

# SYLLABIC INFORMATION RATE: A CROSS-LANGUAGE APPROACH

François Pellegrino

*Dartmouth College  
September, 27 2012*

'Dynamique Du Langage' Laboratory  
CNRS & University of Lyon (Université Lyon 2)  
*ASLAN Advanced Studies on Language Complexity*

# OVERVIEW

- ✓ *Rationale & central hypothesis*
- ✓ *Glimpse of the study (Data, Results)*
- ✓ *Complexity and Information*
- ✓ *Perspectives*

# ACKNOWLEDGEMENT

## ✓ Co-authors

- Yoon-Mi Oh, Egidio Marsico & Christophe Coupé (DDL – University of Lyon – CNRS)



## ✓ Financial support

- Laboratory of Excellence ASLAN (*Advanced studies on language complexity*)
- CNRS – Grant PICS n°05766 ‘Dartmouth-DDL’, co-PIs: I. Chitoran & F. Pellegrino (2011-13)

## ✓ Reference:

- Pellegrino, F., Coupé, C. & Marsico, E., 2011, "A cross-language perspective on speech information rate", *Language*, 87:3, pp. 539-558

# RATIONALE

## ✓ "All languages are equally complex"

- This equal-language-complexity hypothesis is often mentioned, sometimes tested (e.g. Comrie, Planck, Dahl, Shosted), and also sometimes questioned (discussion on creole languages).
- Mostly (but not exclusively) studied at the morpho-syntactic level
- At the phonological level, mainly evaluated by Ian Maddieson

## ✓ A ill-defined problem (in our opinion)

- The basic assumption is that there is no 'simple' vs. 'elaborated' language, and it has drifted to: all linguistic codes (= language) are equally complex.
- From our point of view, what really matters is the capacity to convey information during speech communication, whatever the coding system is.

# RATIONALE (CONT'D)

- ✓ From 'equal overall complexity' to 'equal overall communicative capacity'
  - Language as a communicative device, with a limited 'bandwidth'
- ✓ **Limits on the information rate:**
  - Lower limit: speech communication not efficient enough to be socially useful and acceptable
  - Upper limit: exceeds the human physiological and cognitive capacities (or at least requires an extra-effort)
- ✓ **Consequences**
  - Given that phonological complexity is positively correlated to the quantity of information conveyed per phonological unit,
  - A less complex system will carry less information per unit; it will 'pack' more units per second during communication
  - ➔ There is an interaction between system complexity (paradigmatic dimension) and the speech rate (syntagmatic dimension)
- ✓ **An hypothesis already introduced 50 years ago**
  - ICPHS, 1961. H. Karlgren "Speech Rate and Information Theory"
    - "It is a challenging thought that general optimization rules could be formulated for the relation between speech rate variation and the statistical structure of a language. Judging from my experiments, there are reasons to believe that there is an equilibrium between information value on the one hand and duration and similar qualities of the realization on the other".
  - Also somewhat addressed by C. Hockett (1966)

Karlgren, H., 1961, "Speech Rate and Information Theory", *proc. of 4th ICPHS*, 671-677

HOCKETT, C.F. (1966). *The quantification of functional load: A linguistic problem*. Report Number RM-5168-PR, Rand Corp. Santa Monica

# *INTUITION VS. EMPIRICAL DATA*

# STUDY OVERVIEW

## ✓ Data (1st stage)

- About 3 hours of recording by 62 speakers of 7 languages (English, French, German, Italian, Japanese, Mandarin Chinese, Spanish) + *Vietnamese as a 'reference'*
- 20 texts (5-sentence long) translated into each language
- Postulate: on average, the semantic content is similar among languages

## ✓ Three parameters are estimated for each language $L$

- $ID_L$  Syllabic Information Density (average information carried by each syllable) - normalized
- $SR_L$  Syllabic Speech Rate (average number of syllables uttered per second)
- $IR_L$  Information Rate (average rate of semantic information per second) - normalized

