
  - ✓ High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, 128(2).
- Our group

ASLAN UNIVERSITE DE LYON DDL

METHODOLOGY & DATA

Methodology – Notion of entropy

- Mathematical theory of communication (Shannon, 1948)
  - ✓ A theory of communication (= information transmission)
  - ✓ Quantification of information, entropy, channel capacity and redundancy

- Considering that language *L* is a source of linguistic sequences composed of units (*w*) from a finite set (N_L)
- Entropy *H*(*L*) = Average quantity of information per unit

$$H(L) = -\sum_{i=1}^{N_L} p_{w_i} \log_2\left(p_{w_i}\right)$$

  - ✓ Easy to estimate from the set of units and their probabilities
  - ✓ Probabilities p_{w_i} estimated by their frequency in a relevant corpus
  - ✓ Units may be words, syllables, phonemes, etc.
  - ✓ *More elaborated formulas to take contextual information into account (see Study #2)*

ASLAN UNIVERSITE DE LYON DDL

Methodology – FL estimation

Quantitative entropy-based definition of FL

- ✓ Following Hockett (1966) & Carter (1967)
- ✓ Language *L* considered as a source of sequences of independent words w_i taken from a set N_L
- ✓ FL of a phonological contrast x~y = **quantification of the perturbation induced by merging x and y** in terms of increase of homophony and of changes in the distribution of word frequencies
- ✓ *FL*(*x*,*y*) = relative difference in entropy between the observed state *L* and a fictive state L^*_{xy} in which the contrast is neutralized

$$FL(x,y) = \frac{H(L) - H(L^*_{xy})}{H(L)}$$

Methodology – Toy Language

Observed Lexicon

Form	Frequency
pal	300
pil	200
bal	150
bil	150
pul	100
bul	100
TOTAL	1000

Inventory: /a i u p b l/

$N_L = 6 \quad H(L) = 2.47$

Contrast /a-i/

Form	Frequency
p¤l	300
p¤l	200
b¤l	150
b¤l	150
pul	100
bul	100
TOTAL	1000

Fictive Lexicon

Form	Frequency
p¤l	500
b¤l	300
pul	100
bul	100
TOTAL	1000

$H(L^*_{ai}) = 1.69$

FL(a-i) = (2.47-1.69)/2.47 = **31.8 %**

FL(a-u) = **23.1 %**

FL(i-u) = **21.0 %**

Phoneme /a/

$$FL(x) = \frac{1}{2}\sum_{y} FL(x,y)$$

FL_V=61%

FL(a) = ½ (FL(a-i)+FL(a-u)) = ½ (31.8+23.1)=27.45 %

Material

Language	ISO 639-3 Code	Source
Cantonese	YUE	A linguistic corpus of mid-20th century Hong Kong Cantonese (Research Centre on Linguistics and Language Information Sciences, 2013)
English	ENG	WebCelex (Max Planck Institute for Psycholinguistics, 2013, 2014)
Japanese	JPN	The corpus of spontaneous Japanese (NINJAL, 2011)
Korean	KOR	(Leipzig corpora collection)
Mandarin	CMN	Chinese Internet Corpus (Sharoff et al, 2006)
German	DEU	WebCelex (Max Planck Institute for Psycholinguistics, 2013, 2014)
Swahili	SWH	Gelas, Besacier, & Pellegrino, (2012)
Italian	ITA	PAISÀ Corpus (Lyding et al., 2014)
French	FRA	Lexique 3.80 (New et al., 2001)

20,000 most frequent words (inflected forms) considered, except for Cantonese (5,000 forms) & Italian (15,788 forms)

