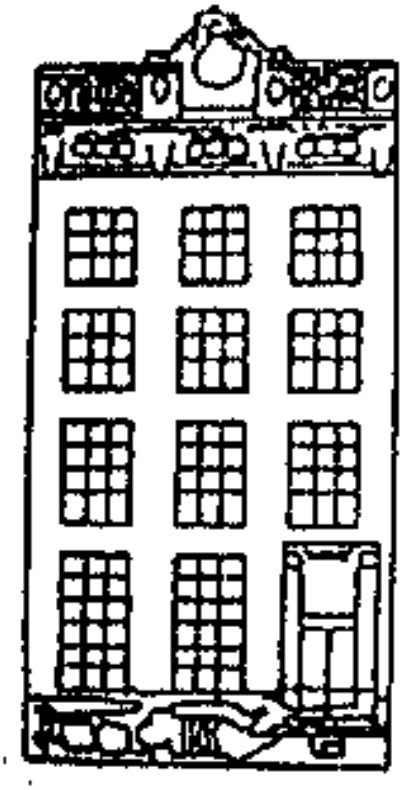
Evidence for Efficiency in Vowel Production

R.J.J.H. van Son and Louis C.W. Pols

University of Amsterdam
Institute of Phonetic Sciences/ACLC
Herengracht 338, 1016 CG Amsterdam
The Netherlands

Rob.van.Son@hum.uva.nl



Introduction

- *Speech is the Missing Information*
  (Lindblom, JASA 1996)
- Trade-off for Efficiency:
  - Minimize Speaking *Effort*
  - Maximize *Intelligibility*
- Compare (Liberman, Lang&Speech 1963):
  - *A stitch in time saves **nine***
  - *The next number is **nine***
- Vowel Reduction is Affected by:
  - Word Frequency
  - Word Predictability
  - Phoneme Predictability**???**

Single Phoneme Information Content, i.e., Redundancy

$$I_s = -\log_2\left(\frac{Frequency([word-onset]+s)}{Frequency([word-onset]+any\ segment)}\right)$$

I_s: Segmental Information in bits

s: Phoneme Segment

[word-onset]: Preceeding Segment Sequence

Correlate I_s to Measures of Reduction

Examples

- ***/a:/*** in ***/x@da:n/*** (Dutch: *gedaan* English: *done*)

  Probability (***/a:/ | /x@d_ /***) =
  Frequency(/x@da:/) = **14946 / 81360** = **0.184**
  Frequency(/x@d */)

  $I_s = Log_2(\mathbf{0.184}) = \mathbf{2.44}$ bits

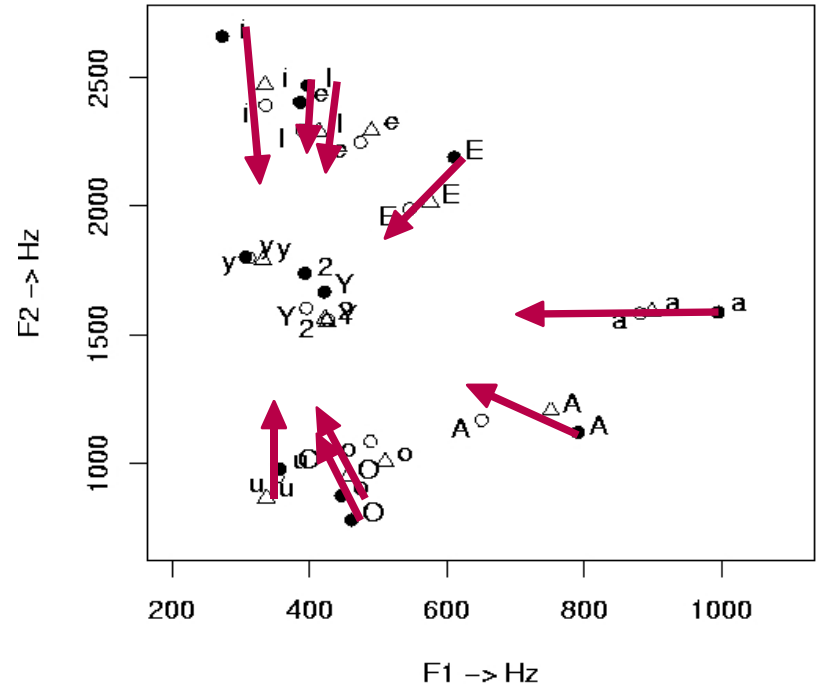
- ***/i/*** in ***/x@dint/*** (Dutch: *gediend* English: *served*)

