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
OUTLINE

- General Background
- Structural organization of phonological information
- Temporal regulation of phonological information
- Conclusion
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)
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
THE LANGUAGE NICHE

Language A Niche

Language B Niche

Language C Niche
The language niche is nested.
The language niche is **nested**

**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 **ubiquitous and highly diverse, and functional** whatever the environment (global fitness)

One of these functions is to **convey information**
The language niche is nested

**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
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)
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
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
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
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

  - 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)
A NEW CENTURY, A NEW DAWN FOR FL...

- **Initiative: D. Surendran and colleagues**

- **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)**

- **Our group**
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 ($\omega$) from a finite set ($N_L$)

Entropy $H(L) = \text{Average quantity of information per unit}$

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

✓ 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)
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 \( x \sim y \) = 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**

<table>
<thead>
<tr>
<th>Form</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>pal</td>
<td>300</td>
</tr>
<tr>
<td>pil</td>
<td>200</td>
</tr>
<tr>
<td>bal</td>
<td>150</td>
</tr>
<tr>
<td>bil</td>
<td>150</td>
</tr>
<tr>
<td>pul</td>
<td>100</td>
</tr>
<tr>
<td>bul</td>
<td>100</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1000</td>
</tr>
</tbody>
</table>

Inventory: /a i u p b l/

\[ N_L = 6 \quad H(L) = 2.47 \]

**Contrast /a-i/**

<table>
<thead>
<tr>
<th>Form</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>pal</td>
<td>300</td>
</tr>
<tr>
<td>pil</td>
<td>200</td>
</tr>
<tr>
<td>bal</td>
<td>150</td>
</tr>
<tr>
<td>bil</td>
<td>150</td>
</tr>
<tr>
<td>pul</td>
<td>100</td>
</tr>
<tr>
<td>bul</td>
<td>100</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1000</td>
</tr>
</tbody>
</table>

**Fluctuation of /a-i/**

\[ H(L_{ai}) = 1.69 \]

\[ FL(a-i) = \frac{(2.47-1.69)}{2.47} = 31.8\% \]

\[ FL(a-u) = 23.1\% \]

\[ FL(i-u) = 21.0\% \]

**Phoneme /a/**

\[ FL(x) = \frac{1}{2} \sum_y FL(x, y) \]

\[ FL(a) = \frac{1}{2} (FL(a-i)+FL(a-u)) = \frac{1}{2} (31.8+23.1)=27.45\% \]

**Fictive Lexicon**

<table>
<thead>
<tr>
<th>Form</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>pal</td>
<td>500</td>
</tr>
<tr>
<td>pil</td>
<td>300</td>
</tr>
<tr>
<td>bal</td>
<td>150</td>
</tr>
<tr>
<td>bil</td>
<td>150</td>
</tr>
<tr>
<td>pul</td>
<td>100</td>
</tr>
<tr>
<td>bul</td>
<td>100</td>
</tr>
<tr>
<td>TOTAL</td>
<td>1000</td>
</tr>
</tbody>
</table>

\[ H(L_{ai}) = 1.69 \]

\[ FL(a-i) = \frac{(2.47-1.69)}{2.47} = 31.8\% \]

**Fluctuation of /a-i/**

\[ FL(a) = \frac{1}{2} (FL(a-i)+FL(a-u)) = \frac{1}{2} (31.8+23.1)=27.45\% \]
# Material

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

20,000 most frequent words (inflected forms) considered, except for Cantonese (5,000 forms) & Italian (15,788 forms)
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)
FINER GRAIN: HOW TO READ A RESULT GRAPH

Units (ranked by decreasing order of FL)
**Most “natural” scale of organization: segments**

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


- Methodological details
- Discussion
  - Lemmas vs inflected wordforms
  - Types vs. Tokens
  - Consonantal bias hypothesis (Nespòr, Peña, & Mehler, 2003)
## Feature Scale