ASLAN UNIVERSITE DE LYON
DDL

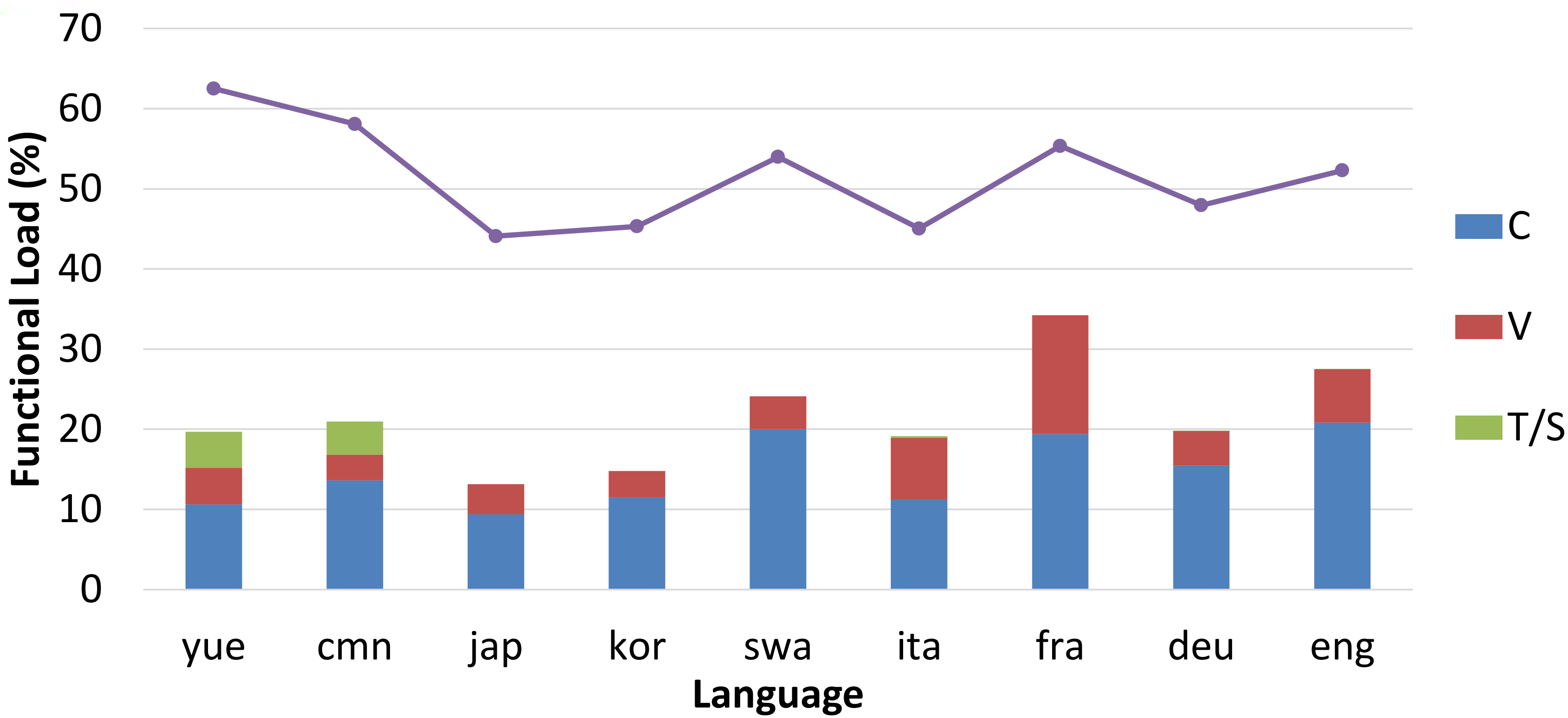
RESULTS

Subsystems

Segments

Features

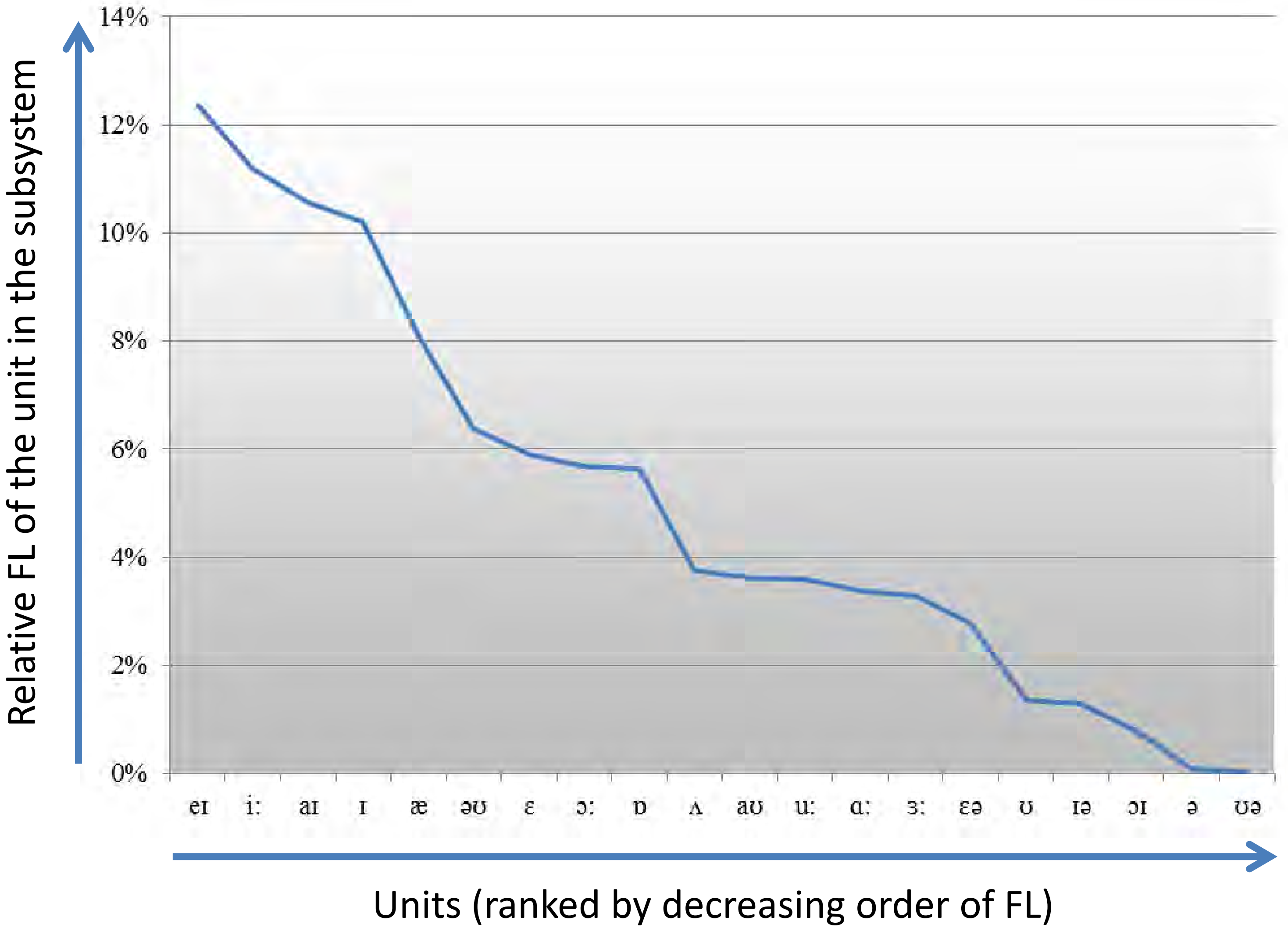
Coarse grain: vowel, consonant, and tone systems



- Variation in phonological FL
- In Mandarin and Cantonese, $FL_V \approx FL_T$
- High FL of vowels in French (and Italian)

ASLAN UNIVERSITE DE LYON DDL

FINER GRAIN : HOW TO READ A RESULT GRAPH



ASLAN UNIVERSITE DE LYON

DDL

Most "natural" scale of organization: segments

Language	ISO 639-3 Code	Phonological system size	
Cantonese	YUE	V C T	13 19 6
English	ENG	V C S	22 28 2
Japanese	JPN	V C	10 16
Korean	KOR	V C	8 22
Mandarin	CMN	V C T	7 25 5
German	DEU	V C S	22 24 1
Swahili	SWH	V C	5 30
Italian	ITA	V C S	8 25 1
French	FRA	V C	15 21

Oh, Coupé, Marsico, & Pellegrino (2015). *J. Pho.*

⌧ Methodological details

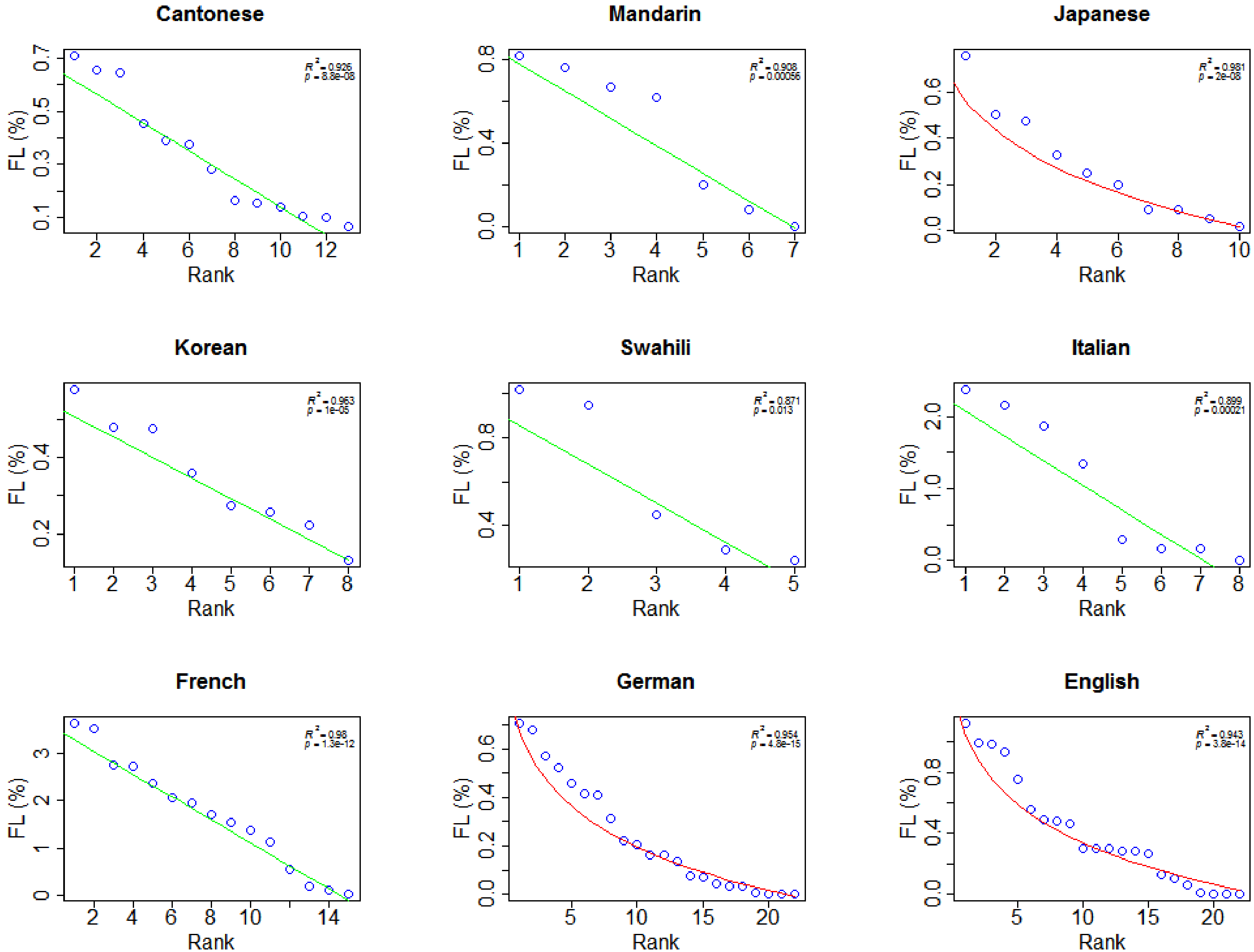
⌧ Discussion

- Lemmas vs inflected wordforms
- Types vs. Tokens
- Consonantal bias hypothesis (Nespor, Peña, & Mehler, 2003)

