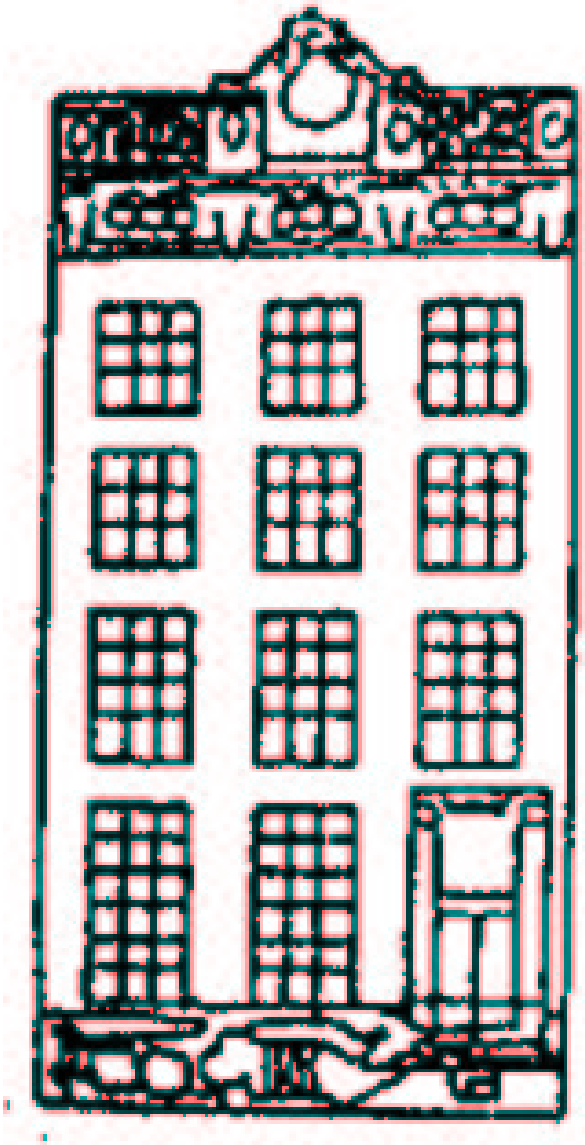
Evidence for Efficiency in vowel production

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***
- Predictability/Redundancy of words affects vowel reduction
- Word−frequency affects vowel reduction
- Does phoneme predictability affect vowel reduction?

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: The segmental information in bits

s: Phoneme segment

[word−onset]: preceeding segment sequence

Example: The relative importance of the /a:/ in /x@da:n/ (Dutch: ’gedaan’ English: ’done’)

Probability(/a:/ | /x@da:n/) =
Frequency(/x@da:/) = **14946/81360** = **0.184** » **2.44** bits
Frequency(/x@d*/)

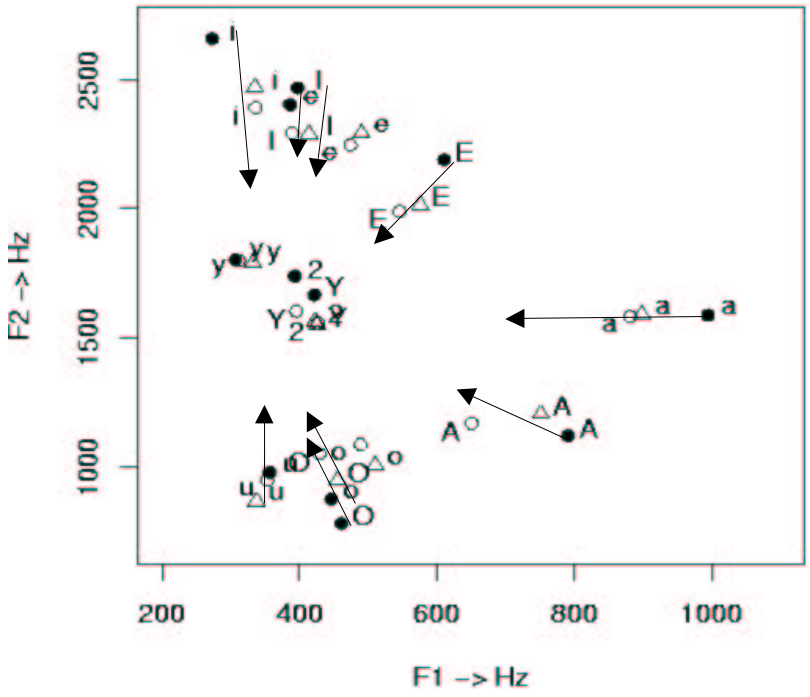
versus the /i/ in /x@dint/ (Dutch: ’gediend’ English: ’served’)