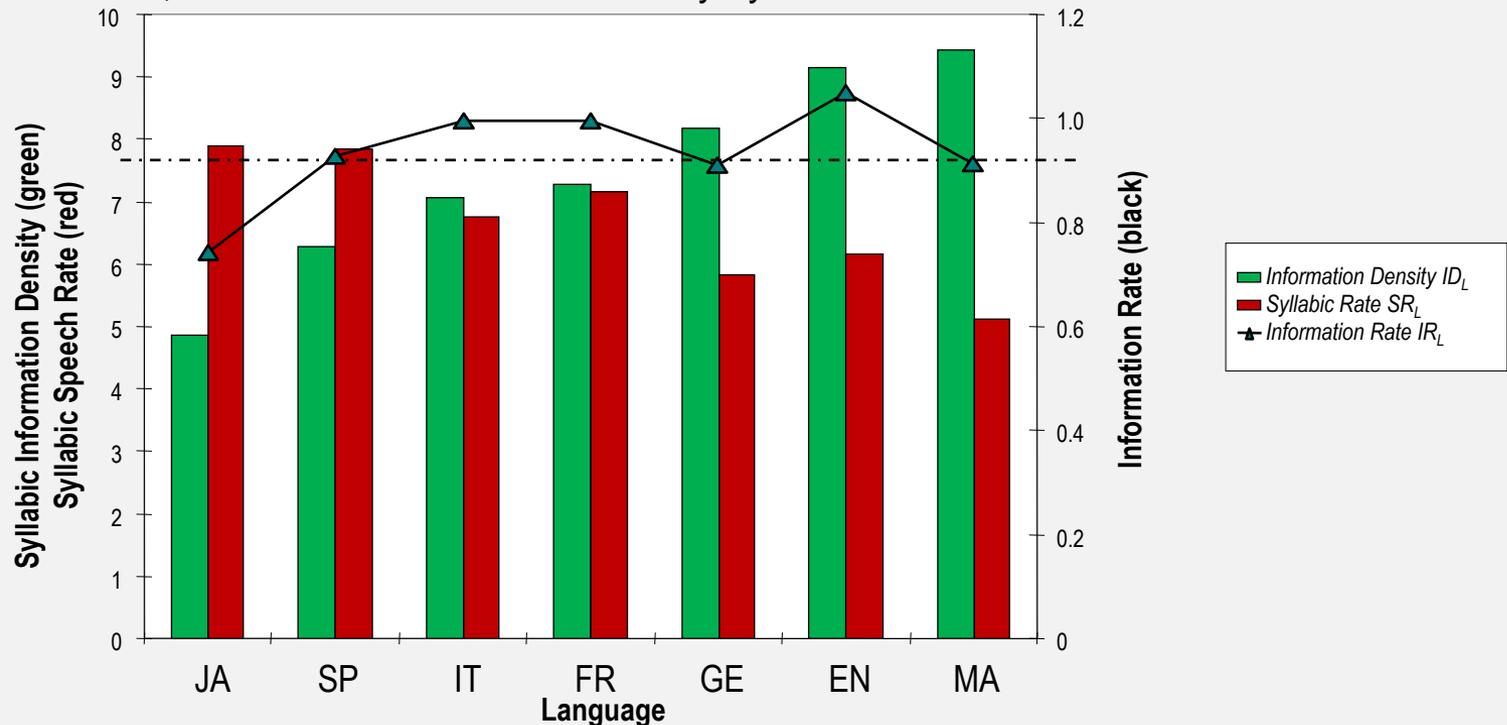
## ✓ A word of caution

- Very small language sample → no typological range

# SPEECH INFORMATION RATE

## ✓ Main results

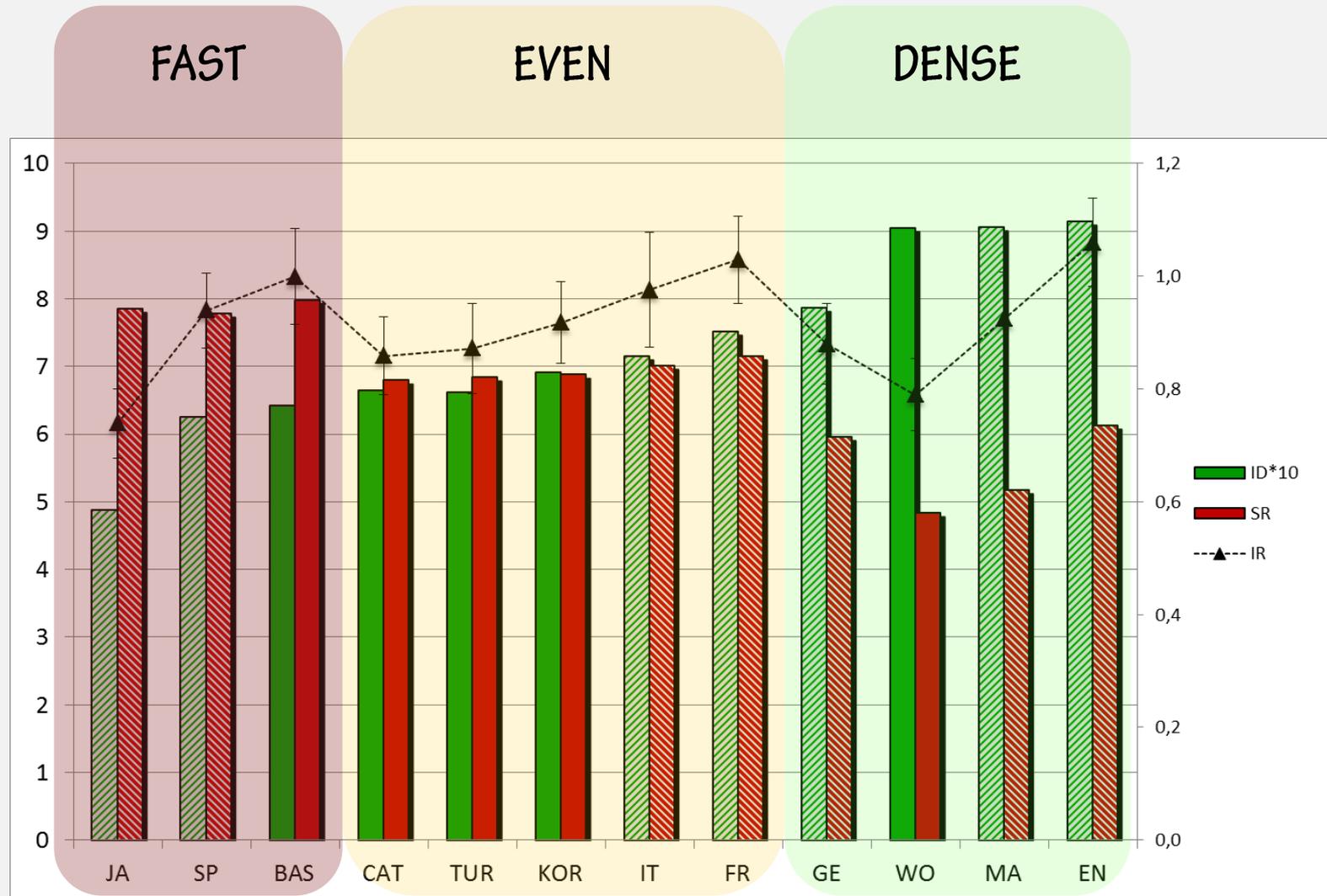
- Information Rate is not constant
  - Statistical analysis reveals significant differences for JA and EN
- An interaction between syllabic information density and speech syllabic rate
  - Correlation  $r = -0.93$ ;  $p < 0.01$
- Different strategies may lead to almost the same information rate
  - Spanish is characterized by a fast rate of low-density syllables
  - Mandarin exhibits a 34% slower syllabic rate with syllables 'denser' by a factor of 49%
  - In the end, their information rates differ only by 4%.