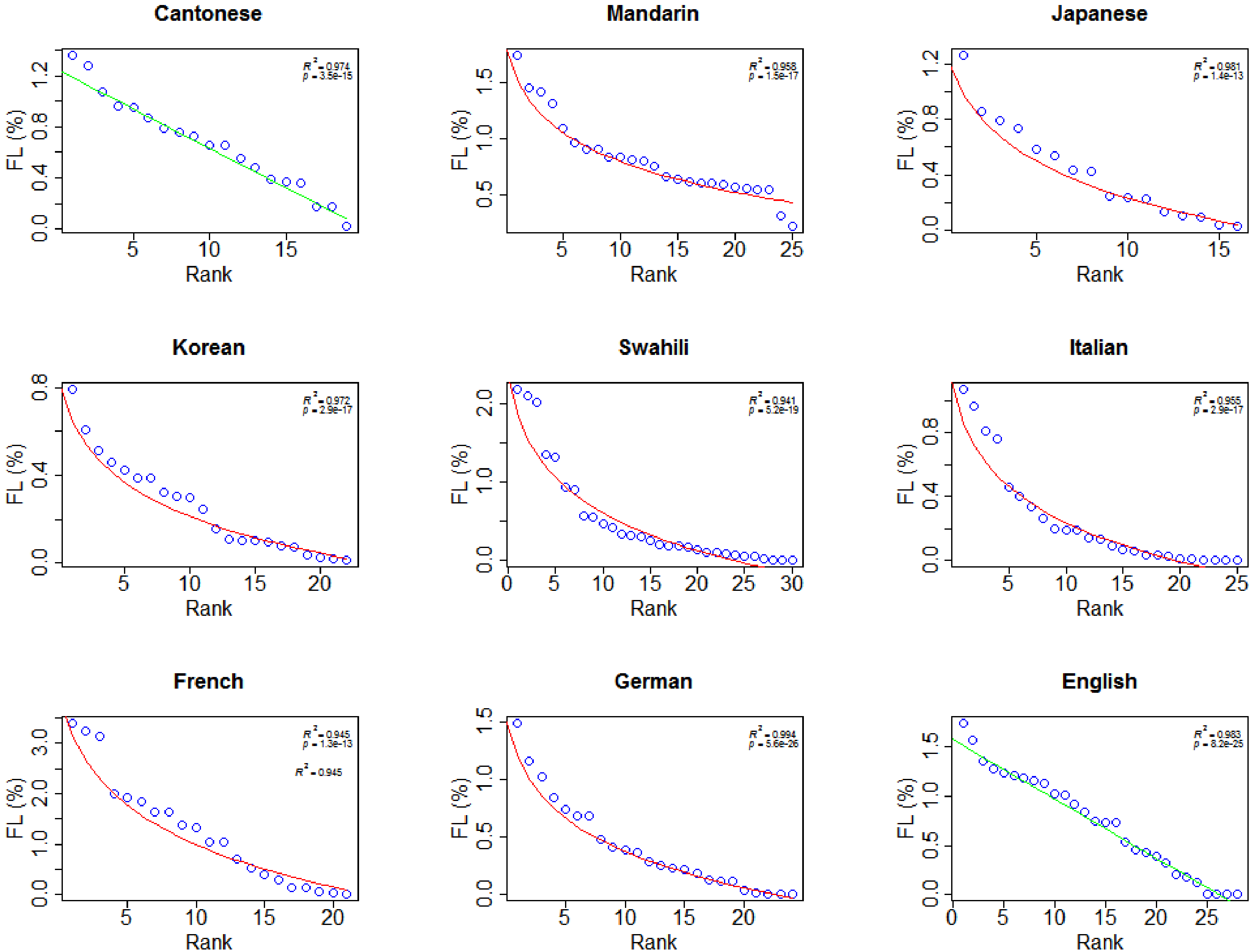
ASLAN UNIVERSITE DE LYON

DDL

Vowel FL


Cantonese
R^2 = 0.926
p = 8.8e-08
Mandarin
R^2 = 0.908
p = 0.00056
Japanese
R^2 = 0.981
p = 2e-08
Korean
R^2 = 0.963
p = 1e-05
Swahili
R^2 = 0.871
p = 0.013
Italian
R^2 = 0.899
p = 0.00021
French
R^2 = 0.98
p = 1.3e-12
German
R^2 = 0.954
p = 4.8e-15
English
R^2 = 0.943
p = 3.8e-14
FL (%)
Rank

Consonant FL



FEATURE SCALE

Language	ISO 639-3 Code	Phonological system		Number of features
Cantonese	YUE	V C	13 19	12 18
English	ENG	V C	22 28	27 19
Japanese	JPN	V C	10 16	10 15
Korean	KOR	V C	8 22	10 17
Mandarin	CMN	V C	7 25	11 19
German	DEU	V C	22 24	21 18
Swahili	SWH	V C	5 30	9 19
Italian	ITA	V C	8 25	10 18
French	FRA	V C	15 21	12 17

E.g.

/i/: high front unrounded

/p/: bilabial voiceless stop

⌦ (Mostly articulatory) description of segments in terms of features based on UPSID (Maddieson & Precoda, 1990) et revised in LAPSyD (Maddieson *et al.*, 2011)

ASLAN UNIVERSITE DE LYON

DDL

FROM FEATURES TO ARTICULATORY DIMENSIONS

i ɯ u

e o

ɛ ʌ

a

Vocalic inventory of Korean

Aperture: 3 sets of mergers

ɛ, e, i (front unrounded)
o, u (back rounded)
ʌ, ɯ (back unrounded)
a

Anteriority: 2 set of mergers

ɛ, ʌ (lower-mid unrounded)
i, ɯ (high unrounded)
e
o
a

E.g. To compute the FL of **aperture**, the actual lexicon is contrasted with a lexicon where 3 simultaneous sets of mergers create homophony and modify the distribution of word frequencies.

ASLAN UNIVERSITE DE LYON DDL

Feature FL: a closer look at dimensions

	V					C			
	Aperture	*Anteriority*	*Roundedness*	*Length*	*Nasalization*	*Place*	*Manner*	*Voicing*	*Aspiration*
yue	0,23	0,15	0,11			1,72	1,07		0,33
cmn	1,02	0,06	0,25			2,07	1,04	0,18	0,67
jap	0,07	0,14		0,21		0,79	0,26	0,36	
kor	0,53	0,08	0,02			0,67	1,01	0,05	0,01
swa	0,26					1,84	1,66	0,13	
ita	1,62					0,22	2,03	0,12	
fra	2,66	0,66	1,50		0,16	1,56	3,04	0,89	
deu	0,31	0,11	0,06	0,03		0,91	2,72	0,11	
eng	1,19	0,11				0,99	2,45	0,59	

Regarding vowels and primary articulatory dimensions, **aperture** carries the heaviest load in 8 of the 9 languages.
Secondary features can also have a high / the highest FL.

Regarding consonants, languages seem to choose **either place or manner** as the primary way to differentiate between words. **Voicing** always comes after except in Japanese.

