

Demand and Supply in the Communication Process: The Case of Lexical Richness and Phonological Features

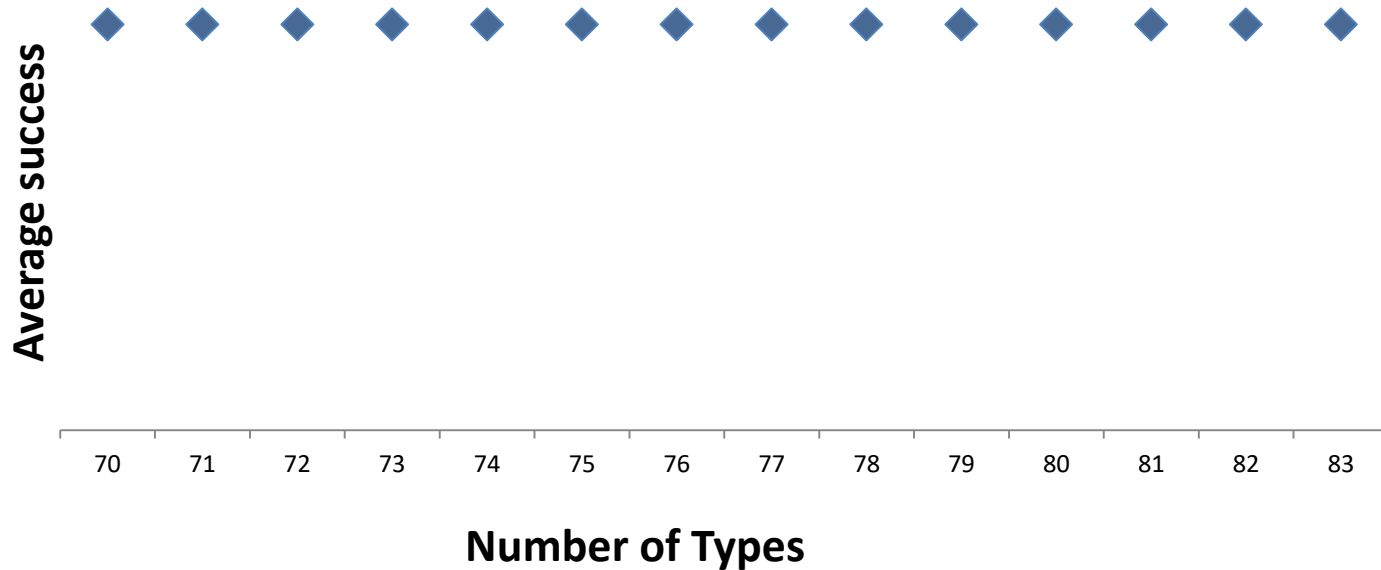
Jiří Milička & Alžběta Růžičková

Introduction

- Features of successful texts
- Lexical variables
- Phonological variables

Hypothesis 0 – No Relation

There is no relation between the lexical richness and the text success at all.



Hypothesis 1

The lexically less rich texts are more popular than the lexically richer ones.