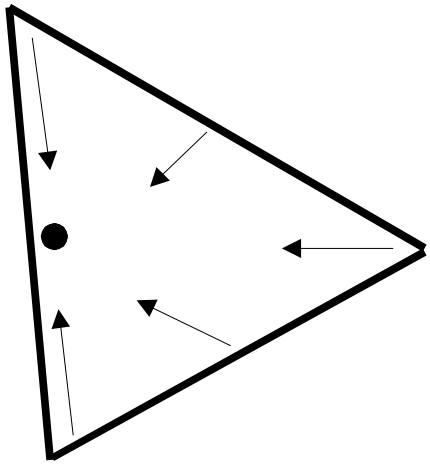
Probability(/i/ | /x@dint/) =
Frequency(/x@di/) = **1225/81360** = **0.015** » **6.05** bits
Frequency(/x@d*/)

Correlate this to measures of reduction

Acoustic measures of vowel reduction

- Duration
- F_1/F_2 contrast:





Take the distance to the "center of reduction" in semitones (this equalizes the variances in F_1 and F_2)

Factors influencing vowel reduction (in Dutch)

Factors to account for:

- Speaker Identity
- Vowel identity
- Speaking style
- Lexical stress
- Prominence

Use quasi−uniform subsets for calculating correlations

Automatic prominence assignment

Rules for prominence (Streefkerk, 2001):

- Each content word receives 2 marks
- Each word from the classes [Noun, Adjective, Numeral, Negation] receives an additional mark
- Polysyllabic content−words from the classes [Pronoun, Verb, Adverb] receive an additional mark
- The first content word in a sentence receives an additional mark (only implemented for the first 3 words)
- Each Noun preceded by an Adjective loses one mark

Function words receive 0, content words 1−4 marks. Only words with 0−3 prominence marks were frequent enough to be used.