<table>
<thead>
<tr>
<th>Language</th>
<th>ISO 639-3 Code</th>
<th>Phonological system</th>
<th>Number of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantonese</td>
<td>YUE</td>
<td>V</td>
<td>13, 22</td>
</tr>
<tr>
<td>English</td>
<td>ENG</td>
<td>V</td>
<td>22, 28</td>
</tr>
<tr>
<td>Japanese</td>
<td>JPN</td>
<td>V</td>
<td>10, 16</td>
</tr>
<tr>
<td>Korean</td>
<td>KOR</td>
<td>V</td>
<td>8, 22</td>
</tr>
<tr>
<td>Mandarin</td>
<td>CMN</td>
<td>V</td>
<td>7, 22</td>
</tr>
<tr>
<td>German</td>
<td>DEU</td>
<td>V</td>
<td>22, 24</td>
</tr>
<tr>
<td>Swahili</td>
<td>SWH</td>
<td>V</td>
<td>5, 30</td>
</tr>
<tr>
<td>Italian</td>
<td>ITA</td>
<td>V</td>
<td>8, 25</td>
</tr>
<tr>
<td>French</td>
<td>FRA</td>
<td>V</td>
<td>15, 21</td>
</tr>
</tbody>
</table>

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)
From features to articulatory dimensions

Aperture: 3 sets of mergers

<table>
<thead>
<tr>
<th>Front unrounded</th>
<th>Back rounded</th>
<th>Back unrounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon, \varepsilon, \imath )</td>
<td>( \mathcal{O}, \mathcal{U} )</td>
<td>( \Lambda, \mathcal{W} )</td>
</tr>
</tbody>
</table>

Anteriority: 2 set of mergers

<table>
<thead>
<tr>
<th>Lower-mid unrounded</th>
<th>High unrounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon, \Lambda )</td>
<td>( \imath, \mathcal{W} )</td>
</tr>
</tbody>
</table>

Vocalic inventory of Korean

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.
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
✓ \( \sim 50\% \) 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**
  - 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

ENCODING
• Resolutions (x2)
• #Grey levels (256 vs. 16)
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)
AVERAGE INFORMATION RATE: OUR DEFINITIONS

Information rate defined as:

**Average Information per syllable** \(\times\) **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 \(\Rightarrow\) **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) \(\Rightarrow\) **What you do say**
**DATA**

**Languages** ($N_L = 17$)


**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
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)
INFORMATION DENSITY (SYLLABIC ENTROPY)

Language $L$

- Source of linguistic sequences of syllables ($\sigma$) drawn from a finite set $\Sigma$ of $N_L$ distinct syllables
- Syllable probabilities estimated from a large corpus

Shannon Entropy (= average information per syllable)

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

- $H_L$ is always inferior to $\log_2(N_L)$ (channel capacity $H_{\text{max}}$)
- $H_L = H_{\text{max}}$ iff the syllables are equiprobable (maximal uncertainty)

But syllables are not independent. Context matters!!

Conditional Entropy

- Average information per syllable, given the context

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

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

$\Rightarrow 0 \leq \text{Conditional Entropy} \leq \text{Shannon Entropy} \leq \log_2(N_L)$
RESULT #1
A BALANCE IN INFORMATION RATE

Speech Rate (#syl/sec) vs. Information Density (bits/syl)

EXISTENCE OF DIFFERENT STRATEGIES
RESULT #2
SPEECH RATE VS INFORMATION RATE DISTRIBUTIONS

SPEECH RATE ➔ Broad distribution

INFO RATE ➔ Compact distribution
PAIRWISE DIVERGENCES BETWEEN LANGUAGES

Jensen-Shannon divergence

- Syllable Info (Social Convention)
- Speech Rate (Individual Behavior)
- Information Rate (Functional Fitness)
**STUDY 2: FACTUAL CONCLUSION**

- Languages exhibit large variation
  - Syllabic Speech Rate
  - Syllabic Information density

- Average **Information Rates** tend to exhibit less 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
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

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: $\delta (<3\text{Hz})$, $\theta (4-8 \text{Hz})$, $\alpha (8-12 \text{Hz})$, $\beta (15-35 \text{Hz})$, $\gamma (25-40 \text{Hz})$
    - 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)
SPEECH RATE AND BRAIN OSCILLATIONS: ILLUSTRATION

Ghitza and Greenberg (2009) ; Ghitza (2014)

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

b) Time-compressed x3

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

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

⇒ Same speech compression but different phasing w.r.t. original timing
**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 $\theta$-syllable long speech fragment per one $\theta_{max}$ cycle. Equivalently, the auditory channel capacity is 9 $\theta$-syllables/s.” (Ghitzan 2014:1)

Average Speech Rate Varies Across Languages

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

Does ‘theta band’ depend on the subject’s mother tongue?
SO WHAT?

Relationship between brain oscillations & speech

- Described and considered in terms of *speech* rate
- Mostly interpreted in terms of *information* 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