Lower repertoire of types

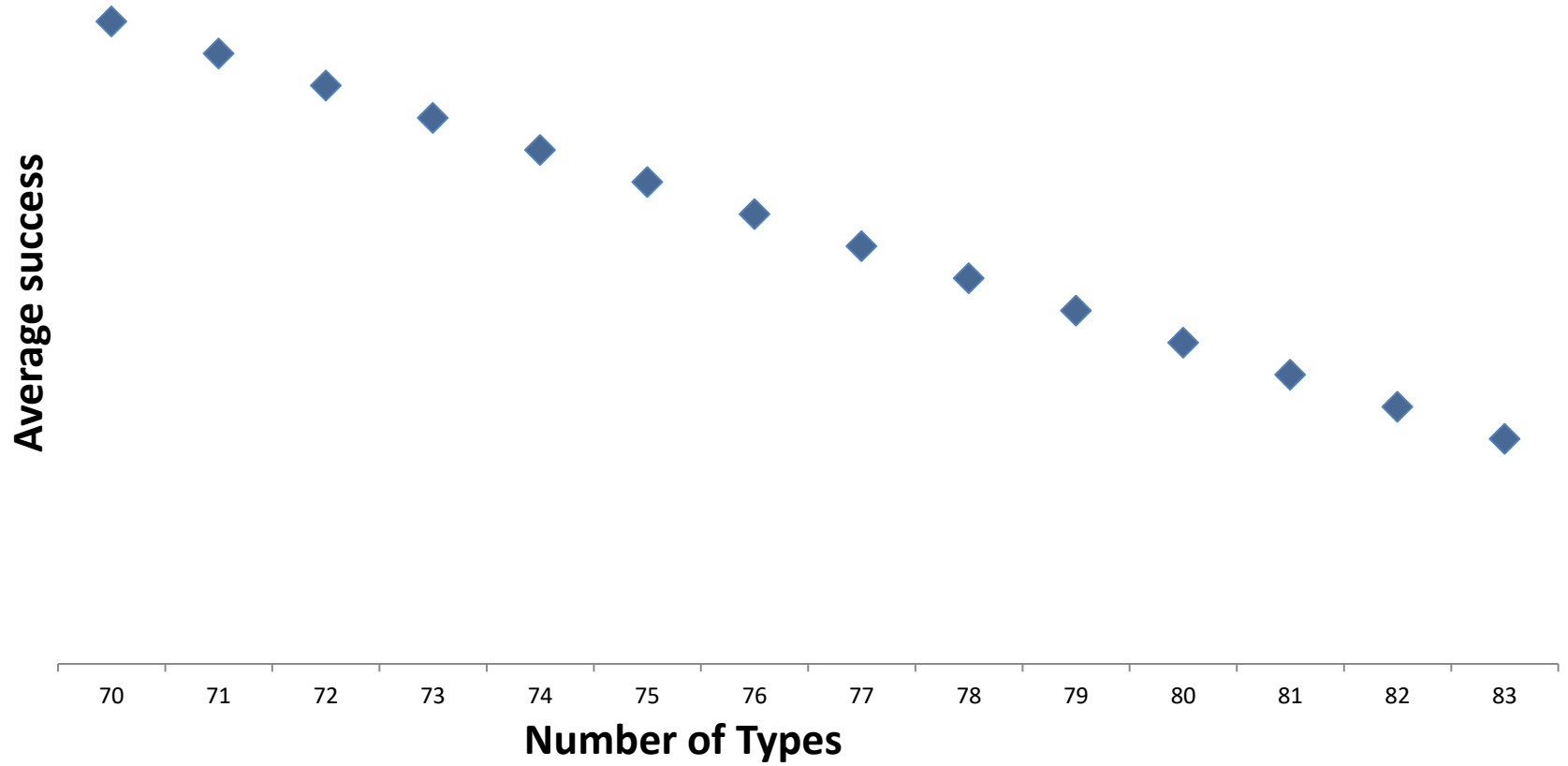


lower entropy rate



lower rate of errors in predicting following word

Hypothesis 1

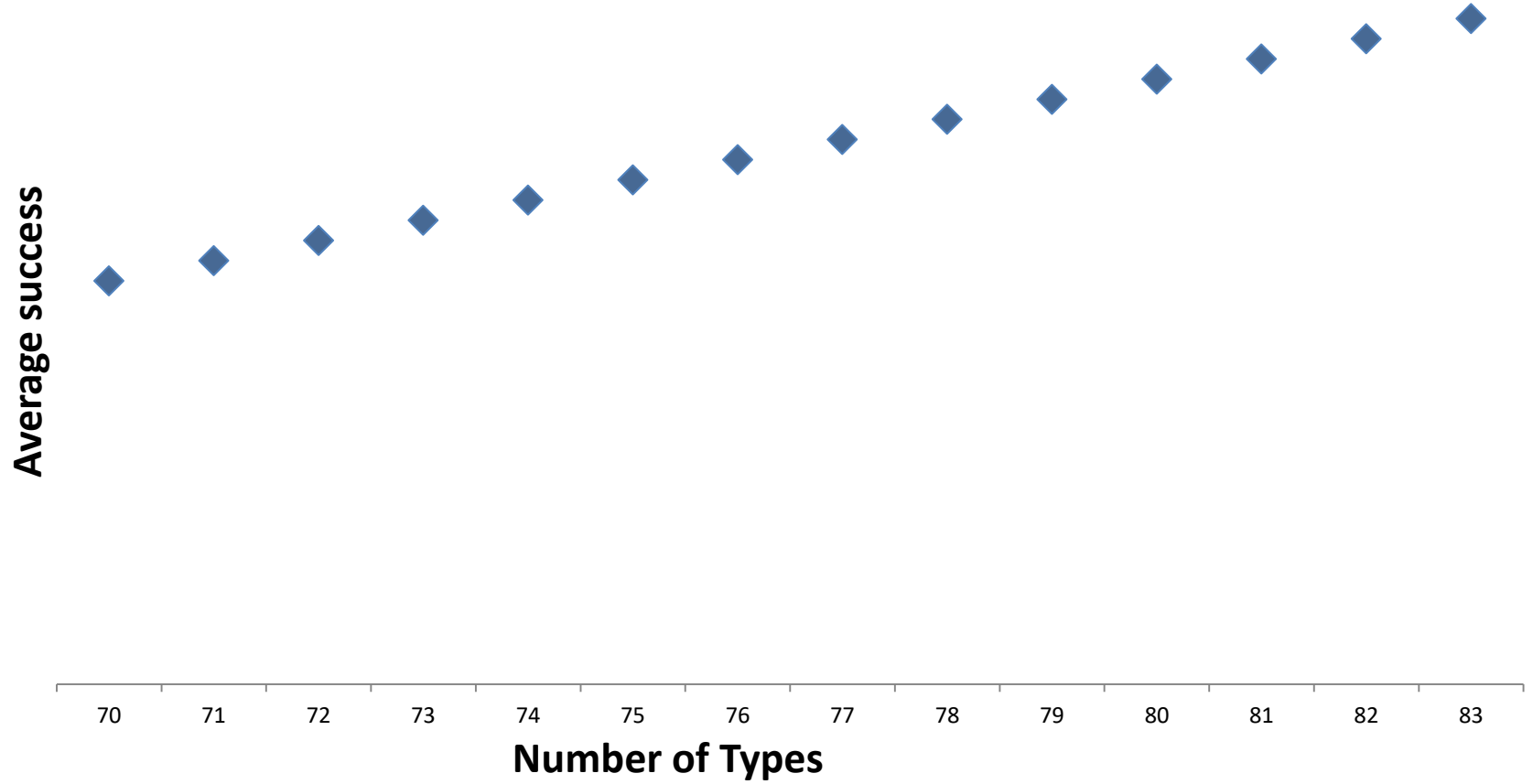


Hypothesis 2

The lexically richer texts are more popular than the lexically simpler ones.

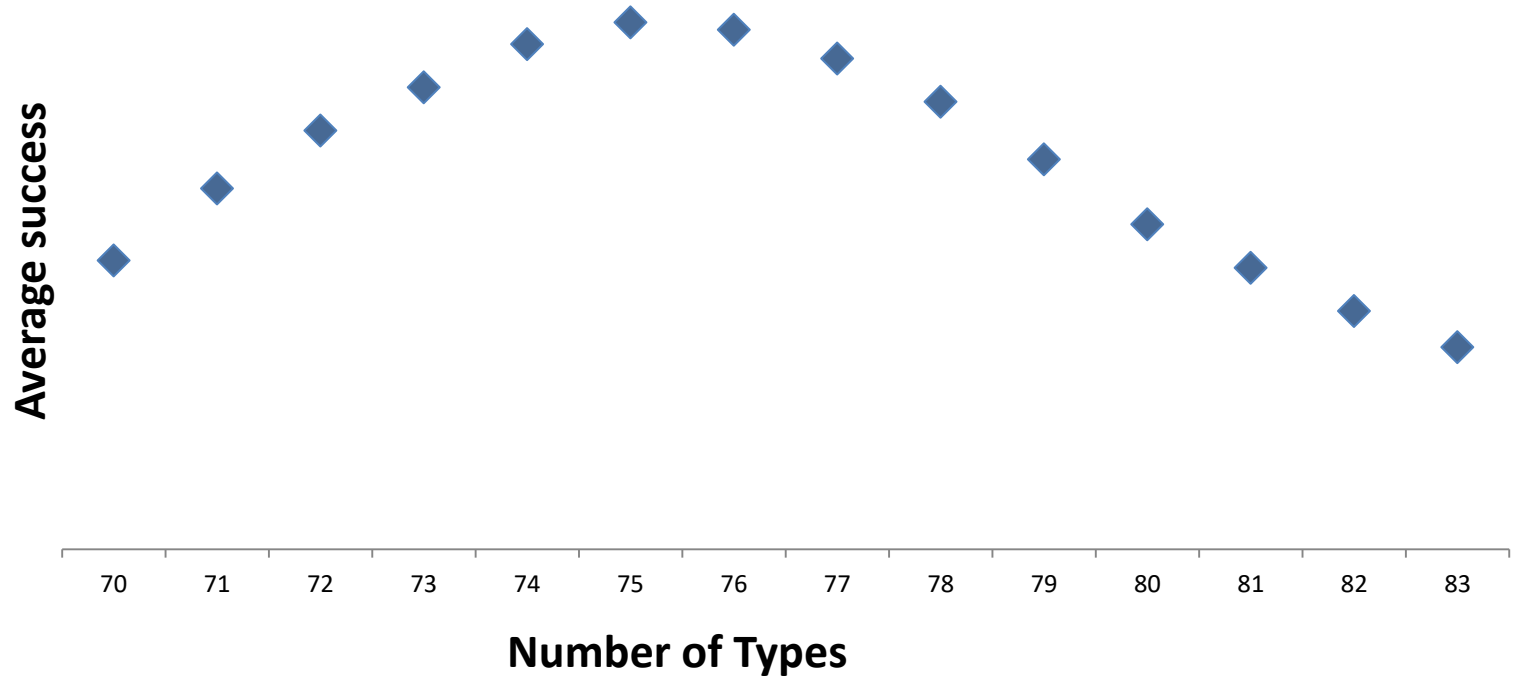
- A. Recipients like language proficiency;
- B. it is difficult to attain the lexical richness level that is ideal for the recipient.