# EXTENSION (WORK IN PROGRESS)

## ✓ Additional languages

- Basque (4), Catalan (10), Korean (9), Turkish (10), Wolof (8), (+ Cantonese, Thai & Khmer)
- 12 languages so far



# THE NOTION OF SYLLABIC INFORMATION

- ✓ **In the previous slide, syllabic information is**
  - an average value computed by dividing the total amount of semantic information of an excerpt by its (phonological) number of syllables
  - Estimated from the temporal (syntagmatic) dimension of speech communication (audio recordings)
  
- ✓ **But syllabic information may also be related to other notions in information theory**
  - The syllable complexity (in terms of number of constituents – *description-based*)
    - A larger repertoire allows a larger transmission channel capacity
    - Even if syllabic complexity may be defined in other ways!
  - The Shannon entropy of the distribution of syllable frequency (*usage-based*)
    - See below
  
- ✓ **What we have done next**
  - Gather large text corpora and syllabicate them (more than 100k syllables)
  - Estimate statistics on syllable structure and usage
    - Average syllabic complexity (in terms of # of segments + tone if present)
    - Syllabic entropy

# SYLLABIC ENTROPY

- ✓ Considering that language  $L$  is a source of linguistic sequences composed of syllables ( $\sigma$ ) from a finite set ( $N_L$ )
- ✓ Assuming that the syllables are independent from each other
- ✓ Entropy of  $L$  is  $H_L$

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

## ✓ Comments

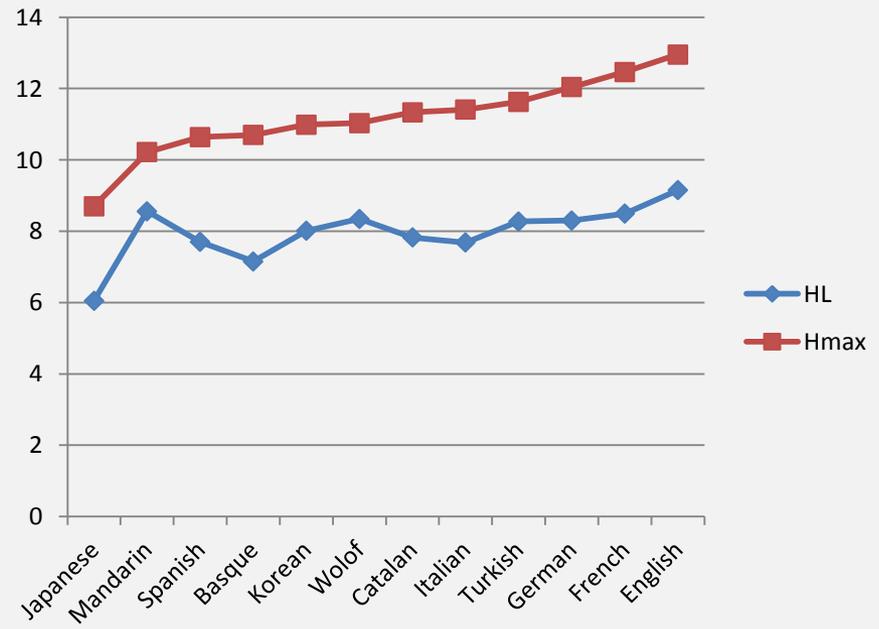
- $H_L$  is always inferior to  $\log_2(N_L)$  (channel capacity  $H_{max}$ )
- $H_L = H_{max}$  iff the syllables are equiprobable.
- We consider here the unconditional entropy → dependencies between adjacent syllables are not taken into account

# SYLLABIC COMPLEXITY AND SYLLABIC ENTROPY

| Language | Syl. Inventory Size | Syllabic complexity (Type) | Syllabic complexity (Token) | Syllabic Entropy $H_L$ |
|----------|---------------------|----------------------------|-----------------------------|------------------------|
| Japanese | 416                 | 2.65                       | 1.93                        | 6.05                   |
| Mandarin | 1 191               | 3.87                       | 3.58                        | 8.56                   |
| Spanish  | 1 593               | 3.30                       | 2.40                        | 7.70                   |
| Basque   | 1 658               | 3.15                       | 2.04                        | 7.15                   |
| Korean   | 2 026               | 2.83                       | 2.48                        | 8.01                   |
| Wolof    | 2 094               | 3.48                       | 2.49                        | 8.34                   |
| Catalan  | 2 582               | 3.20                       | 2.29                        | 7.83                   |
| Italian  | 2 719               | 3.50                       | 2.30                        | 7.68                   |
| Turkish  | 3 163               | 3.12                       | 2.20                        | 8.27                   |
| German   | 4 207               | 3.70                       | 2.68                        | 8.30                   |
| French   | 5 646               | 3.50                       | 2.21                        | 8.49                   |
| English  | 7 931               | 3.70                       | 2.48                        | 9.15                   |

✓ How much of the potential of a given syllabic inventory is really used?