Study 1: Discussion

- Cross-linguistic trends
  - ✓ Whatever the organizational scale, FL is not evenly distributed
  - ✓ Distribution of heavy vs. light load units may be more (for Vs) or less (for Cs) linear
  - ✓ ~50% of the lexical distinctions rely on infra-syllabic components (nice balance between localized short-term and longer term information)
  - ✓ Importance of coronal consonants and low vowels [not shown here]
- Cross-linguistic diversity
  - ✓ Languages differ in their heavy-loaded units [not shown here]
- But…
  - ✓ Is this entropy-based definition of FL relevant?
  - ✓ How about other methodologies?
    - Absolute number of minimal pairs (Wedel, *et al.*, 2013)
    - Relative number of minimal pairs w.r.t. the expected ones (Martin & Peperkamp, 2017)

Study 1: Speculative Conclusion

- Strong tendency toward an uneven distribution of FL
- Uneven distributions often found in languages (Zipf's law, etc.)
  - ✓ Structure of the lexicon and morphology (preferential binding, etc.)
    - Distinctiveness vs. efficiency (Kello & Beltz, 2009)
    - Notion of kernel word network (Ferrer i Cancho & Solé, 2001; Dorogovtsev & Mendes, 2001)
  - ✓ (Maximum) re-use of phonological 'chunks' ?
- Existence of a kernel phonemic network?
  - ✓ Core heavy-load phonemes and contrasts vs. others (more peripheral)
- In this view the latter are not useless
  - ✓ Probably reflect the adaptive nature of the language
  - ✓ Probably useful in terms of information distribution for cognitive processing

Study 2:
Temporal regulation of phonological information

CAVEAT

- About Average information rate. Loosely related to
  - ✓ Local information modulation hypotheses : Uniform Information Density (Jaeger, 2010) and Smooth Signal Redundancy (Aylett & Turk, 2004)
  - ✓ Within-language variation (Cohen Priva, 2017)

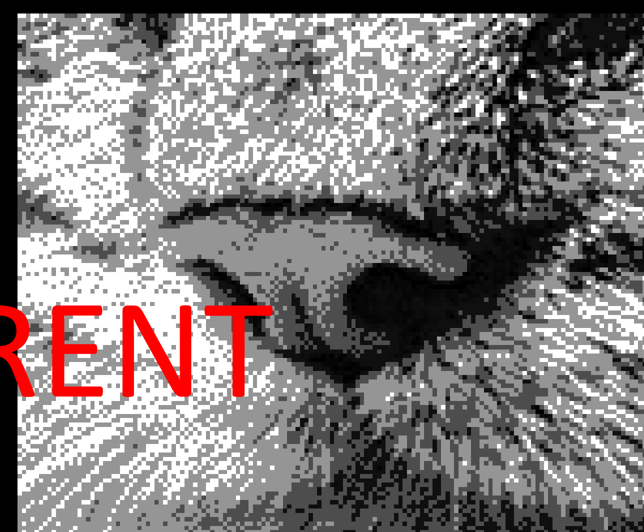
- Definition of information based on Shannon's theory
  - ✓ A quantitative approach to information encoding
  - ✓ Only loosely connected to semantics and meaning

SEMANTICS

SAME

ENCODING

• Resolutions (x2)
• #Grey levels (256 vs. 16)

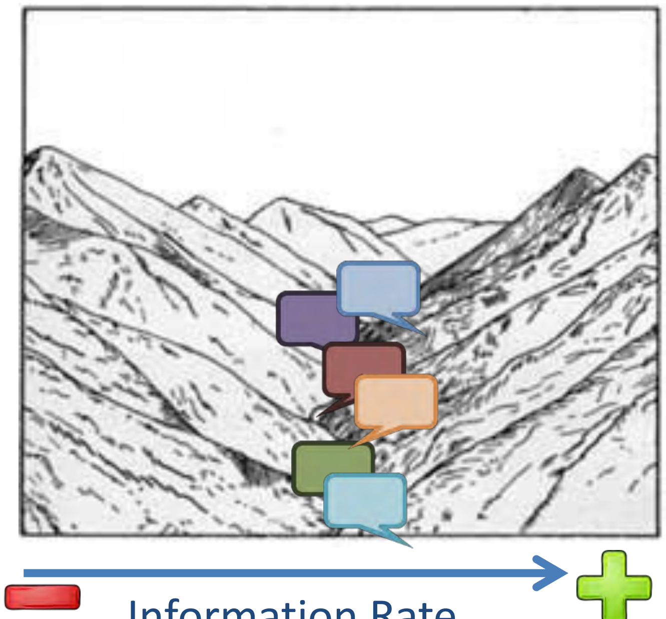


DIFFERENT

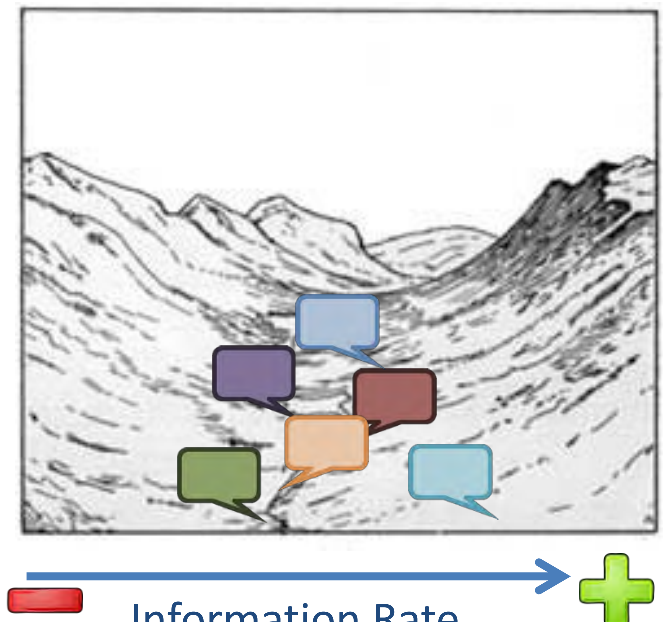
Research Question

- How Average Information Rates vary across languages?
- What does the information rate landscape look like?

V-Shape (strongly constrained)



U-Shape (weakly constrained)



Flat (unconstrained)

ASLAN UNIVERSITE DE LYON

DDL

AVERAGE INFORMATION RATE: OUR DEFINITIONS

- Information rate defined as :
  Average Information per syllable x Average Speech Rate

- Speech Rate (in syllables/sec)
  - ✓ Estimated from a parallel speech corpus recorded in several languages
- Information per syllable

"[T]his word information in communication theory relates not so much to what you do say, as to what you could say. That is, information is a measure of one's freedom of choice when one selects a message" (Weaver, in Shannon & Weaver, 1949; emphasis added)

- ✓ **Paradigmatic approach** (in bits/sec): Syllabic Entropy ➔ ***WHAT YOU COULD SAY*** (estimated from large written corpora or lexicons)
- ✓ **Syntagmatic approach** (normalized, unitless): Syllabic Information Density w.r.t. a reference language (same corpus as for Speech Rate) ➔ ***WHAT YOU DO SAY***

ASLAN UNIVERSITE DE LYON DDL

DATA

- Languages ($N_L = 17$)
  - ✓ **Basque** EUS, **British English** ENG, **Cantonese** YUE, **Catalan** CAT, **Finnish** FIN, **French** FRA, **German** DEU, **Hungarian** HUN, **Italian** ITA, **Japanese** JAP, **Korean** KOR, **Mandarin Chinese** CMN, **Serbian** SRP, **Spanish** SPA, **Thai** THA, **Turkish** TUR, **Vietnamese** VIE
- Parallel Speech Corpus (Speech Rate)
  - ✓ 17 languages
  - ✓ Postulate: on average, the semantic content is similar across languages
  - ✓ 15 texts (~3'25 per speaker)
  - ✓ 10 speakers (5 females) per text, recorded with ROCMe! (Ferragne et al. ,2012)
- Text Corpus (Paradigmatic Information)
  - ✓ Large corpora (from 130k tokens in Cantonese to almost 1G tokens in Spanish)
  - ✓ Various Sources (Celex, Lexique, Leipzig Corpora Collection, etc.)
  - ✓ Semi-Automatic Syllabification

ASLAN UNIVERSITE DE LYON DDL

Main Parameters (for each language *L*)