Hypothesis 2



Hypothesis 3

- The mixture of the 2 previously mentioned effects



Data

- blog.idnes.cz
- 10 year span
- 344,000 texts (181,450,572 tokens)
- 517,819,018 views
- 2787 females (68280 texts); 5655 males (213602 texts).
- People of all age categories (roughly guessing according to the pictures of authors)
- Blogs without a specific author (i.e. institutions, anonymous...) were omitted.

Method



Method



Likes per view

0.05

0.04

0.01

0.1

0.08

0.001

Method



Likes per view	0.05	0.04	0.01	0.1	0.08	0.001
Proportion of vowels	42.5	42.2	42.1	41.0	41.6	41.1

Method



Likes per view

0.05 0.04 0.01 0.1 0.08 0.001

Proportion of vowels

42.5 42.2 42.1 41.0 41.6 41.1

Bootstrap resampling
(to get confidence intervals)

Method

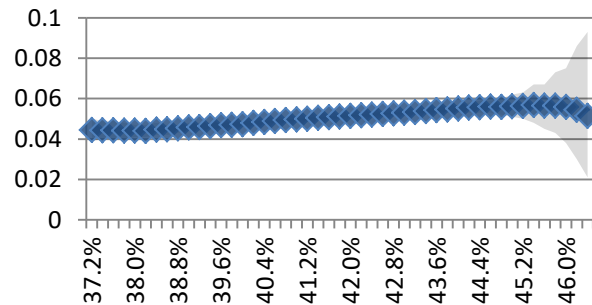


Likes per view

0.05 0.04 0.01 0.1 0.08 0.001

Proportion of vowels

42.5 42.2 42.1 41.0 41.6 41.1



Bootstrap resampling
(to get confidence intervals)

Method

- We need a variable that is text length independent

Method

Long text



tendency towards average values

Method

Long text  tendency to succeed



tendency towards average values

Method

tendency to succeed



average values

Method

tendency to succeed



average values

Method

- We need a variable that is text length independent
- And the deviation of the variable should be also text length independent

Method

Likes per view

0.05

0.04

0.01

0.1

0.08

0.001



Method

Likes per view

0.05

0.04

0.01

0.1

0.08

0.001



Method

Likes per view

0.05

0.04

0.01

0.1

0.08

0.001



43.7

42.2

41.1

40.0

40.6

43.1

Method

Likes per view

0.05

0.04

0.01

0.1

0.08

0.001



40.0 → 0.1

40.6 → 0.08

41.1 → 0.01

42.2 → 0.04

43.1 → 0.05

43.7 → 0.001

Method

Likes per view

0.05

0.04

0.01

0.1

0.08

0.001



40.0 → 0.1

40.6 → 0.08

41.1 → 0.01

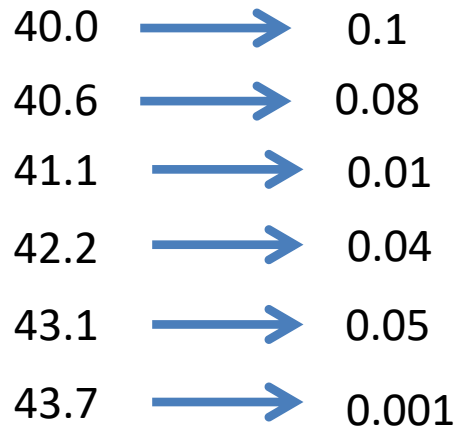
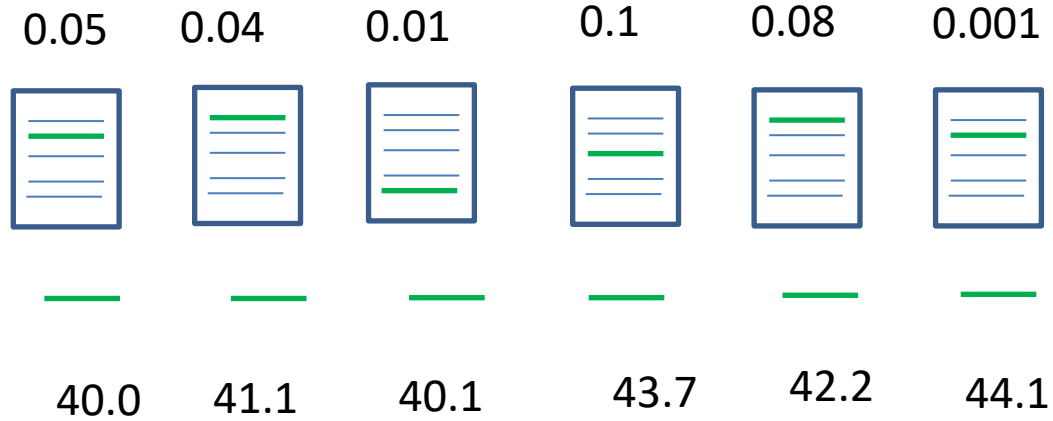
42.2 → 0.04

43.1 → 0.05

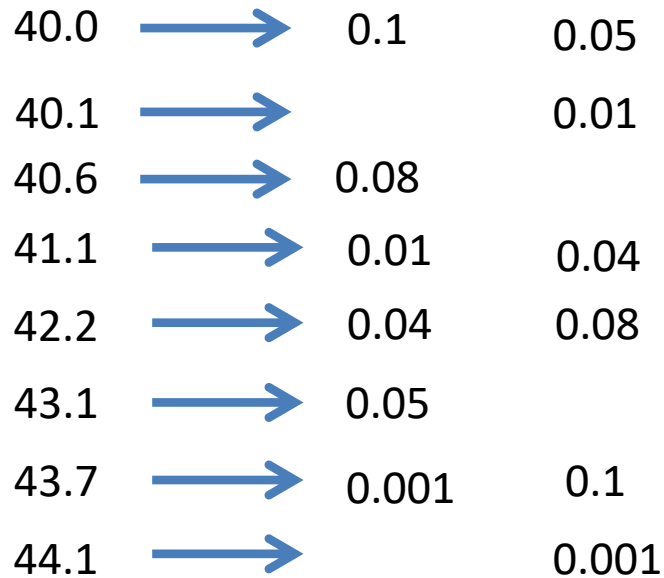
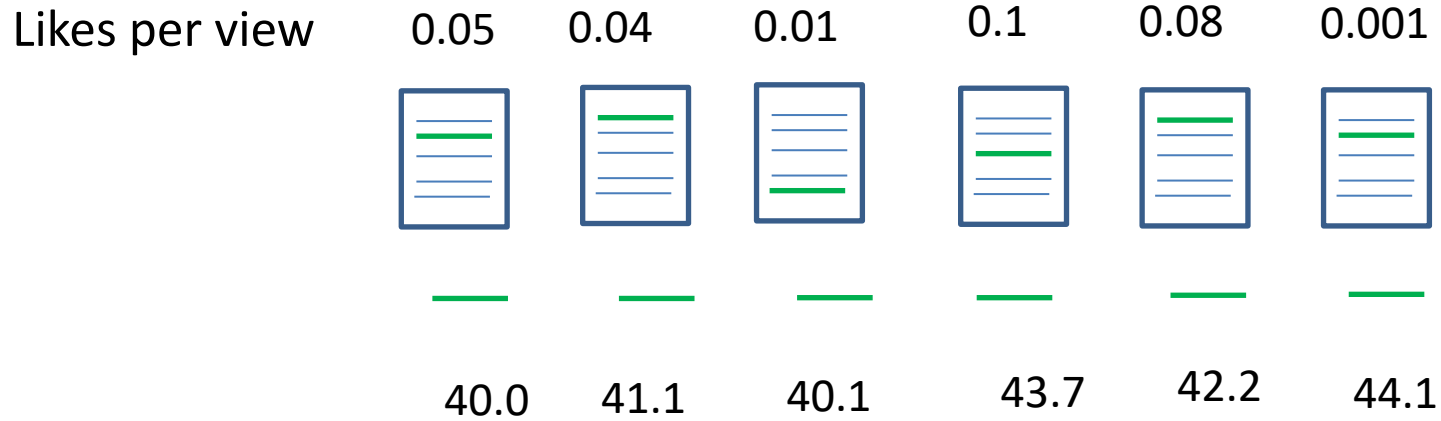
43.7 → 0.001