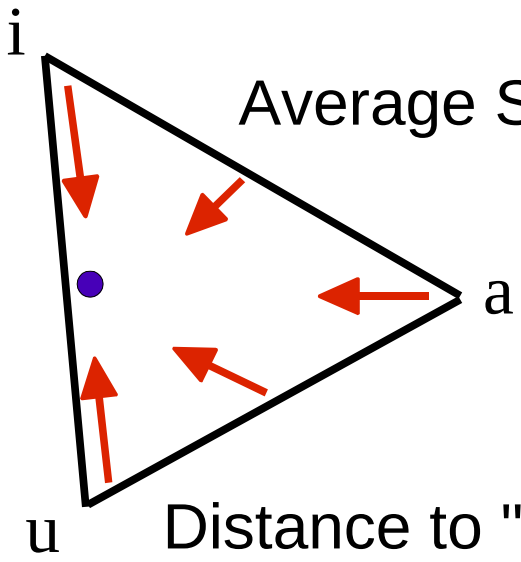
  Probability (***/i/ | /x@d_ /***) =
  Frequency(/x@di/) = **1225 / 81360** = **0.015**
  Frequency(/x@d */)

  $I_s = Log_2(\mathbf{0.015}) = \mathbf{6.05}$ bits

Acoustic Measures of Vowel Reduction

- Duration
- F_1/F_2 contrast:





Average Spectral Vowel Reduction

Distance to "Center of Reduction" in Semitones, excluding *SCHWA* (Equalizes the Variances in F_1 and F_2)

Factors Influencing Vowel Reduction
(in Dutch)

Account for:

- Speaker Identity
- Vowel Identity
- Speaking Style
- Lexical Stress (CELEX word list)
- Prominence (Automatic 0-3)

Use *Quasi-Uniform* Subsets for Calculating Correlations

Rules for Automatic Prominence Assignment

(Streefkerk, 2001/2002)

Based on:

- Parts-of-Speech (POS)
- Wordlength
- Position

Agrees with Human Transcribers:
Cohen's Kappa = **0.62**

- Function Words:
  Prominence = 0
- Content Words:
  Prominence from 1-3
  (weak -> strong)

SPEECH

50 kWord *IFAcorpus*

4 male + 4 female speakers (15-66 yoa)
40,385 vowels

Speaking Styles:

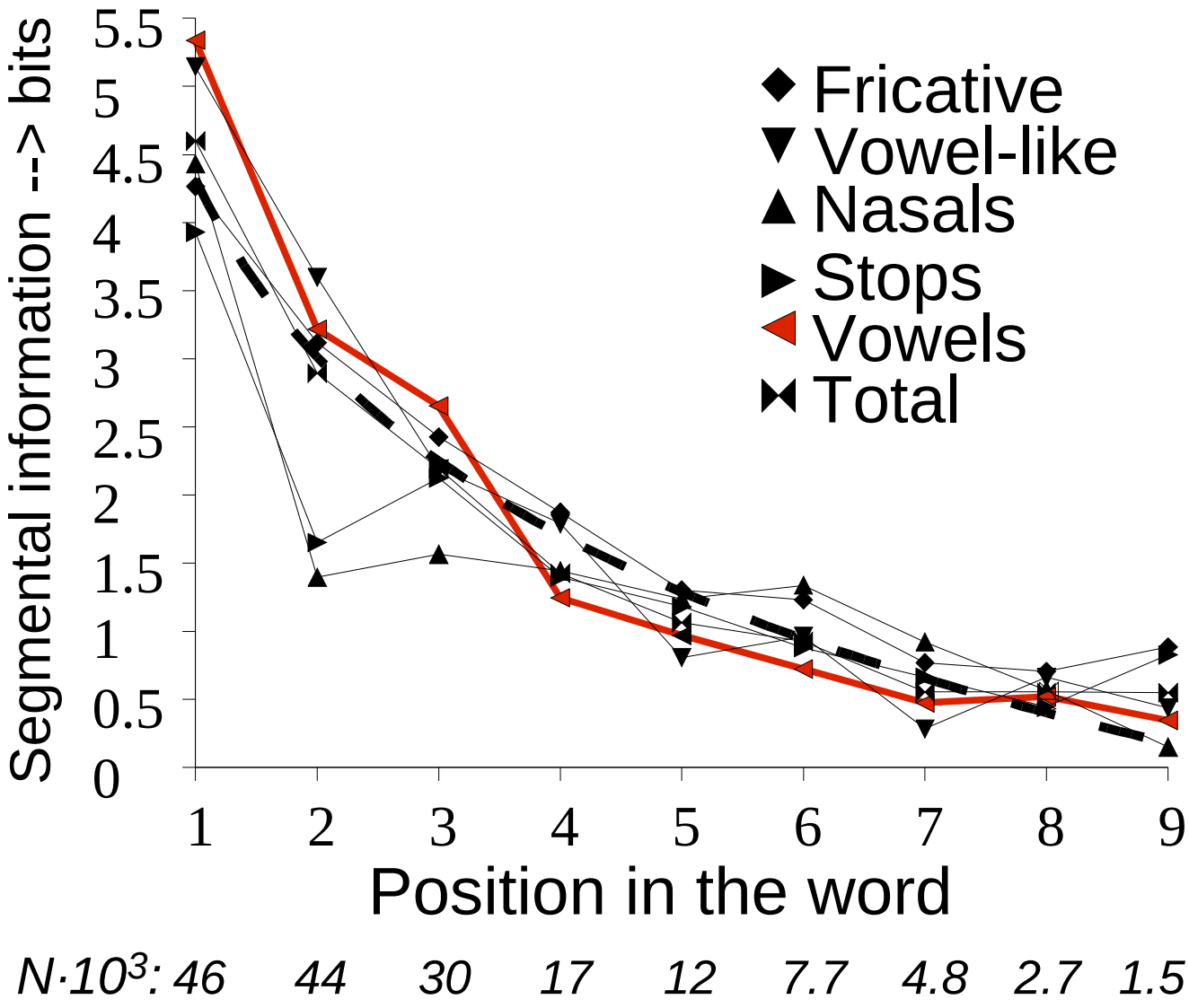
Spontaneous

- **Informal**: Elicited story about a vacation trip (face to face)
- **Retold**: Previously read story retold in an empty room

Read from a Cueing Screen

- **Text**: Long text
- **Sentences**: Isolated sentences
- **Pseudo Sentences**: Strings of randomly picked words

INFORMATION IN PHONEMES versus Position in the Word



Segmental Information versus Position in the Word grouped by Manner of Articulation

Preliminary Results

(see ICSLP2002 proceedings)

Redundancy and Reduction are Correlated (R~0.07, p<0.001)

But *Not* for:

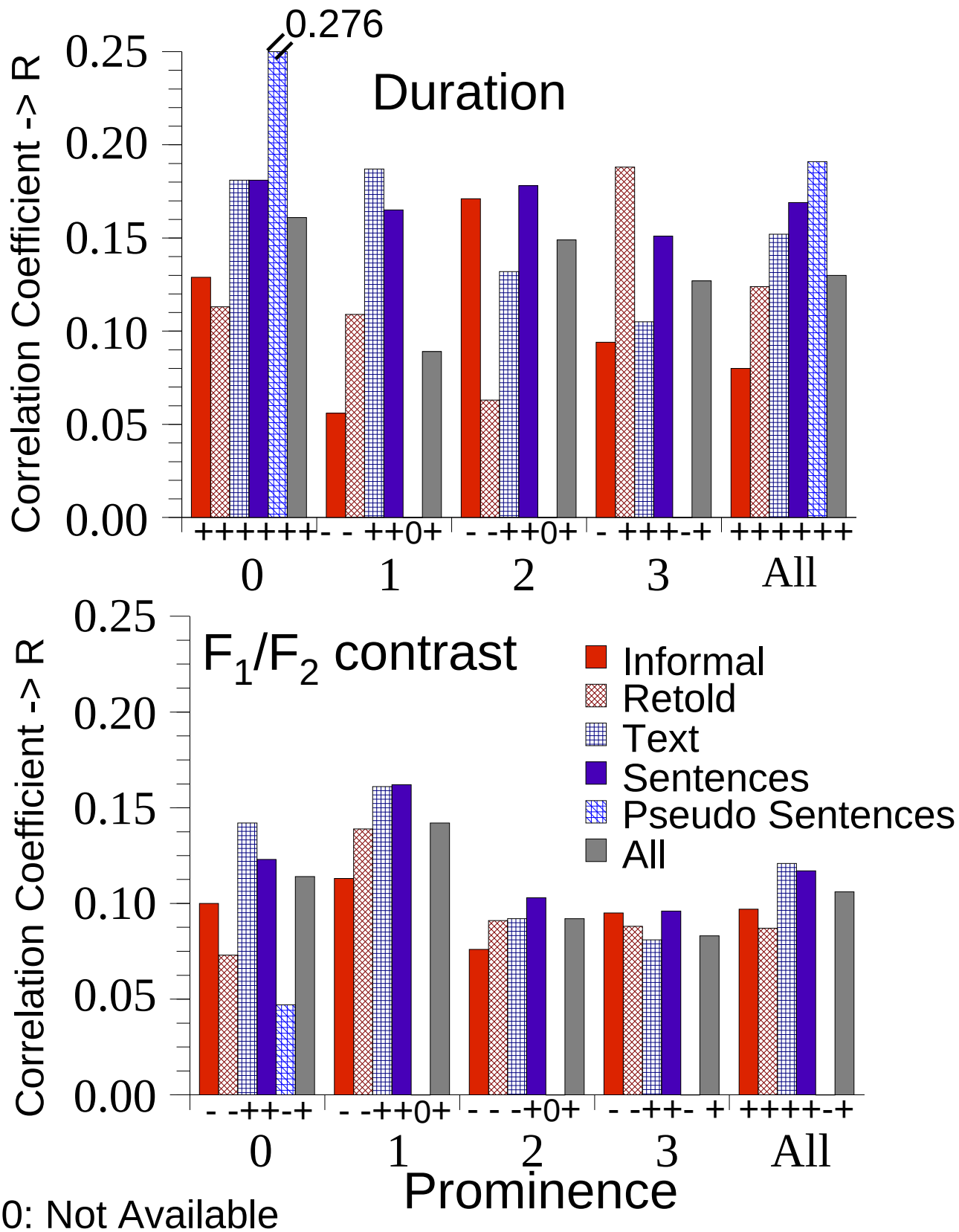
- Redundant Vowels ($I_s \leqslant 2$ bits)
  - *A Floor in Reduction*
- Low-Frequency (Rare) Words
  - *Context predicts Rare Words*
    (e.g., ***ocean*** after ***Pacific*** or ***Atlantic***)

Solution:

- Ignore Redundant Vowels
  ($I_s \leqslant 2$ bits)
- Correct for Predictability in Context, i.e., *Context Distinctiveness*
  (e.g., ***oceaan***: I_s=**16**, CD=**7.5**, diff=**8.6** bits)