- ***SR_L*** Syllabic Speech Rate
  - ✓ Average number of syllables per second
    - Number of Phonological Syllables (from canonical transcription) per second
    - Articulation Rate (pauses are discarded)
- Syllabic Entropy (Shannon's communication theory)
  - ✓ Average amount of information carried by each syllable
  - ✓ Paradigmatic dimension (*what could be said*)
- ***ID_L*** Syllabic Information Density
  - ✓ Syntagmatic dimension (*what is said*)
  - ✓ Number of syllables compared to Vietnamese (taken as reference)

ASLAN UNIVERSITE DE LYON DDL

INFORMATION DENSITY (SYLLABIC ENTROPY)

- Language *L*
  - ✓ Source of linguistic sequences of syllables (σ) drawn from a finite set Σ of N_L distinct syllables
  - ✓ Syllable probabilities estimated from a large corpus
- Shannon Entropy (= average information per syllable)

$$H_L \triangleq -\sum_{i=1}^{N_L} p_{\sigma_i} \log_2\left(p_{\sigma_i}\right)$$

  - ✓ H_L is always inferior to $\log_2(N_L)$ (channel capacity H_{max})
  - ✓ $H_L = H_{max}$ iff the syllables are equiprobable (maximal uncertainty)
- But syllables are not independent. Context matters!!!
- Conditional Entropy
  - ✓ Average information per syllable, given the context

$$CE_L \triangleq -\sum_{c \in C} p(C) \sum_{i=1}^{N_L} p\left(\sigma_i | c\right) \log_2\left(p\left(\sigma_i | c\right)\right)$$

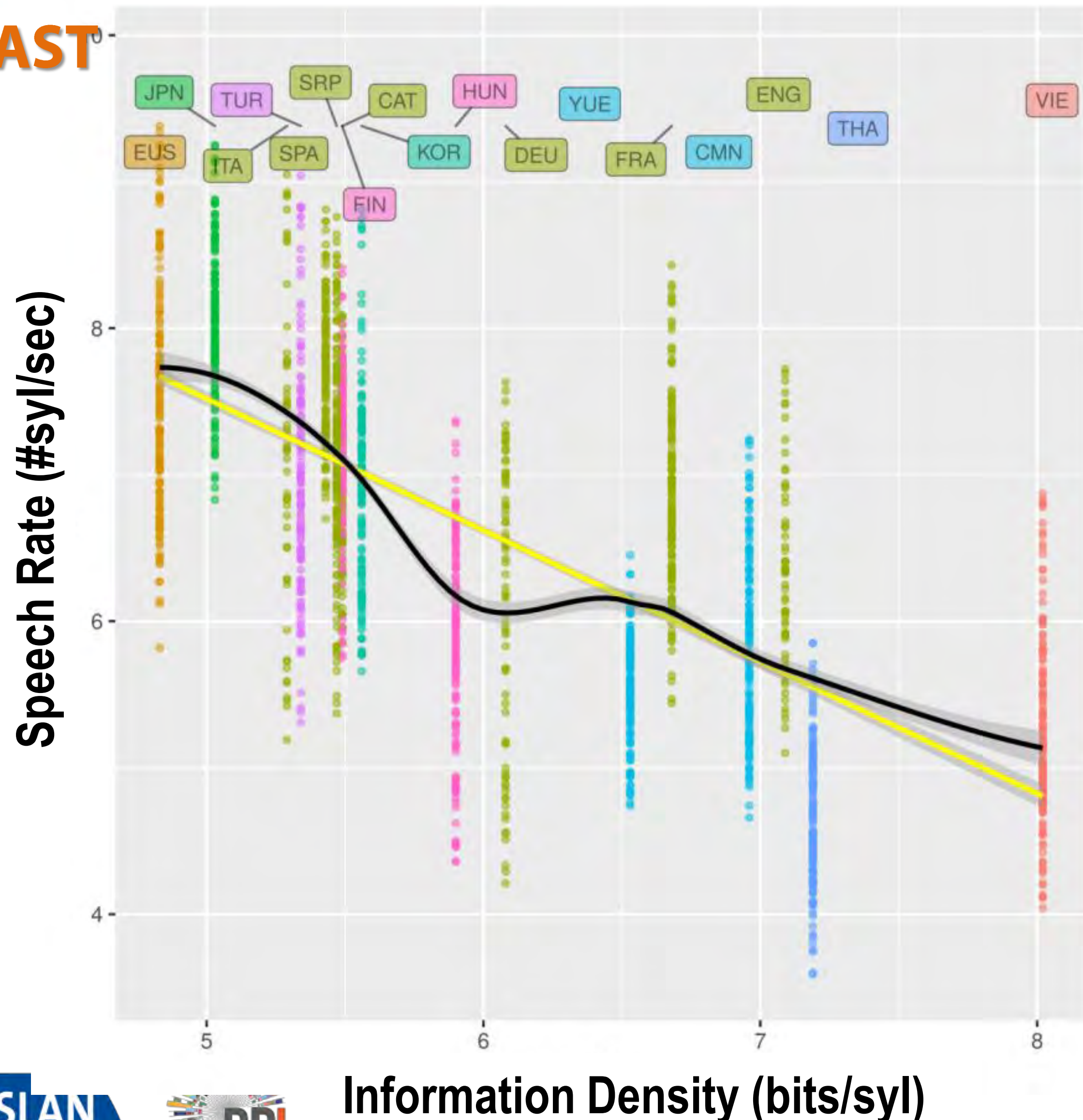
  - ✓ In this study, Context = previous syllable in the sequence (WITHIN WORD)