Method

Likes per view

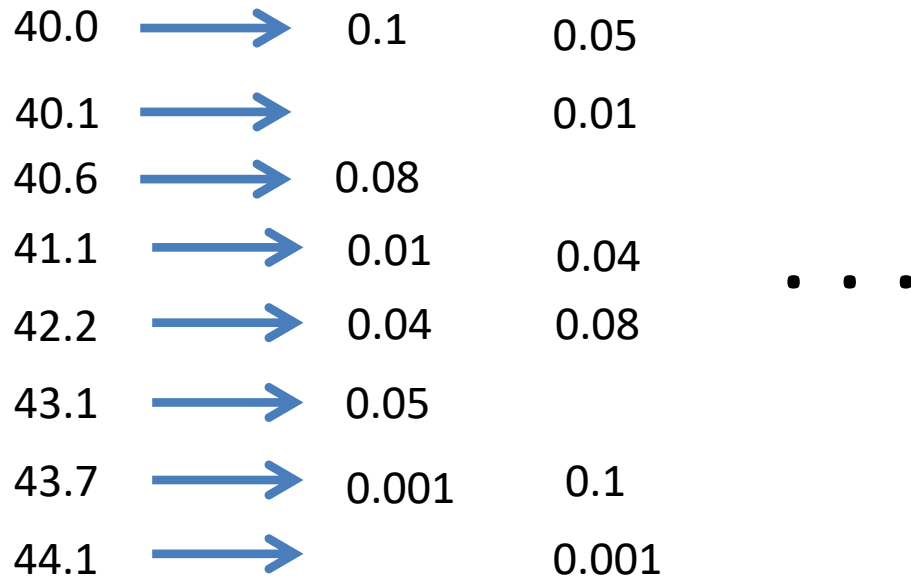
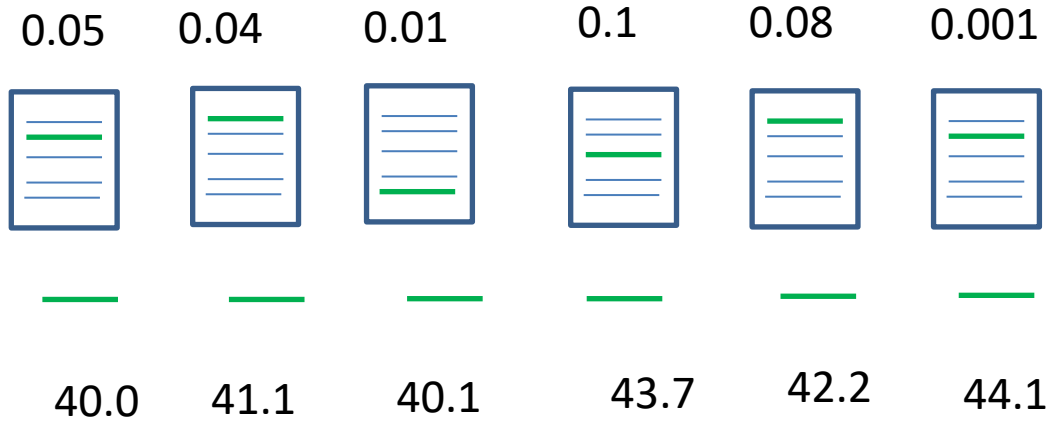


Method



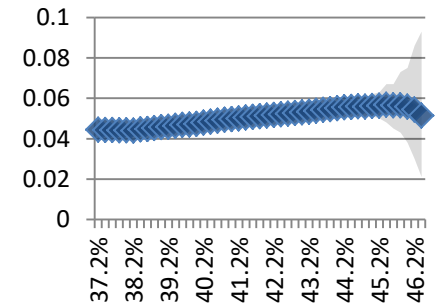
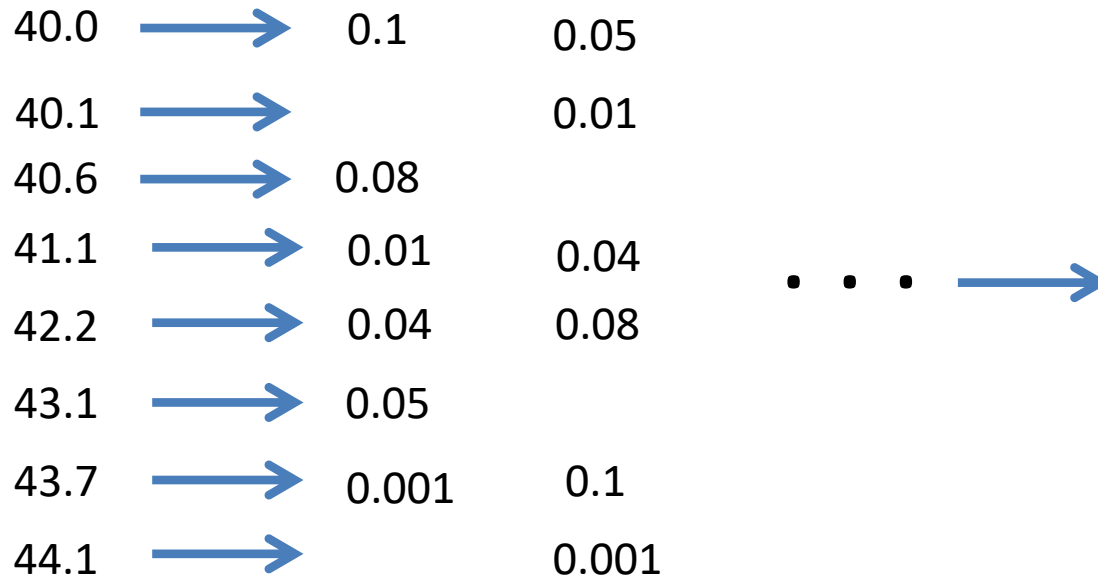
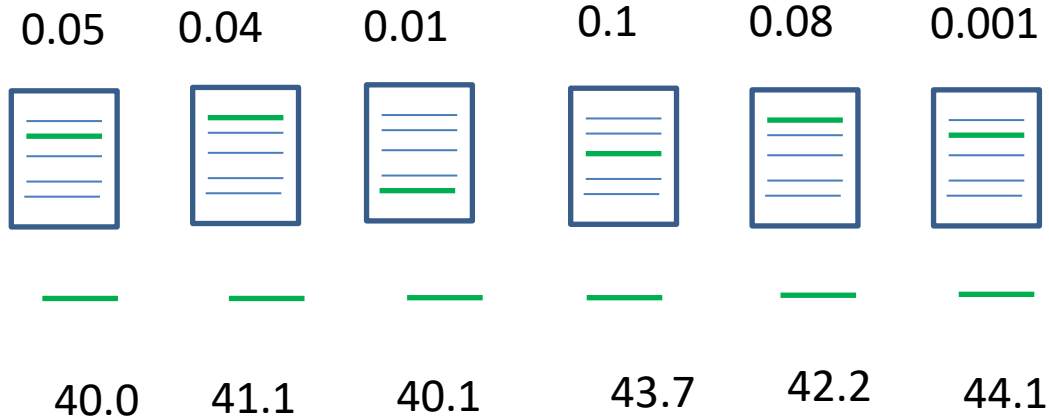
Method