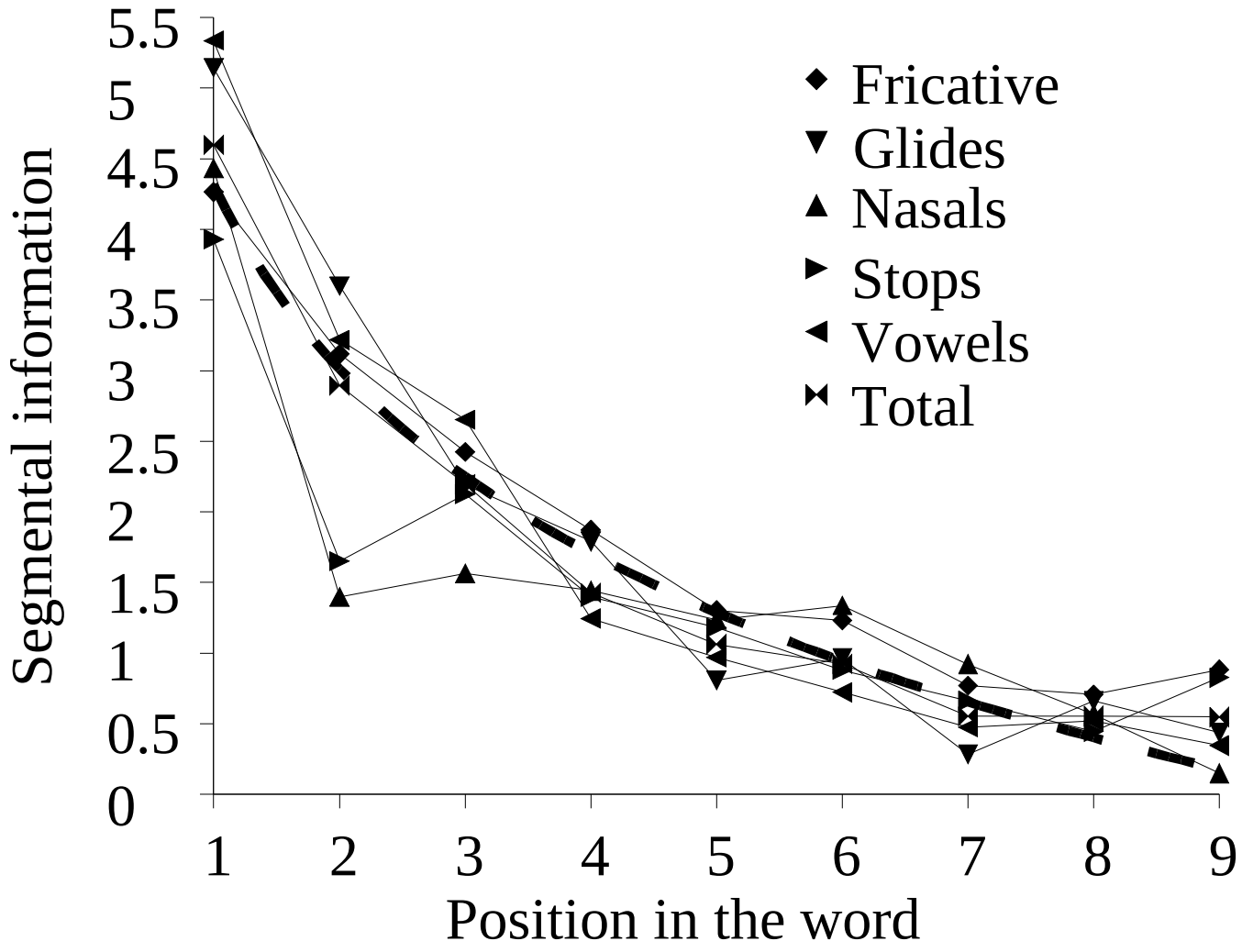
Speech

4 male + 4 female speakers

Speaking styles are:

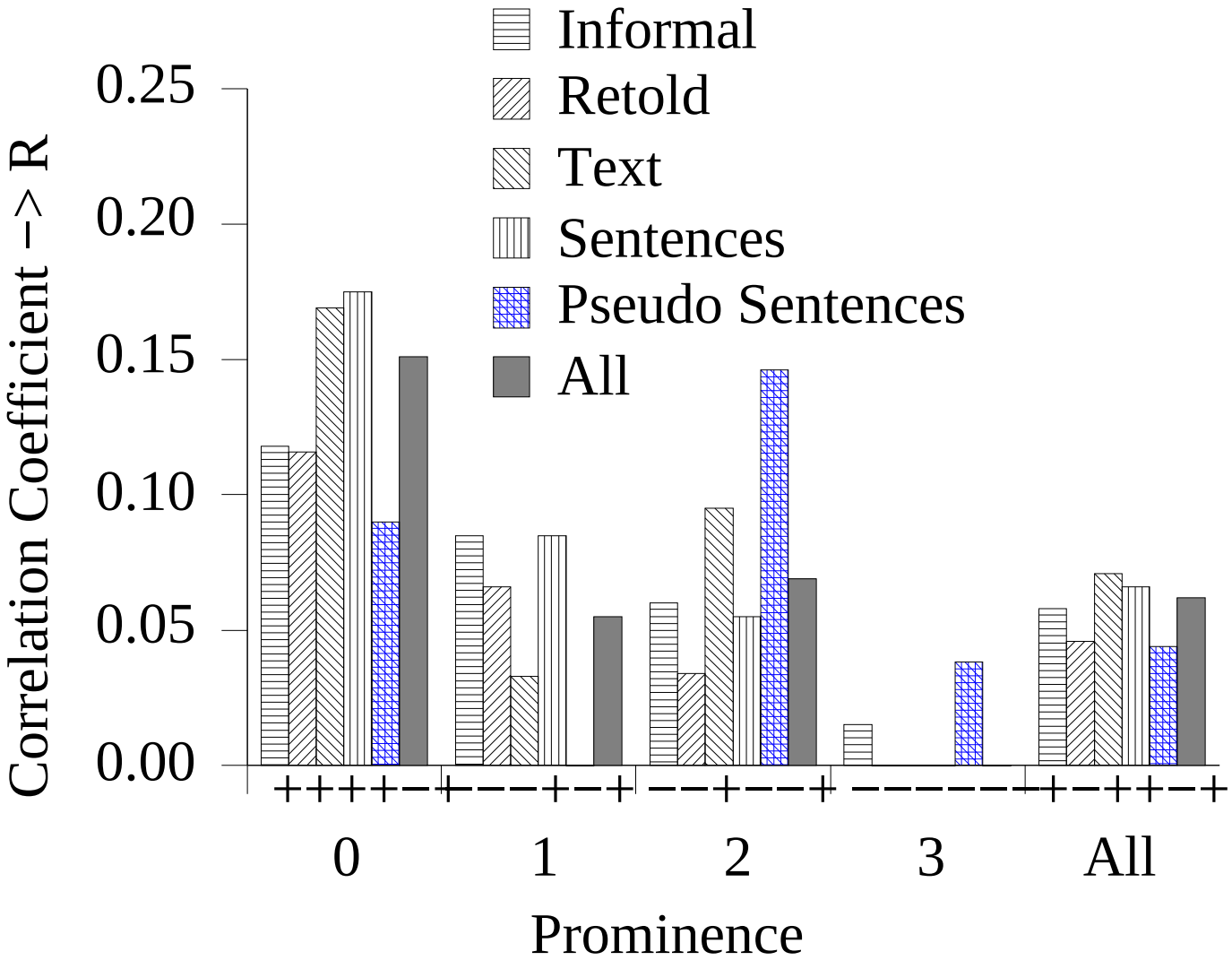
- Informal: An elicited story about a vacation trip told to an interviewer (face to face)
- Retold: A previously read story (a fixed fairy tale or the vacation trip) retold in an empty room
- Read: A long text read from a cueing screen
- Sentence: Isolated sentences read from a cueing screen
- Pseudo Sentence: Non−sentences, constructed by stringing randomly picked words, read from a cueing screen