➔ $0 \leq$ Conditional Entropy $\leq$ Shannon Entropy $\leq \log_2(N_L)$

ASLAN UNIVERSITE DE LYON DDL

Result #1
A balance in information rate

FAST



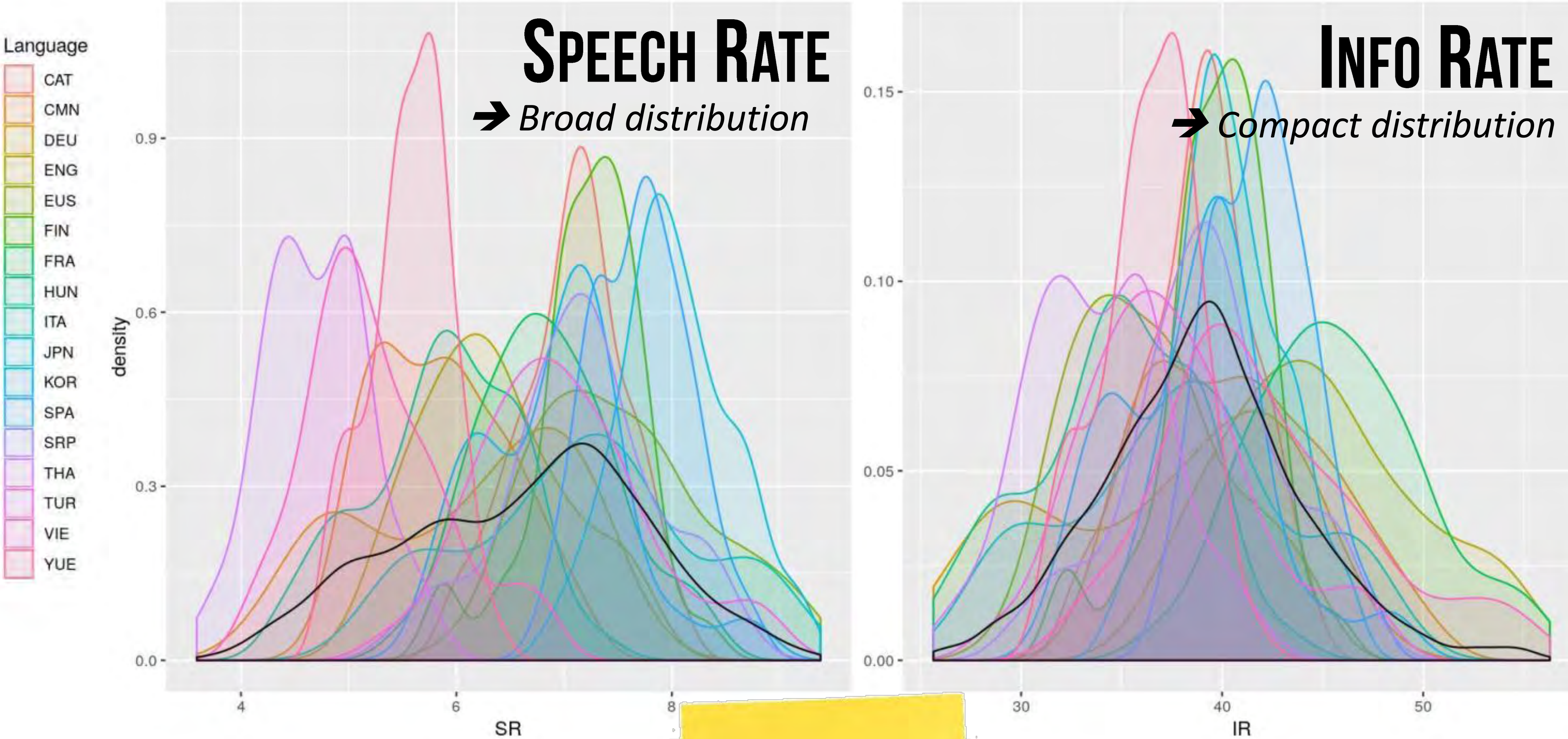
Existence of Different Strategies

DENSE

ASLAN UNIVERSITE DE LYON

DDL

Result #2
Speech Rate vs Information Rate Distributions

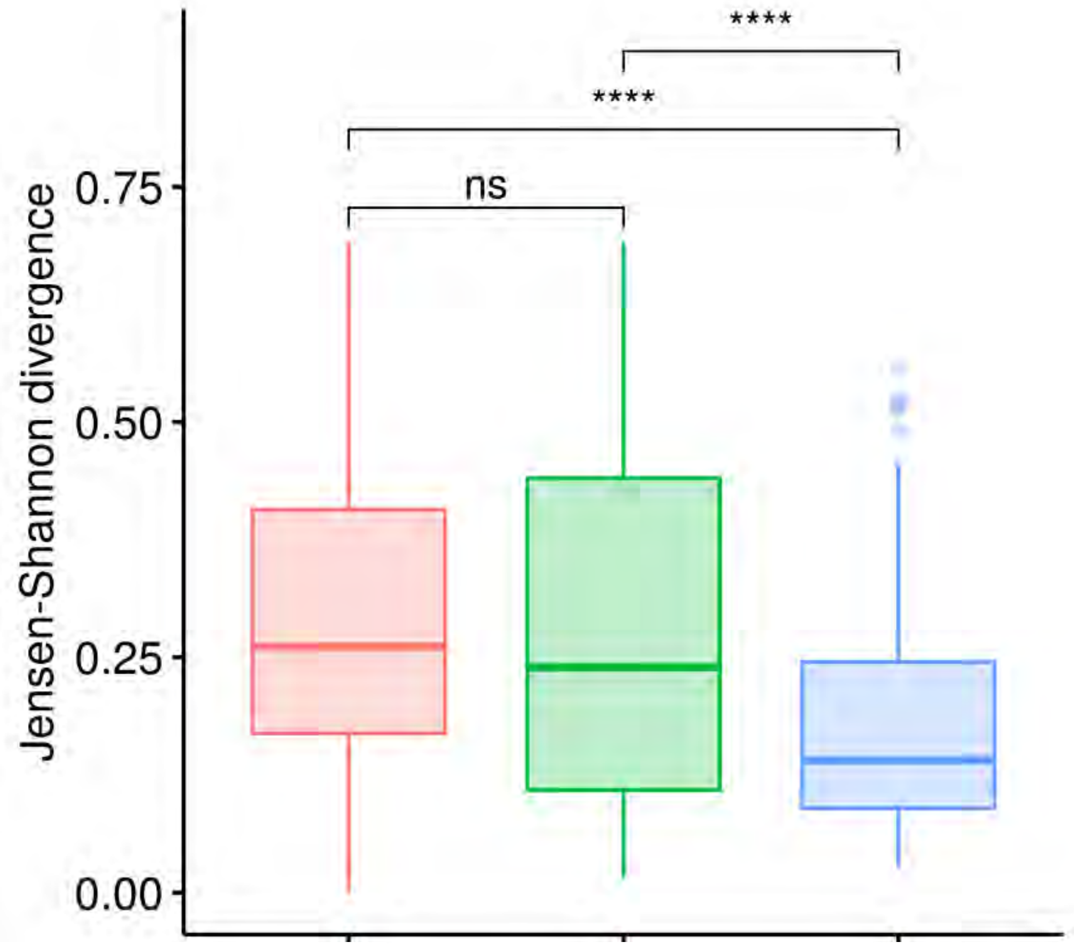


Existence of an attractor

ASLAN UNIVERSITE DE LYON

DDL

Pairwise divergences between languages



Syllable Info
(Social convention)

Speech Rate
(Individual behavior)

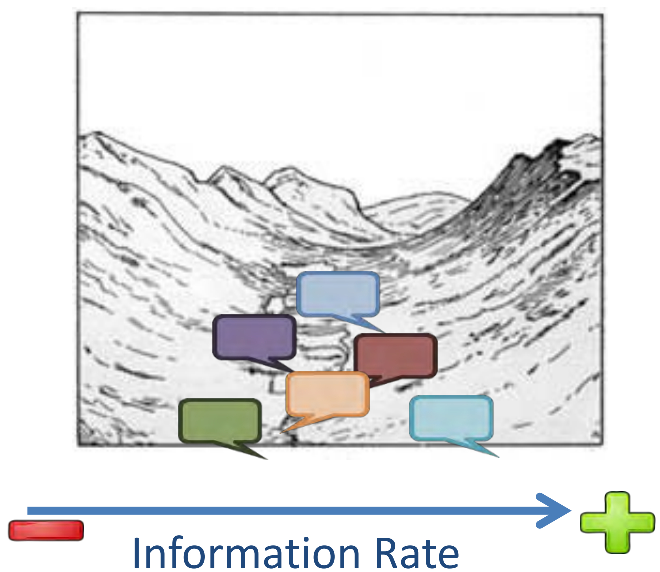
Information Rate
(Functional fitness)

ASLAN UNIVERSITE DE LYON

DDL

Study 2: Factual Conclusion

- Languages exhibit large variation
  - ✓ **Syllabic Speech Rate**
  - ✓ **Syllabic Information density**
- Average **Information Rates** tend to exhibit **<u>less</u>** cross-linguistic variation than average **Speech Rates**
  - ✓ Support the idea that average information rates are **(weakly) constrained** (U-shaped valley)
  - ✓ Information Rate as a better candidate than speech rate for cognitive universals?



Study 2: Speculative Conclusion

Open questions: Language Evolution

- Speculation: Individuals continuously monitor (consciously – or more likely not) and adapt their speech rate to the specific linguistic and communicative context.

- Prediction: when a language change drifts the information rate away from the optimal range, compensatory mechanisms that affect speech rate (e.g., coarticulation) may bring the average information rate back towards optimal regions.