Likes per view

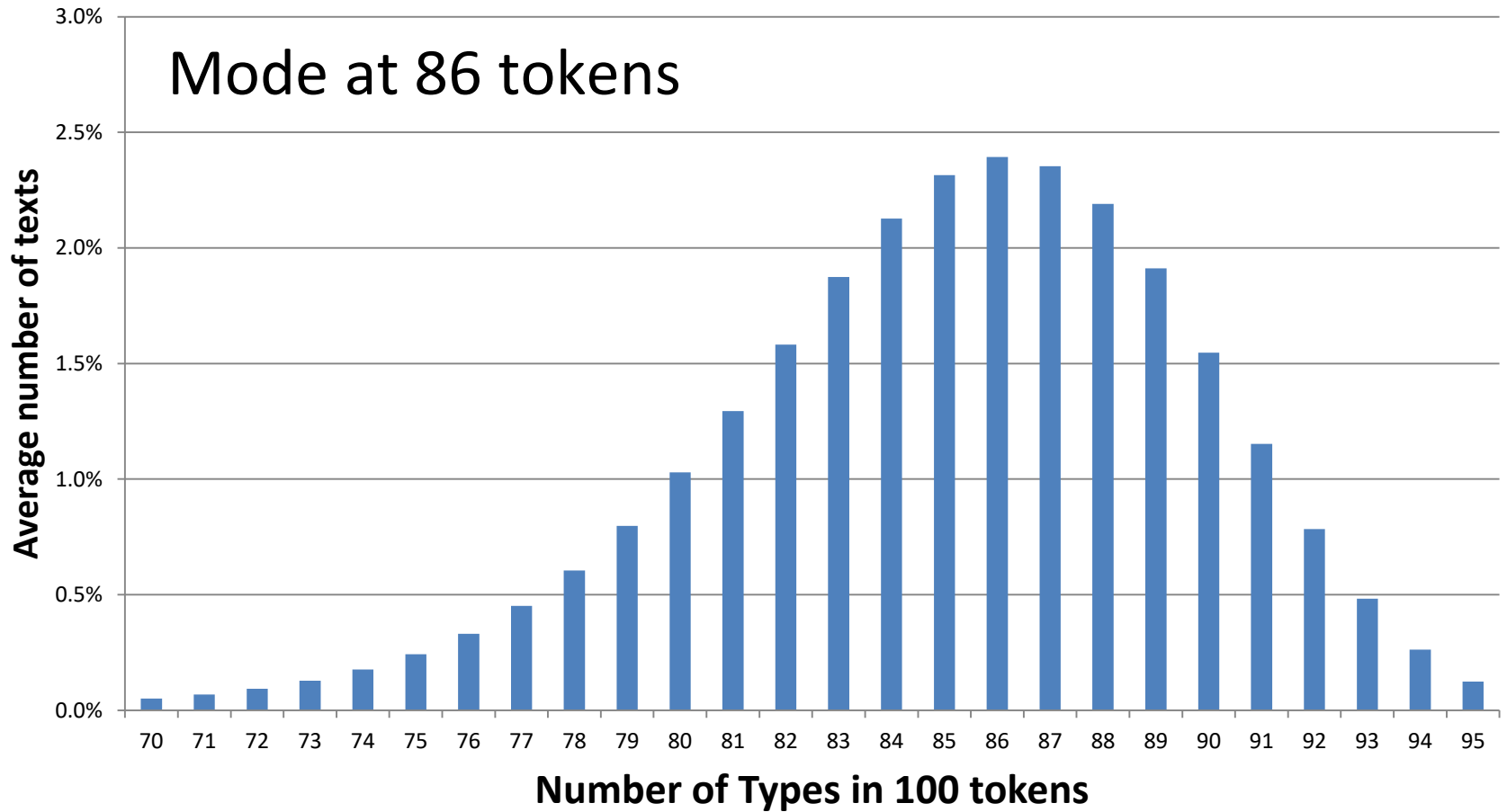


Method

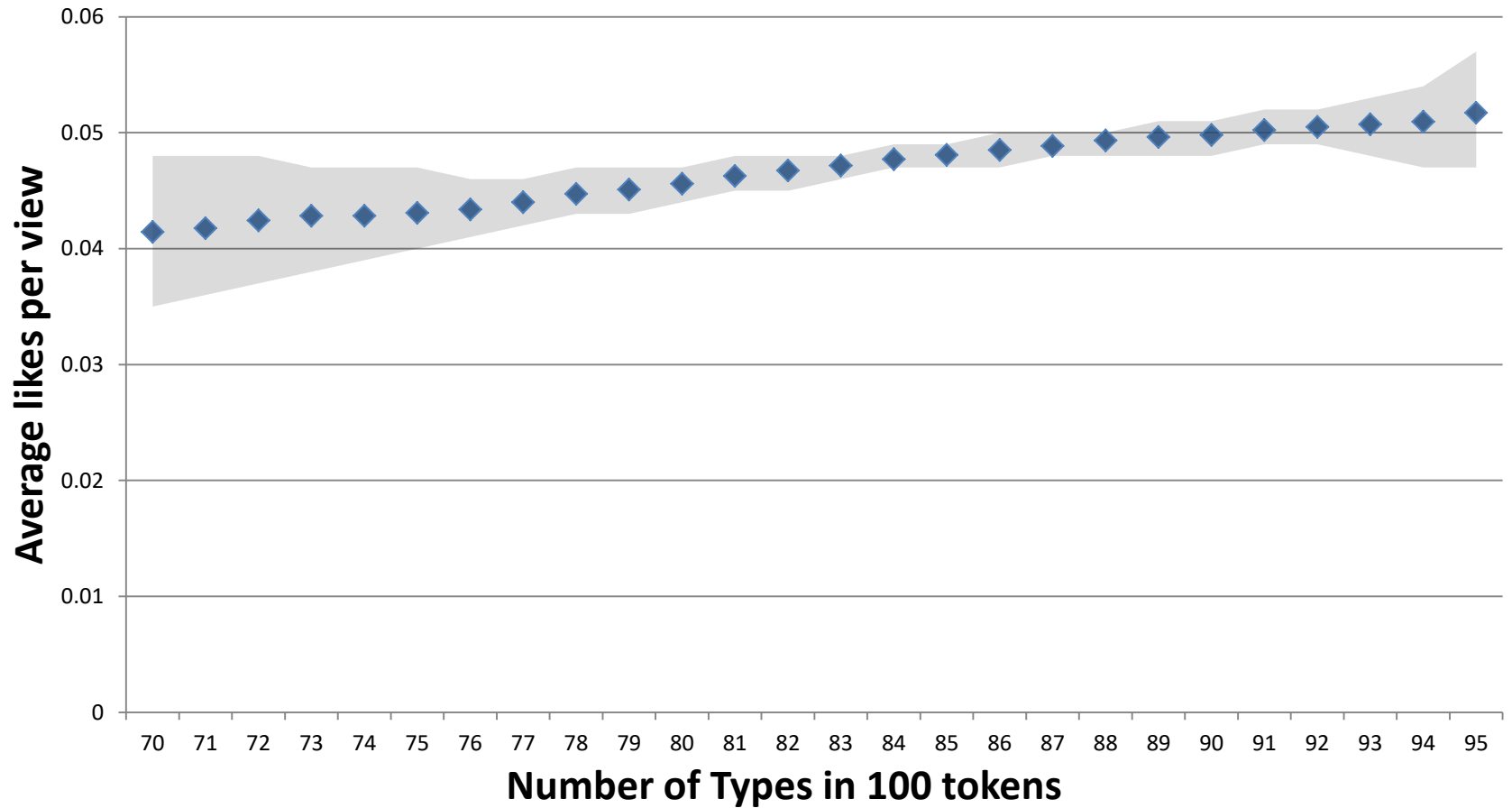