INFORMATION IN PHONEMES versus position in the word



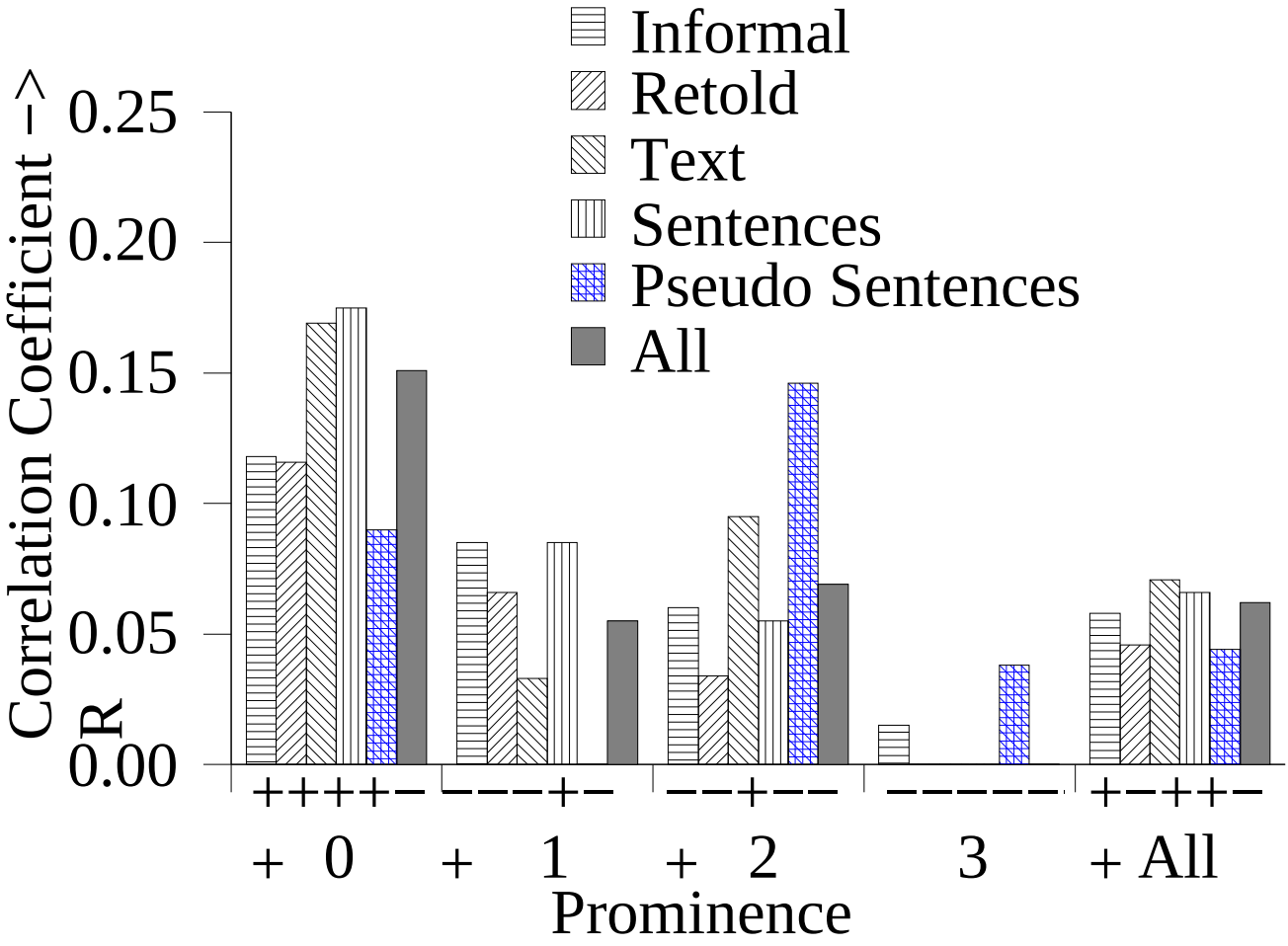
Relation between segmental information and the position in the word grouped by manner of articulation for comparison. The pooled values (Total) have been fitted with a logarithmic line (dashed line).

DURATION VERSUS INFORMATION CONTENT



Correlation strength between segmental information and vowel duration.
Plotted is a breakdown on style and prominence marks. Speaker, lexical stress, and vowel identity are also accounted for.
+: $p < 0.001$, −: not significant.
Total N =40,385 tokens