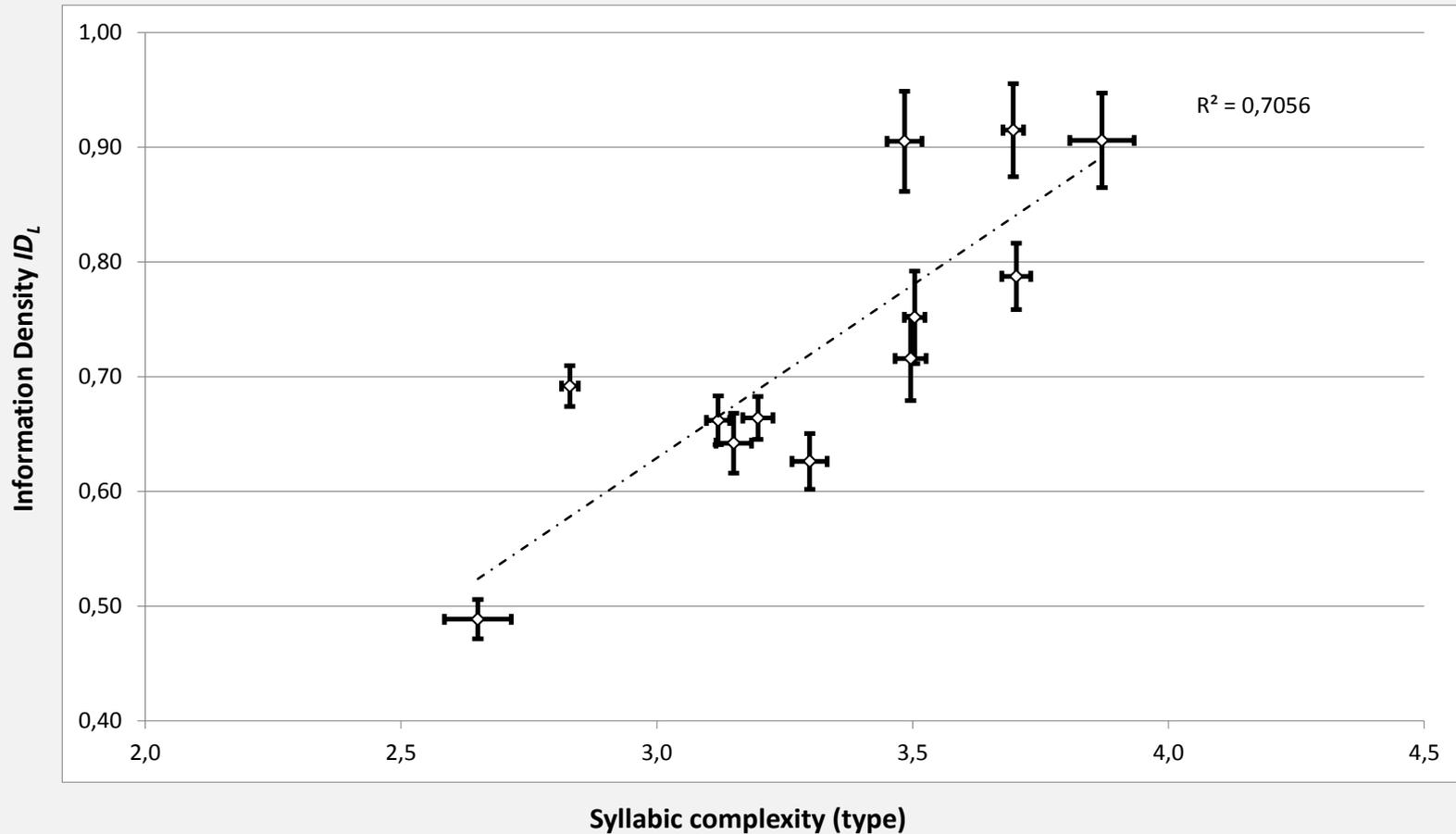
- $H_L \sim 3/4 H_{max}$
- Uneven (Zipf-law) distribution of syllable frequency
- Can a language reach its maximum theoretical capacity?



# COMPLEXITY AND INFORMATION DENSITY

✓ Without taking usage-frequency into account (*type*)

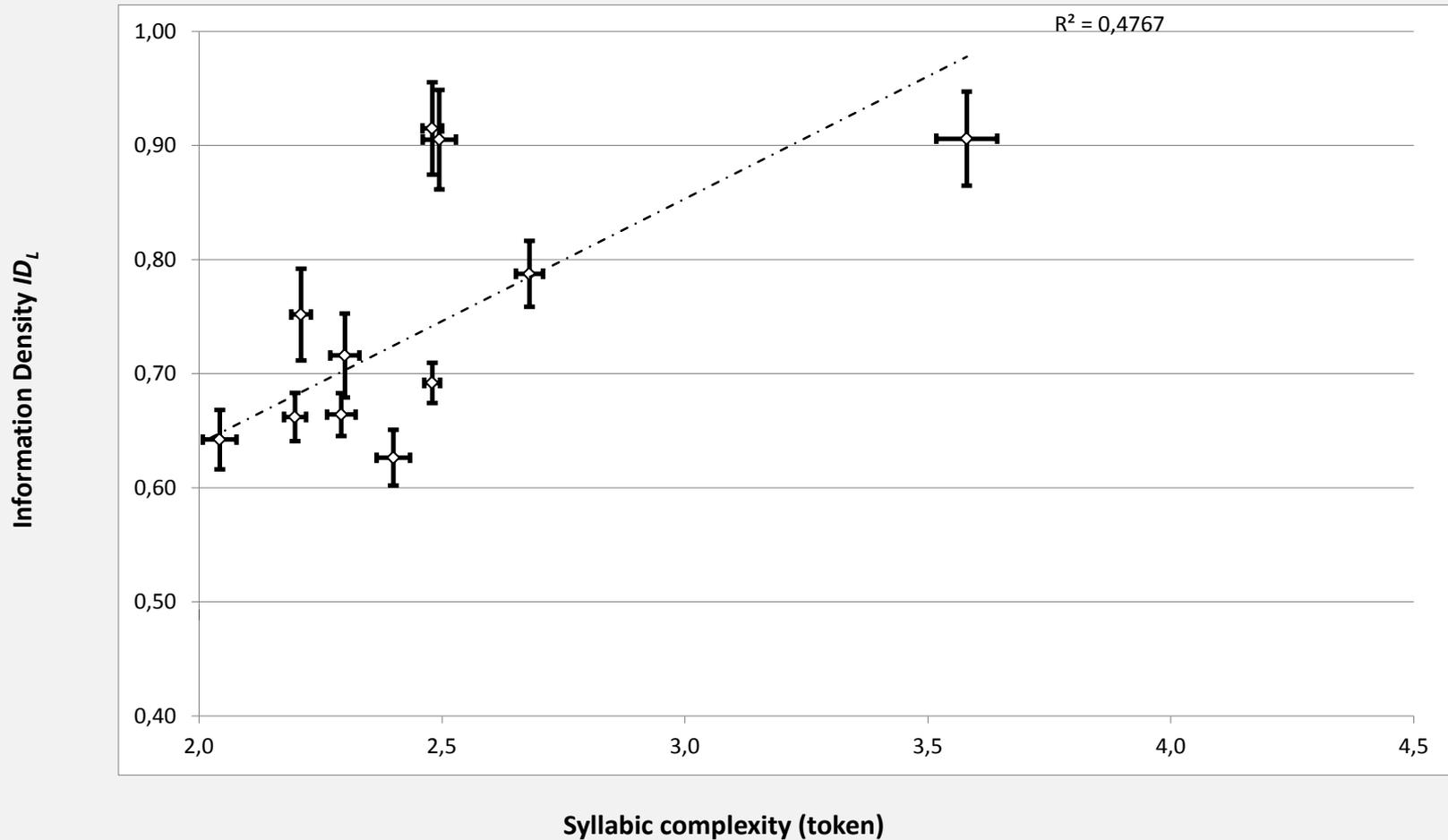
➤ Pearson  $R^2 = 0.71$



# COMPLEXITY AND INFORMATION DENSITY

✓ Taking usage-frequency into account (*token*)