Likes per view



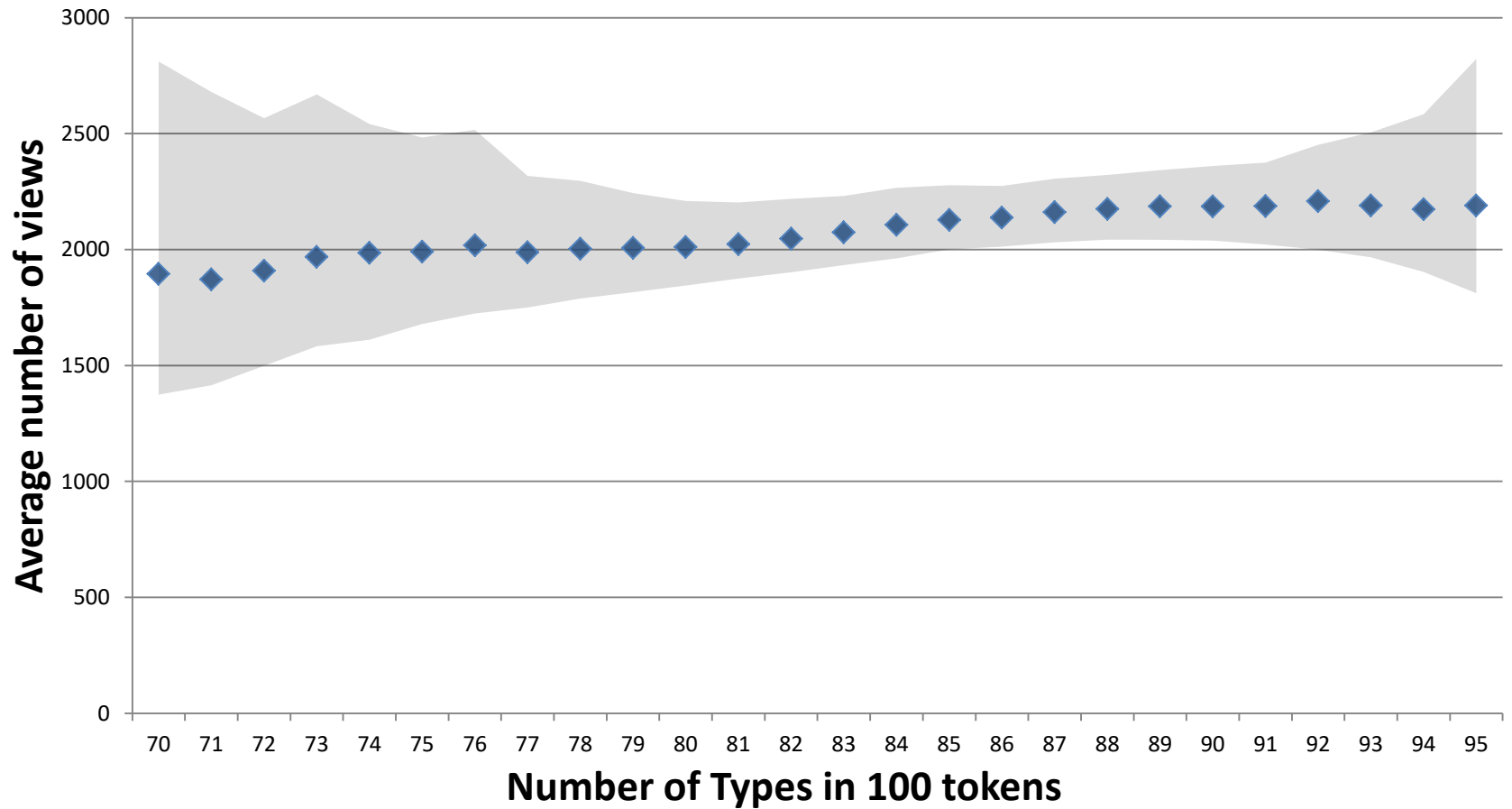
Number of Types



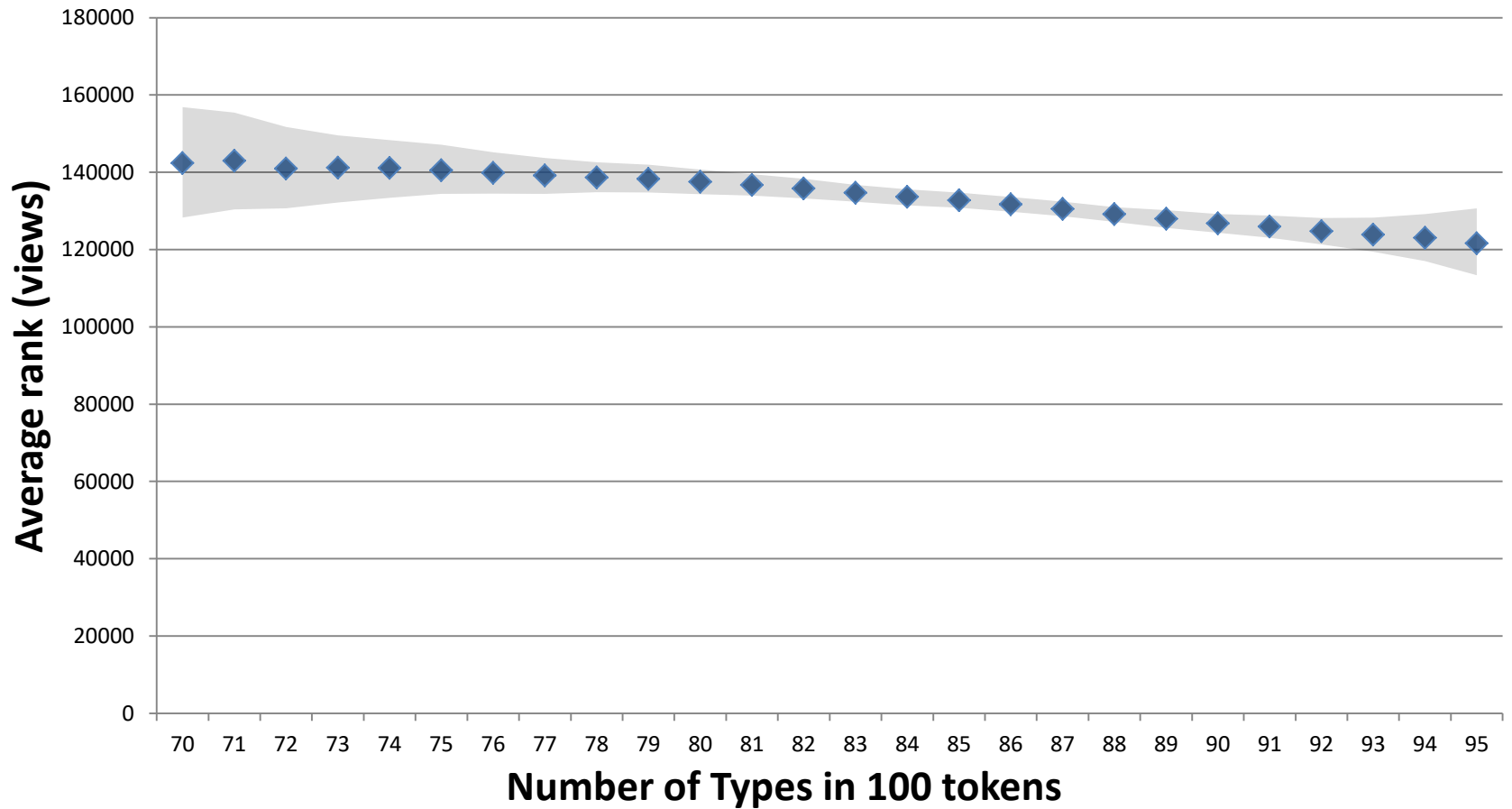