F_1/F_2 CONTRAST VERSUS INFORMATION CONTENT



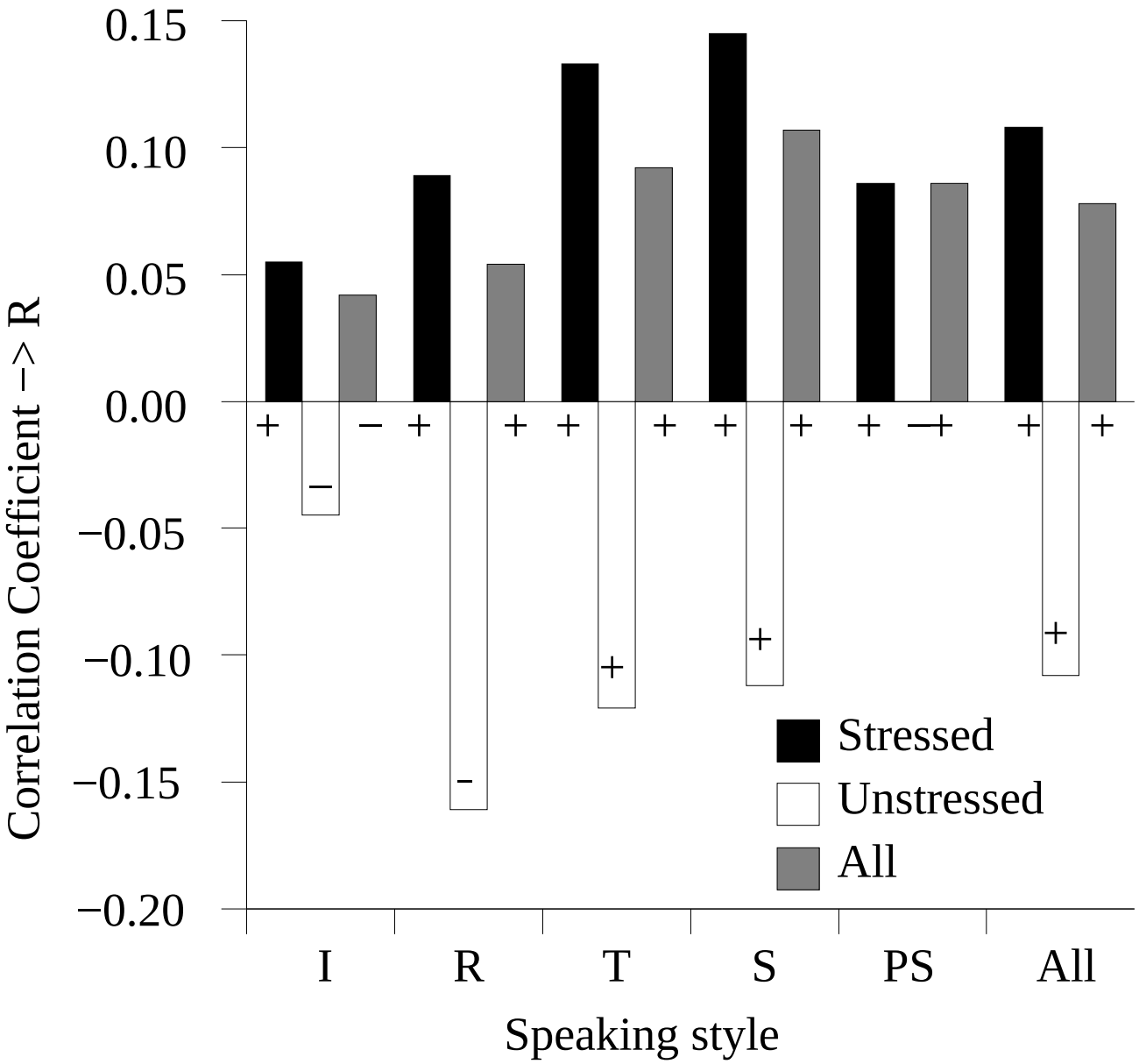
Correlation strength between segmental information and F_1/F_2 contrast.

Plotted is a breakdown on style and prominence marks. Speaker, lexical stress, and vowel identity are also accounted for.

+: $p < 0.001$, −: not significant.