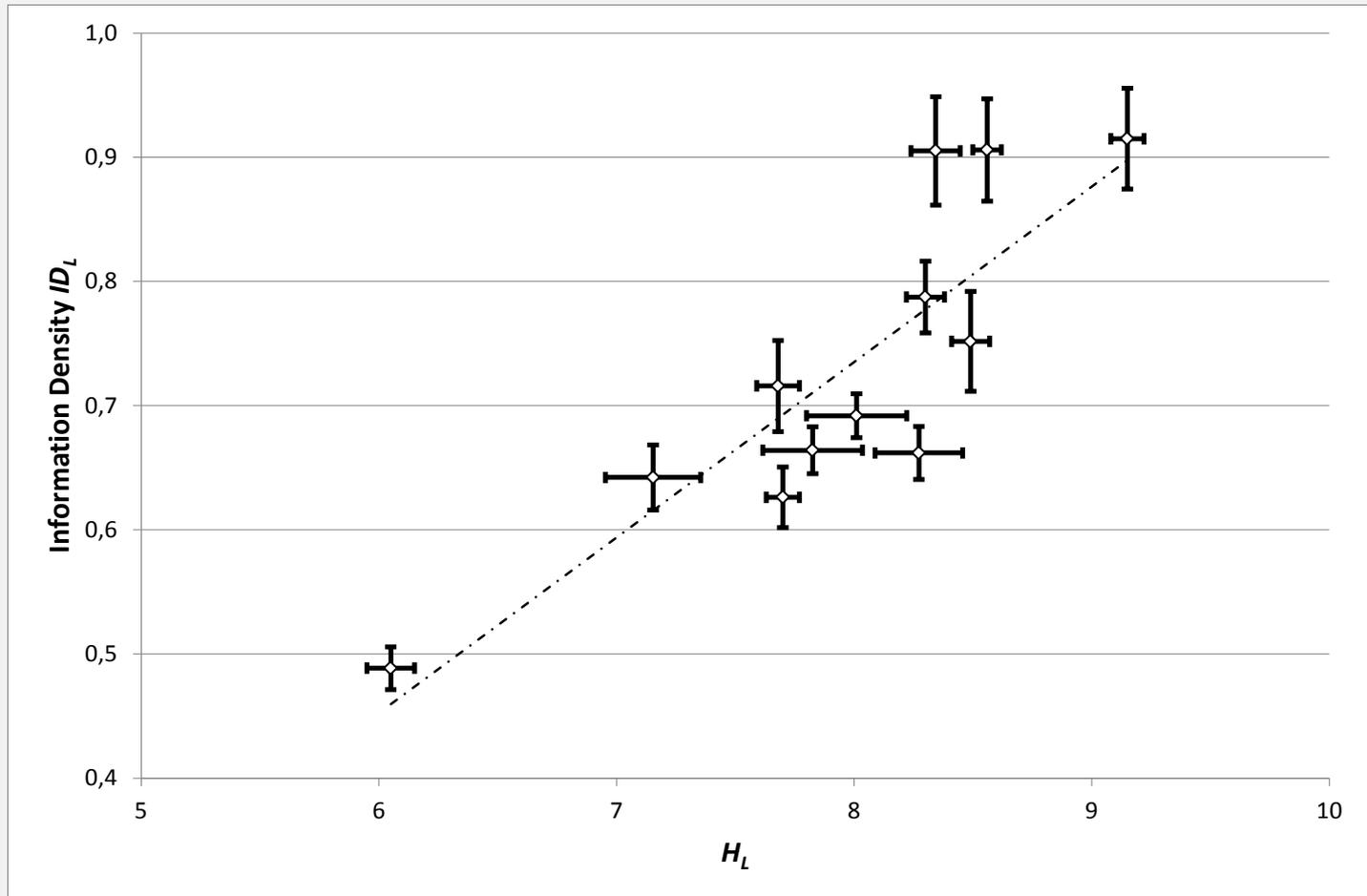
➤ Pearson  $R^2 = 0.48$



# ENTROPY AND INFORMATION DENSITY

## ✓ $H_L$ as a predictor of $ID$

- Linear regression ( $R^2 = .72$ ,  $p < .001$   $\propto$   $\rho^2 = .75$ ,  $p < .001$ )
- Exponential regression ( $R^2 = .79$ )



# COMPLEXITY AND INFORMATION

## ✓ Two paradigmatic measures of phonological complexity

- Based on syllabic frequencies of use from large written corpora (Celex, etc.)
  1. Average syllabic complexity in terms of # of constituents (# of segment + 1 for tone)
    - A classical 'linguistic' way to measure syllabic complexity
  2. Shannon entropy computed from the distribution of syllable frequencies
    - A 'non linguistic' way to measure the information conveyed by a source of communication

## ✓ Comments

- Taking tone into account is essential for Mandarin
- Syllabic Complexity is a very good predictor of Syllabic Information Density
- Syllabic entropy seems to be better and it doesn't require counting the number of internal constituents of the syllable