Number of Types



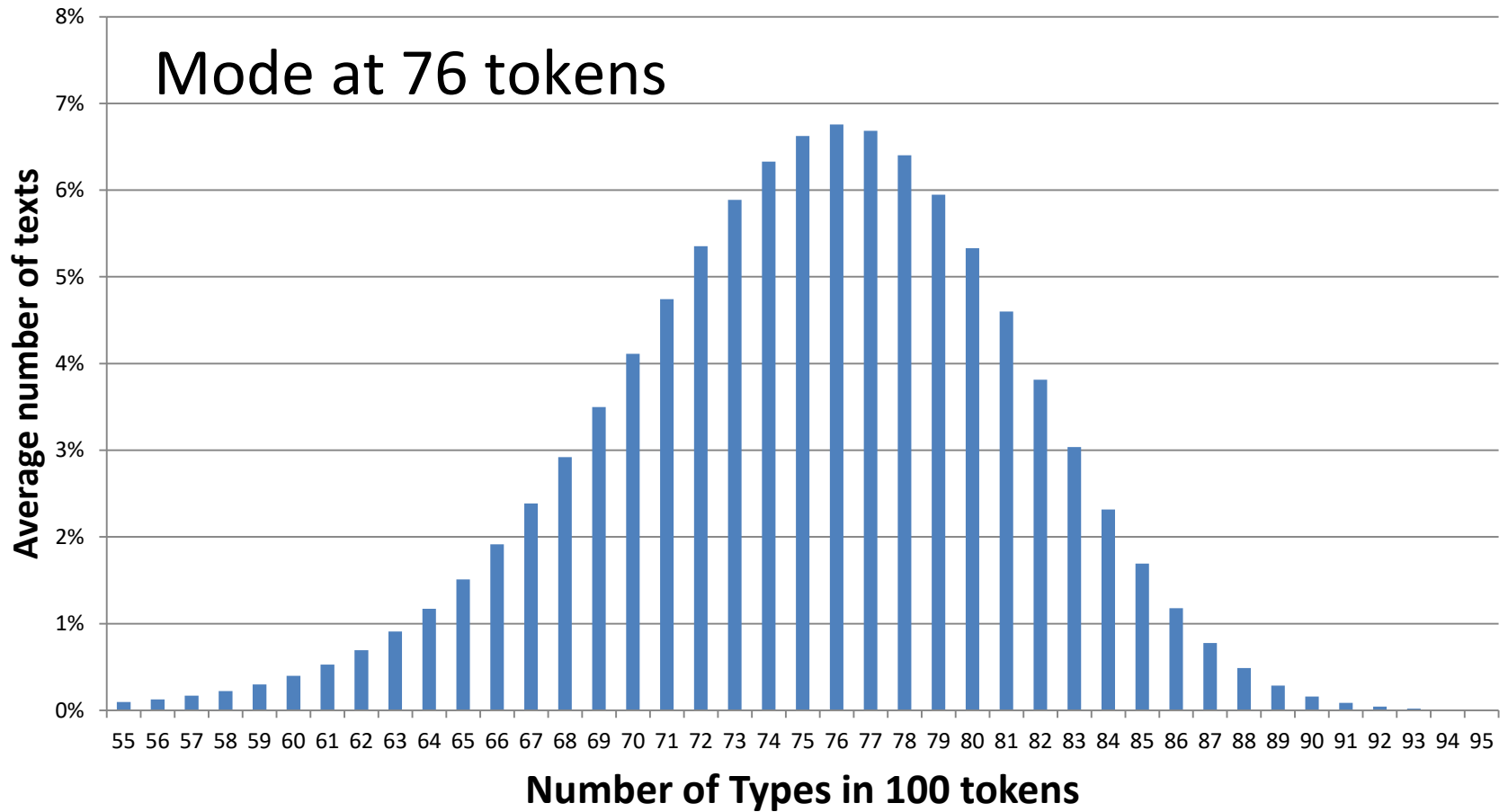
Number of Types



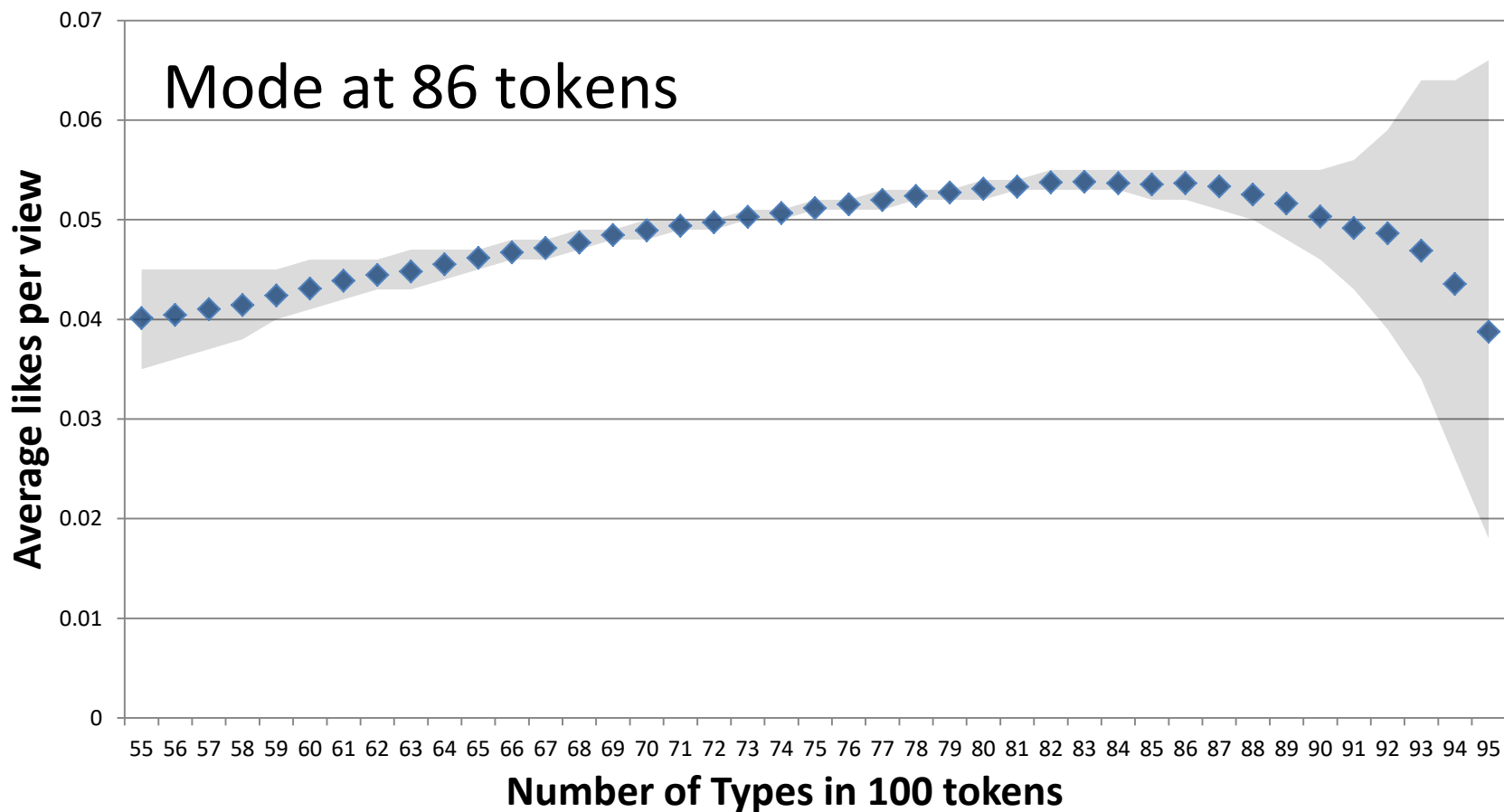