SPEAKING STYLE

Reduction versus Information Content corrected for *Context Distinctiveness*



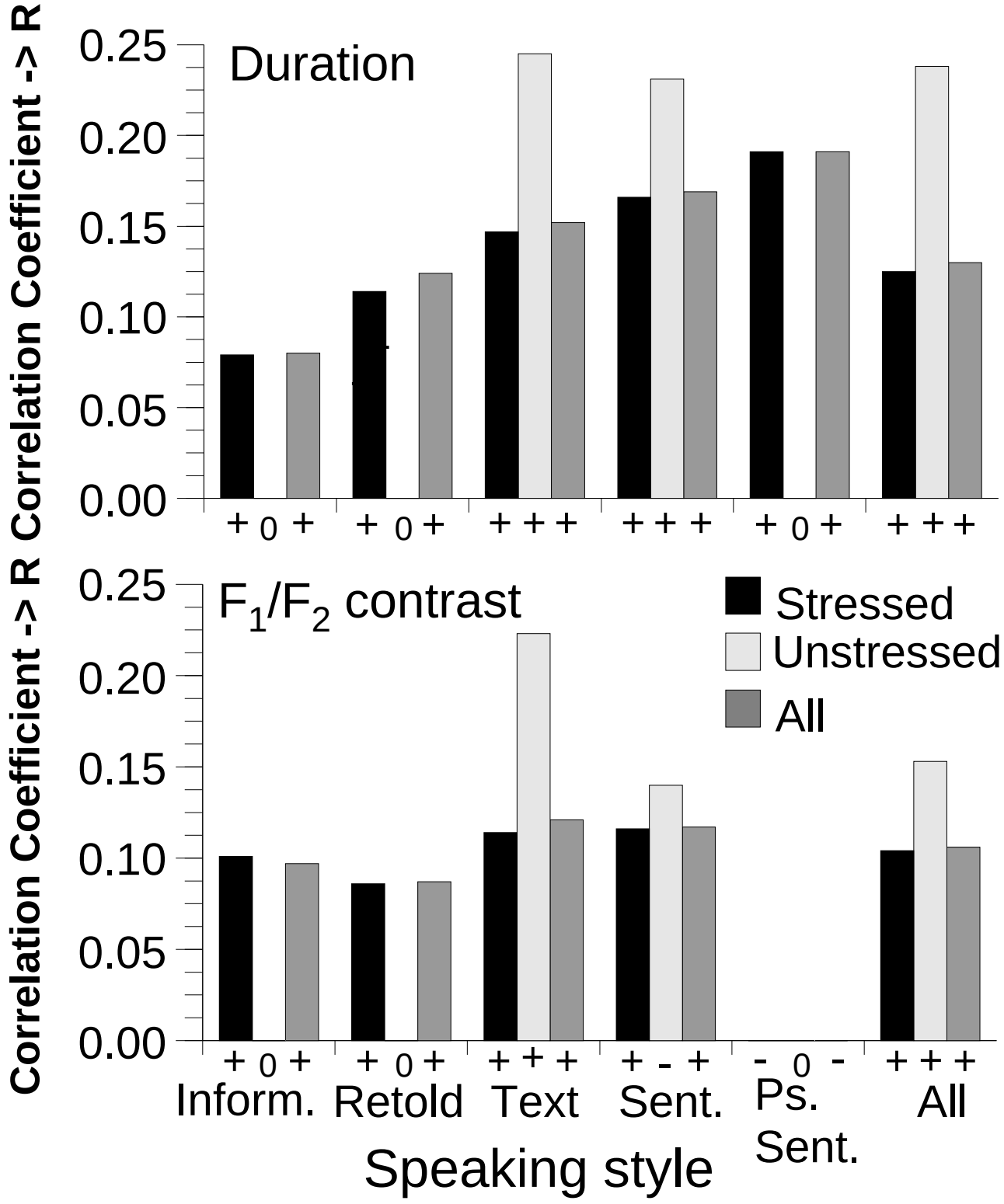
0: Not Available

+: $p < 0.001$, -: not significant

N=25,379

LEXICAL STRESS

Reduction versus Information Content
corrected for *Context Distinctiveness*



0: Not Available

+: $p < 0.001$, -: not significant

N=25,379

Discussion

- Acoustic Reduction correlates with Segmental Redundancy
- There is a Maximum Reduction for Redundant Vowels
- Word-Context is Accounted for i.e., *Context Distinctiveness*
- Strongest Effects for Read Speech
  (**but**: Prominence was modeled after Read Sentences)

Conclusions

- Reduction Increases when Vowels are more Redundant
- Vowel Production seems to be **Efficient at the Segmental Level**
- Holds for both **Duration** and **Spectral Contrast**
- Segmented Speech Corpora are Useful