# CONCLUSION AND PERSPECTIVES

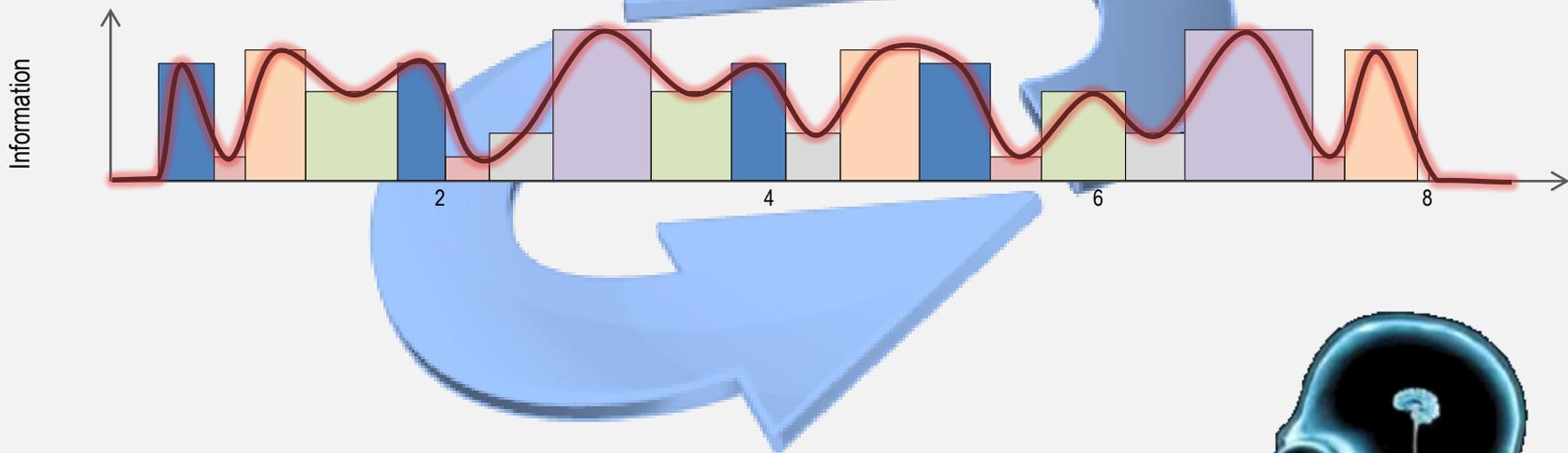
## ✓ Back to our hypotheses

- ✗ ➤ Speech information rate hover over the same values among languages
  - ✓ A) Phonological complexity is positively correlated to the quantity of information conveyed per phonological unit
  - ✓ B) A less complex system will carry less information per unit; it will 'pack' more units per second during communication
  - ✓ ⇒ There is an interaction between system complexity (paradigmatic dimension) and the speech rate (syntagmatic dimension)

## ✓ Perspectives

- Small sample → no typological range
  - Extend the language sample
    - Typological diversity
    - Related dialects
  - Check if the tendencies are robust
- Bilingual speakers (BA/SP, CA/SP, MA/KO)
- Explain this trade-off between speed and density
  - Speech communication efficiency
  - Sociological constraints
  - Cognitive constraints (information processing in the brain)
- Explore its relationship with linguistic strategies
  - Study focused on the 'bricks' of speech (syllables), neglecting phonetic aspects
  - Extension to morpho-syntactic significance of those bricks

# Languages of the World



# THE RELEVANCE OF STUDYING DEGRADED SPEECH

## ✓ *Speech in noise / Cocktail Party effect*

- Basic distinction between energetic and informational masking
- But much more to explore!
- Suggestion : interferent SaSa-like signal

## ✓ *Accelerated speech*

- Interaction with brain oscillations
  - Ghitza & Greenberg, 2009 (Phonetica)
  - Peelle & Davis, 2012 (frontiers in Psychology)
- The ODYSSEE PROJECT (PI Véronique Boulenger, DDL, Lyon)
  - ODYSSEE - Oscillatory Dynamics of Speech Sensorimotor Ensembles
- ➔ Investigate the neurocognitive mechanisms underlying adaptation to natural fast speech in children and adults
- ➔ Provide evidence for the existence of audio-articulatory mappings during speech perception by examining the dynamics of cortical oscillations using MEG.

✓ ...

# PERCEPTION OF TIME-REVERSED SPEECH

## ✓ Background

- Reversed Speech is commonly used as control condition in fMRI.
- It is considered as a speech-analog NON LINGUISTIC condition
- Probably not true

## ✓ Reversed speech in French

- Phonemes from reversed speech may be partially identified par phoneticians: ~90% for nasals, liquids, fricatives and oral vowels (Pellegrino et al. 2010)
- Reversed speech may be recognized as words (in progress) : 'mal/ (evil' ) reversed into a signal analog to /lam/ ('blade'), recognized as a word in a way similar to an original [lam].
- ➔ Reversed speech probably activates areas involved in language processing
- Perception of words and non words with various time-reversals
  - Large individual variation in performances in non words
  - Correlated to the lateralization of the efferent auditory system (and thus to its functionality)
  - EEG study revealed different patterns according to the time span of the perturbation

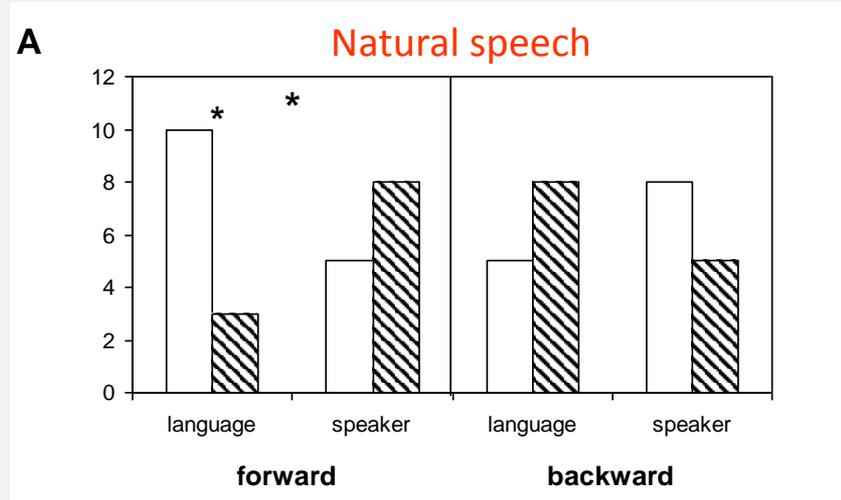
## ✓ Rhythm and Reversed Speech

- Time-reversal modifies the perception of duration
- Time-reversal hinders the capacity of non human mammals to discriminate between rhythm classes