General Conclusion

General Conclusion

- Information-oriented approaches shed some lateral light on the organization and usage of phonological systems
- Robust and multiscale trend towards uneven FL distribution within phonological systems (whatever the descriptive scale)
- Languages exhibit different strategies (Fast vs. Dense) in information encoding ; those differences tend to compensate
- Existence of an attractor in terms of Information Rate is likely

General Conclusion (cont'd)

- Function of distinct syllables (/phonemes/…) may be twofold
  - ✓ Carry information (high-FL units)
  - ✓ Provide a pseudo-rhythmic carrier facilitating neurocognitive information processing

- Cross-language comparisons are essential in research on brain oscillations

- More studies needed to connect the dots…

THANK YOU

ASLAN
UNIVERSITE DE LYON

DDL

Acknowledgement

- ✓ Yoon Mi OH (University of Canterbury, NZ)
- ✓ Egidio MARSICO (DDL – University of Lyon – CNRS)
- ✓ Christophe COUPÉ (Hong Kong University)
- ✓ Dan DEDIU (DDL – University of Lyon – CNRS)

Financial support

- ✓ Collegium of Lyon & EURIAS network
- ✓ Laboratory of Excellence ASLAN (ANR-10-LABX-0081 / ANR-11-IDEX-0007)

Presentation based on

- ✓ Coupé, C., Oh, Y.M., Dediu, D. & Pellegrino, F. (2019). Different languages, similar encoding efficiency: comparable information rates across the human communicative niche", *Science Advances*, Sept. 2019.
- ✓ Oh, Y. M. (2015). *Linguistic complexity and information: quantitative approaches* (Doctoral dissertation, Uni. Lyon, France).
- ✓ Oh, Y.M., Coupé, C., Marsico, E. & Pellegrino, F. (2015). "Bridging phonological system and lexicon: insights from a corpus study of functional load". *J. Pho*, 53.
- ✓ Oh, Y.M., Pellegrino, F., Coupé, C. & Marsico, E., (2013). "Cross-language comparison of functional load for vowels, consonants, and tones". *Interspeech 2013*.
- ✓ Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language*, 87(3).

ASLAN UNIVERSITE DE LYON

DDL

INFORMATION RATE & BRAIN OSCILLATIONS

Brain Oscillations in a nutshell

- Neuronal populations synchronize their activity
  - ✓ Between them for efficient neuronal communication between regions
  - ✓ With external rhythmic information to enhance information processing (entrainment)
  - ✓ Speech as a (quasi-)rhythmic signal
    - Temporal Regularities
    - Deviation from regularities
- Neural oscillations
  - ✓ Neuron oscillatory activity detected by electroencephalography (EEG) in cortex
  - ✓ Characterized by the energy in several frequency bands
    - Approximately: δ(<3Hz) θ(4-8 Hz) α(8-12 Hz) β(15-35 Hz) γ(25-40 Hz)
    - Several hypotheses w.r.t. the role of these oscillations (incl. information transfer…)

Brain Oscillations & Speech

- Growing literature relating speech to neural oscillations
  - ✓ Synchronization of theta oscillation with syllabic rate (Ghitza, 2011; Peelle & Davis, 2012)
  - ✓ Cascaded oscillators (Ghitza, 2011) and Theta-Gamma Nesting (Giraud & Poeppel, 2012)
    - Oscillations in the theta band (=syllabic scale) control gamma oscillations (= phonetic scale)
  - ✓ Open issues
    - Exogenous entrainment or endogenous synchronization? (see Meyer, et al., 2019)
    - Top-down vs. Bottom-up processes?
- Importance of cortical oscillations and theta-gamma coupling
  - ✓ "This phase modulation effectively encodes a prediction of when important events (. . .) are likely to occur, and acts to increase sensitivity to these relevant acoustic cues" (Peelle & Davis, 2012:1)

ASLAN UNIVERSITE DE LYON DDL

SPEECH RATE AND BRAIN OSCILLATIONS: ILLUSTRATION

- Ghitza and Greenberg (2009) ; Ghitza (2014)

Original Paper

Phonetica 2009;66:113–126
DOI: 10.1159/000208934

Received: September 10, 2008
Accepted: January 8, 2009

On the Possible Role of Brain Rhythms in Speech Perception: Intelligibility of Time-Compressed Speech with Periodic and Aperiodic Insertions of Silence

Oded Ghitza[a,b] Steven Greenberg[c]

[a]Sensimetrics Corp., Malden, Mass., [b]Boston University, Boston, Mass., and
[c]Steven Greenberg, Silicon Speech, Santa Venetia, Calif., USA

frontiers in PSYCHOLOGY

ORIGINAL RESEARCH ARTICLE
published: 04 July 2014
doi: 10.3389/fpsyg.2014.00652

Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech

Oded Ghitza *

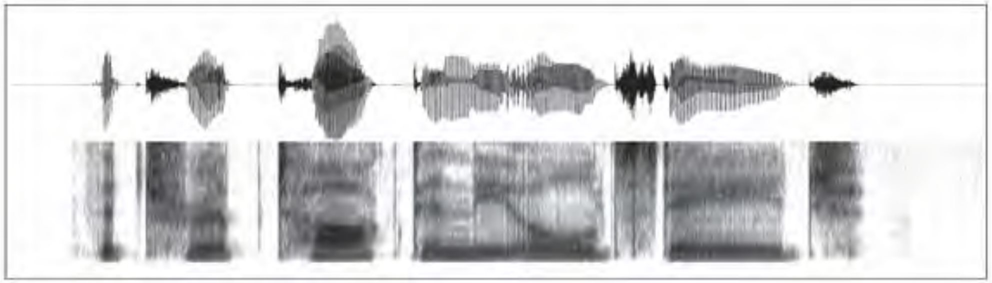
Department of Biomedical Engineering, Hearing Research Center, Boston University, Boston, MA, USA

ASLAN UNIVERSITE DE LYON

DDL

GHITZA & GREENBERG, *PHONETICA* (2009)

a) Original: *The trip talked in the old stage*

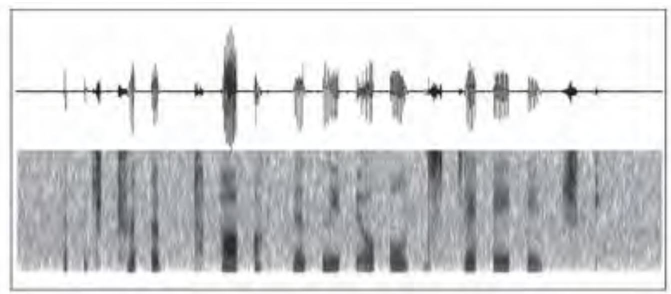


a

b) Time-compressed x3

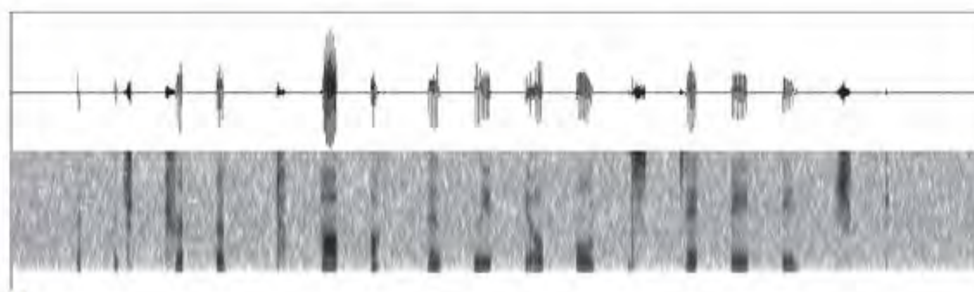
b

c) Same as b)
+ **40ms** speech intervals separated by **40ms** silence intervals
+ Background noise masking



c

d) Same as b)
+ **40ms** speech intervals separated by **80ms** silence intervals
+ Background noise masking