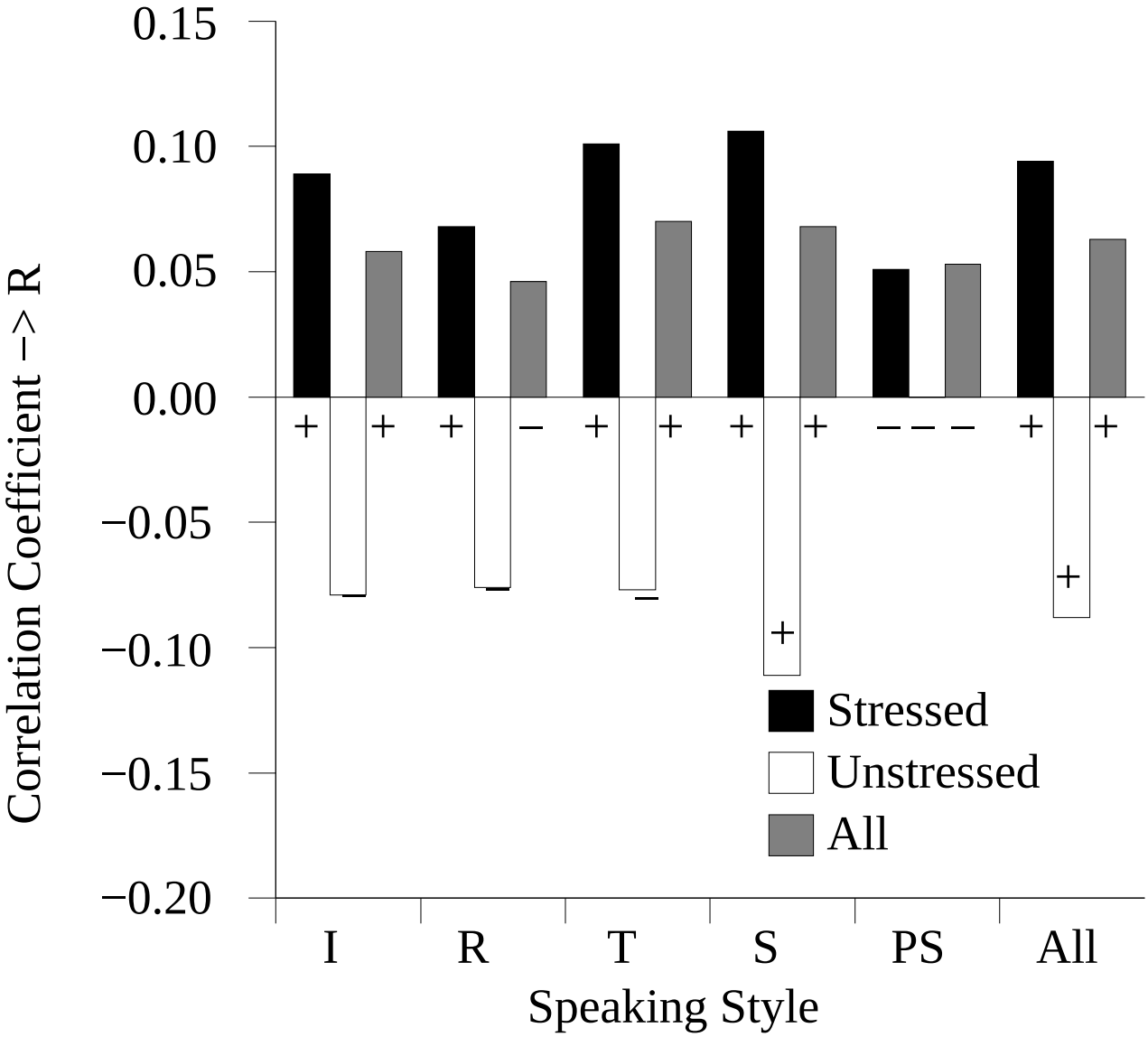
Total N =40,385 tokens

Correlation strength between segmental information and duration



Breakdown on speaking style and lexical stress.
+: $p < 0.001$, −: not significant.
Total N =40,385 tokens

Correlation strength between segmental information and F_1/F_2 contrast



Breakdown on speaking style and lexical stress.
+: $p < 0.001$, −: not significant.
Total N =40,385 tokens

Discussion

- Segmental redundancy correlates with acoustic reduction
- The effects are strongest for read speech (**but**: prominence marks were modeled after read sentences)
- The anomalous semantic content of the pseudo−sentences might interfere with normal speech planning
- The effect of redundancy is strongest on function words (prominence 0) and lowest on words with the highest prominence markings (the difference disappears when we repeat the analysis on only the high−frequency words)
- This suggests that the processing demands for assembling (low−frequency) words on−line can interfere with efficiency in speaking.

Conclusions

- Reduction increases when phonemes are more redundant: Speech production seems to be efficient at the segmental level
- This holds for both duration and spectral contrast
- Segmented speech corpora are useful