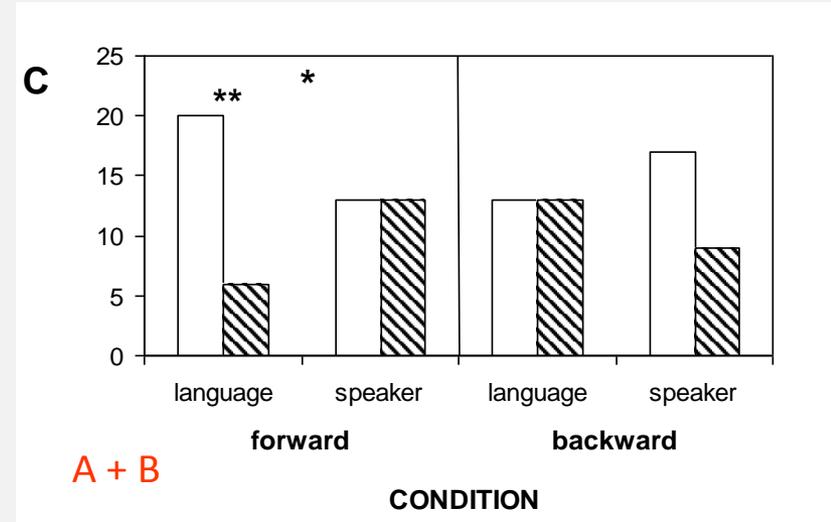
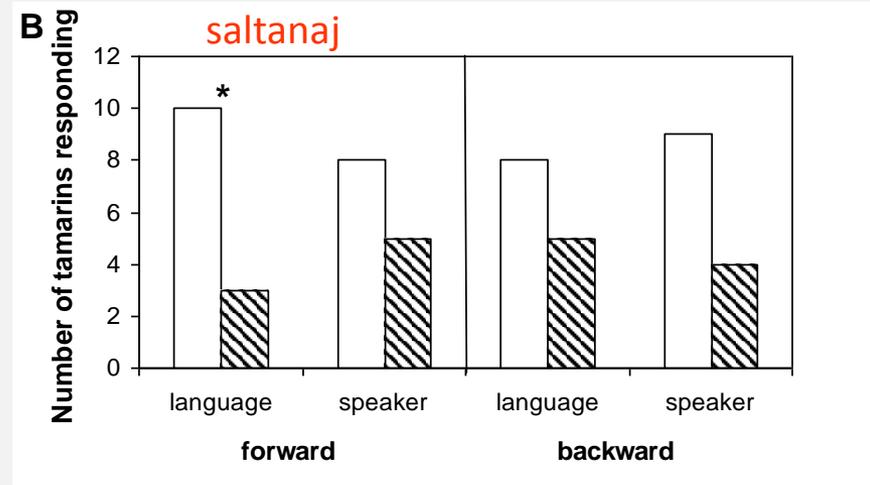
# NON-HUMAN EXPERIMENTS ON REVERSED SPEECH

## ✓ Cotton-top tamarin monkeys, (Ramus et al. 2000)

- Dutch vs. Japanese
- Natural speech / saltanaj



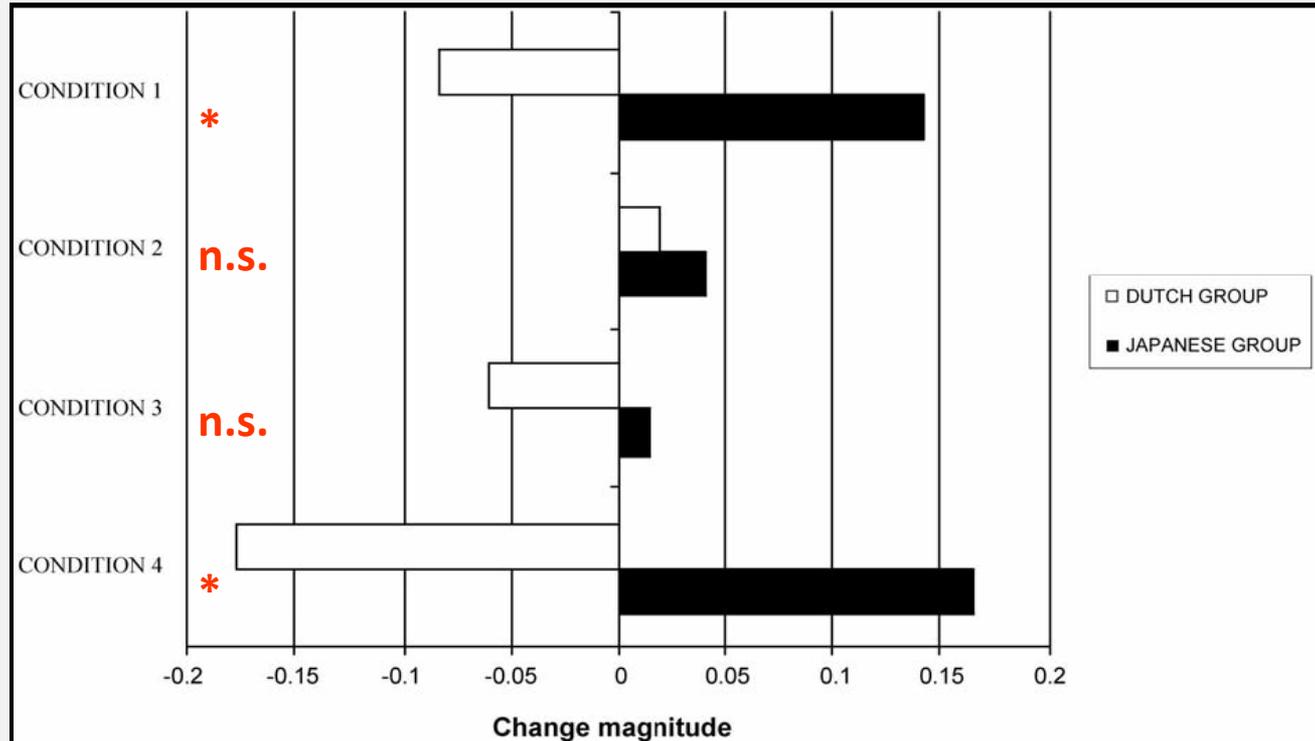
White bar = detected change  
Dashed bar = undetected change



# NON-HUMAN EXPERIMENTS ON REVERSED SPEECH

## ✓ Rats (Toro et al., 2005)

- Discrimination between Dutch and Japanese
- Material
  - saltanaj, forward (condition 1) et backward (C2)
  - Natural speech (several female speakers (C3) ;
  - or 1 female speaker (C4))



From Toro et al., 2005. Mean discrimination ratio (JA) – Mean discrimination Ration (DU)