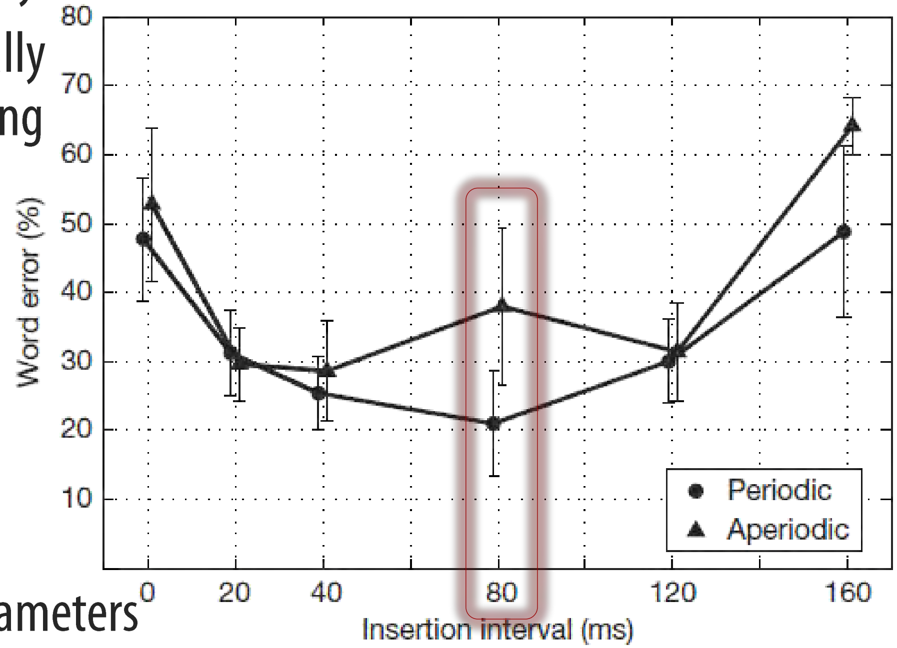
d

➔ Same speech compression but different phasing w.r.t. original timing

GHITZA & GREENBERG, *PHONETICA* (2009)

Results: U shape

- ✓ Variations in intelligibility although the phonetic degradation is constant
- ✓ Best intelligibility when the original rhythm is restored
- ✓ Especially when silences are periodically inserted (i.e. when the original phasing is restored)

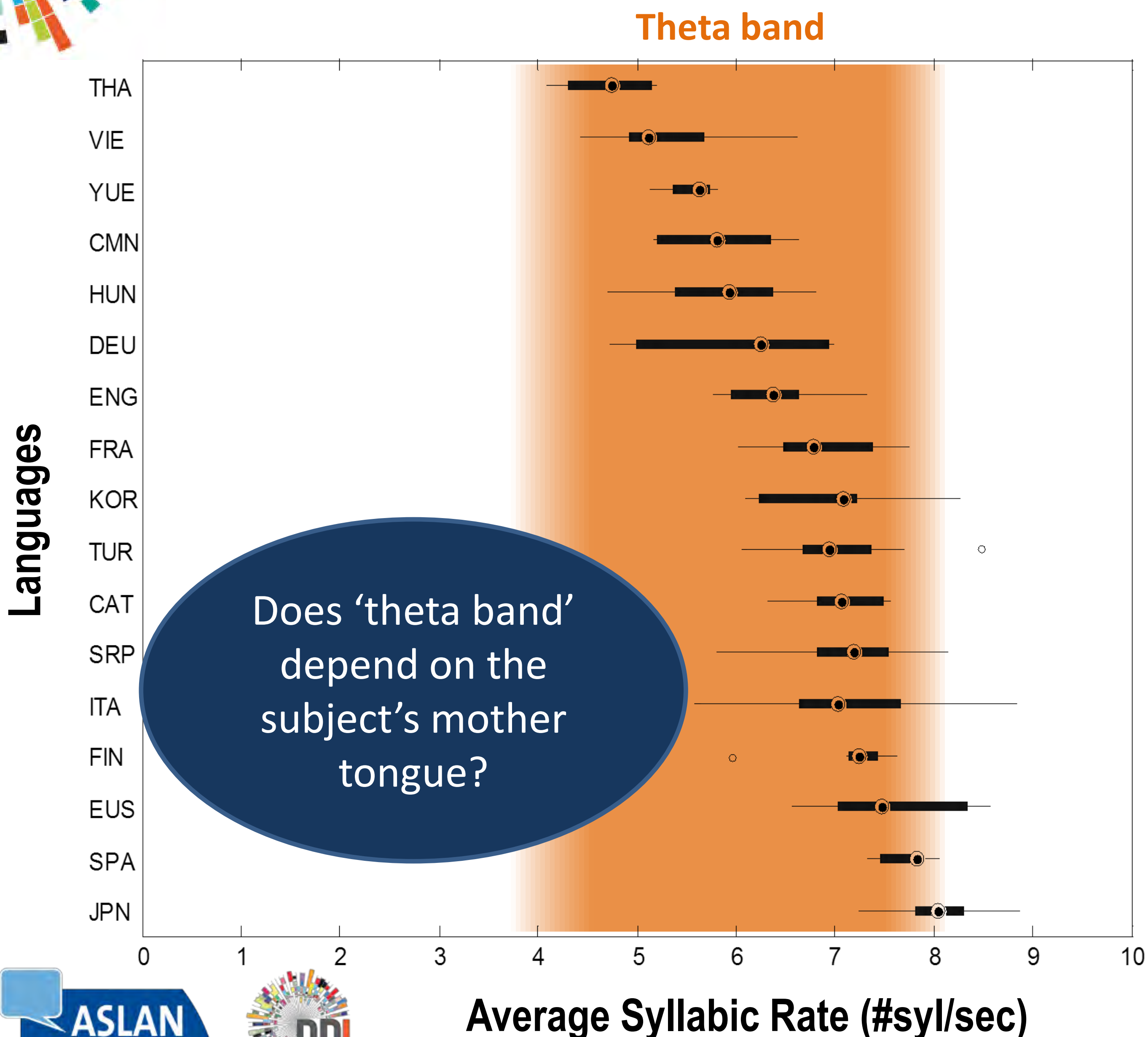


Interpretation

- ✓ WHAT you hear and WHEN you hear it
- ✓ Speech tracking at theta-rhythm
- ✓ Ghitza (2014)
- ✓ More thorough exploration of the parameters
- ✓ "The maximum information transfer rate through the auditory channel is the information in one uncompressed θ-syllable long speech fragment per one θ_{max} cycle. Equivalently, the auditory channel capacity is 9 θ-syllables/s." (Ghitzan 2014:1)

GHITZA, O. (2014). "Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech". *Frontiers in psychology*, 5.

AVERAGE SPEECH RATE VARIES ACROSS LANGUAGES



Does 'theta band' depend on the subject's mother tongue?

On average, Speech Rate is 70% faster in **Japanese** than in **Thai**

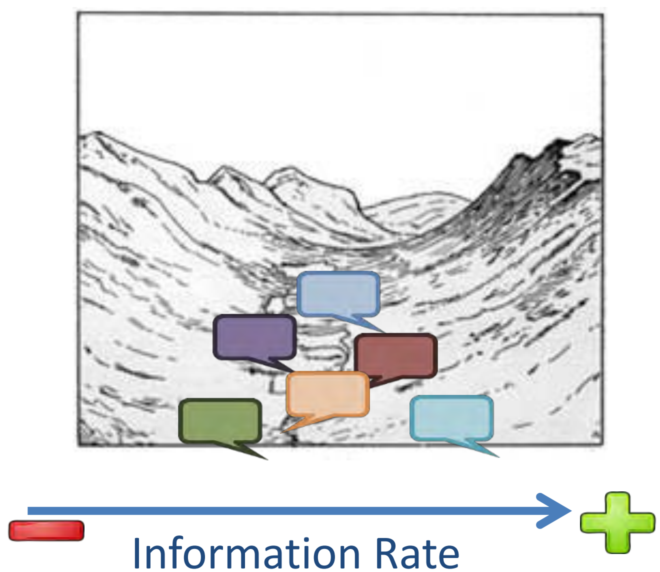
ASLAN UNIVERSITE DE LYON

DDL

So What?

- Relationship between brain oscillations & speech
  - ✓ Described and considered in terms of <u>*speech*</u> rate
  - ✓ Mostly interpreted in terms of <u>*information*</u> timing and processing
  - ➔ Information rate should be taken into account
- A cognitive sweet spot for Information Rate?

  - ✓ LOW: Not efficient enough (social function) and highly demanding (working memory)
  - ✓ HIGH: Demanding on the human physiological/cognitive capacity to process in real time



ASLAN UNIVERSITE DE LYON
DDL