# REVERSED SPEECH PERCEPTION AND INDIVIDUAL VARIATION

## ✓ Laterality & Medial olivocochlear bundle

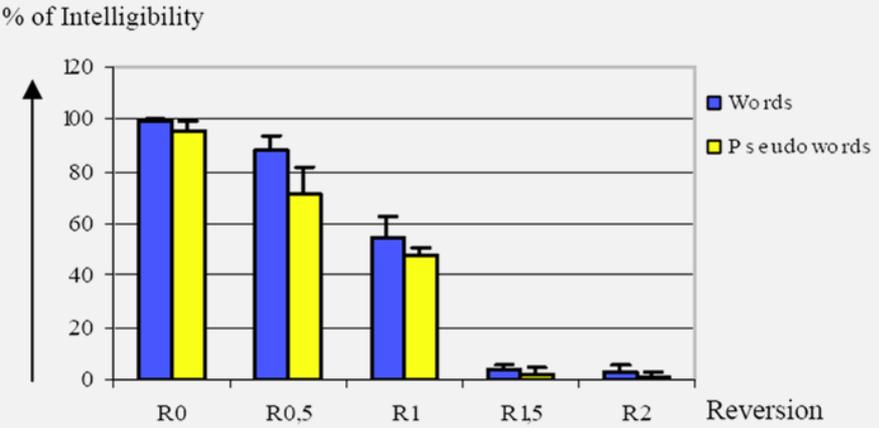


Figure 1: Rate of intelligibility for words and pseudowords plotted against the size of the reversion.

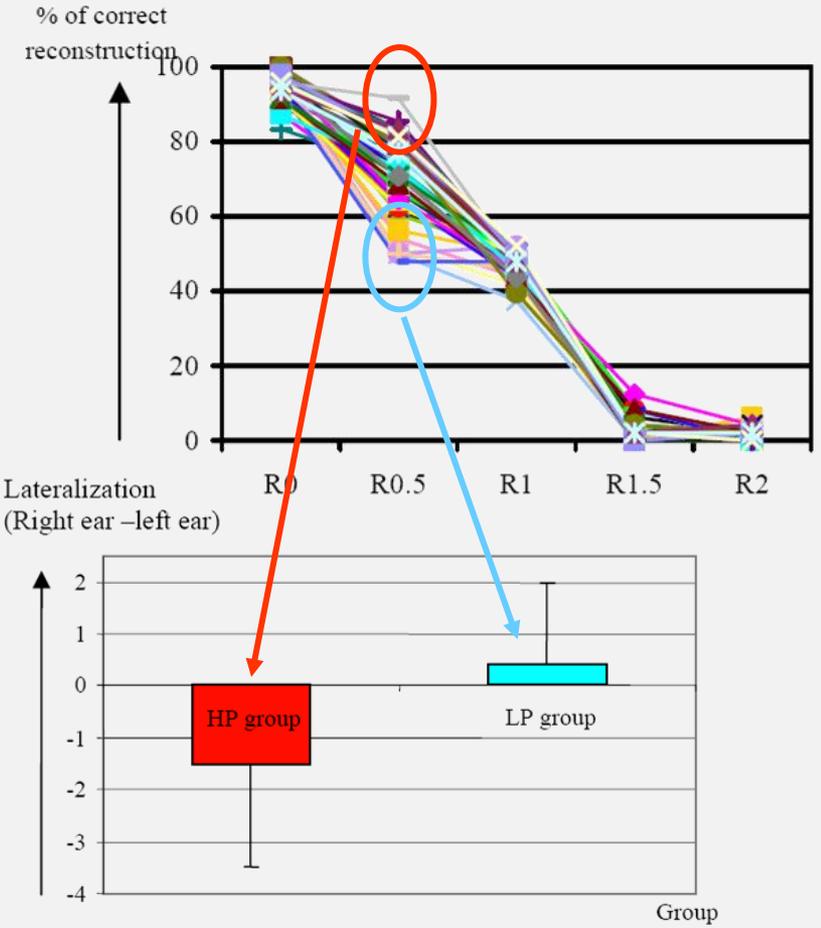
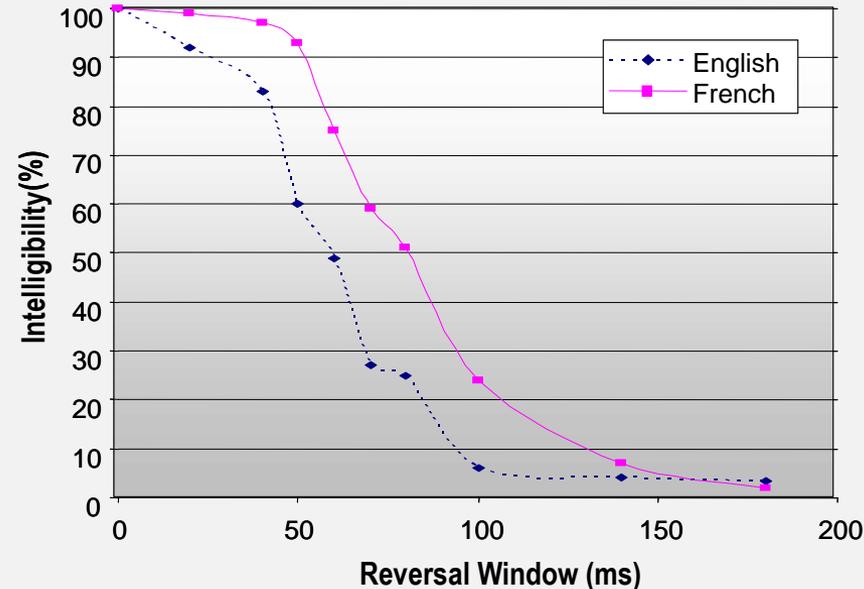


Figure 3: Lateralization for the two groups of subjects.

# PERCEPTION OF TIME-REVERSED SPEECH

## ✓ Interest of cross-language comparison



Adapted from Greenberg & Arai, (2001)  
and Magrin-Chagnolleau *et al.*, (2002).

Greenberg, S. & Arai, T. (2001). The relation between speech intelligibility and the complex modulation spectrum, Proceedings of the 7th Eurospeech Conference on Speech Communication and Technology (Eurospeech-2001)

Magrin-Chagnolleau, I., Barkat, M., Meunier, F. (2002). Intelligibility of reverse speech in French: A perceptual study. Proceedings of the 7th International Conference on Spoken Language Processing, Interspeech 2002, 3, p. 1669, Denver, Colorado.

THANK YOU!