Number of Types



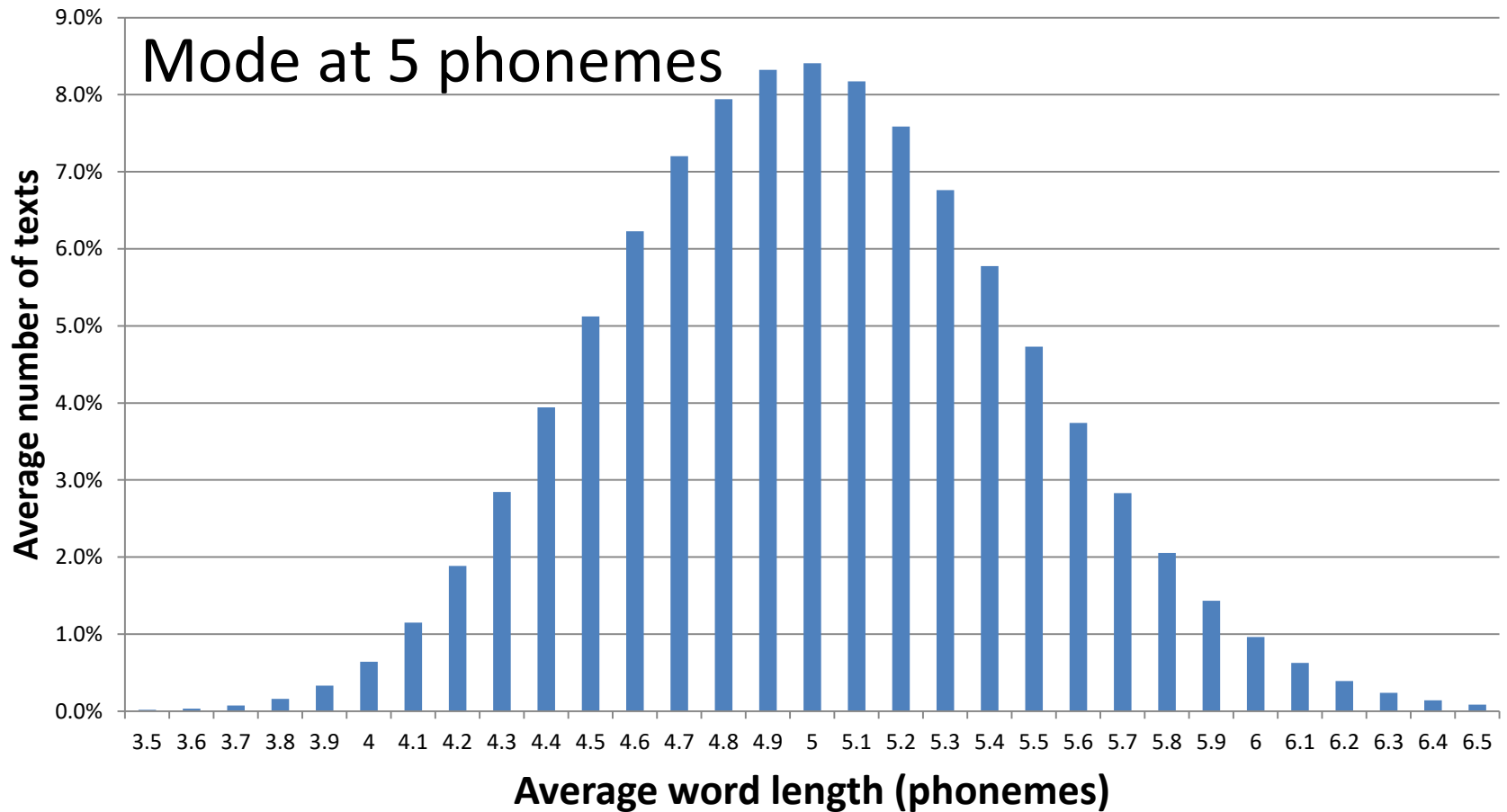
Number of Types (Lemmatized)



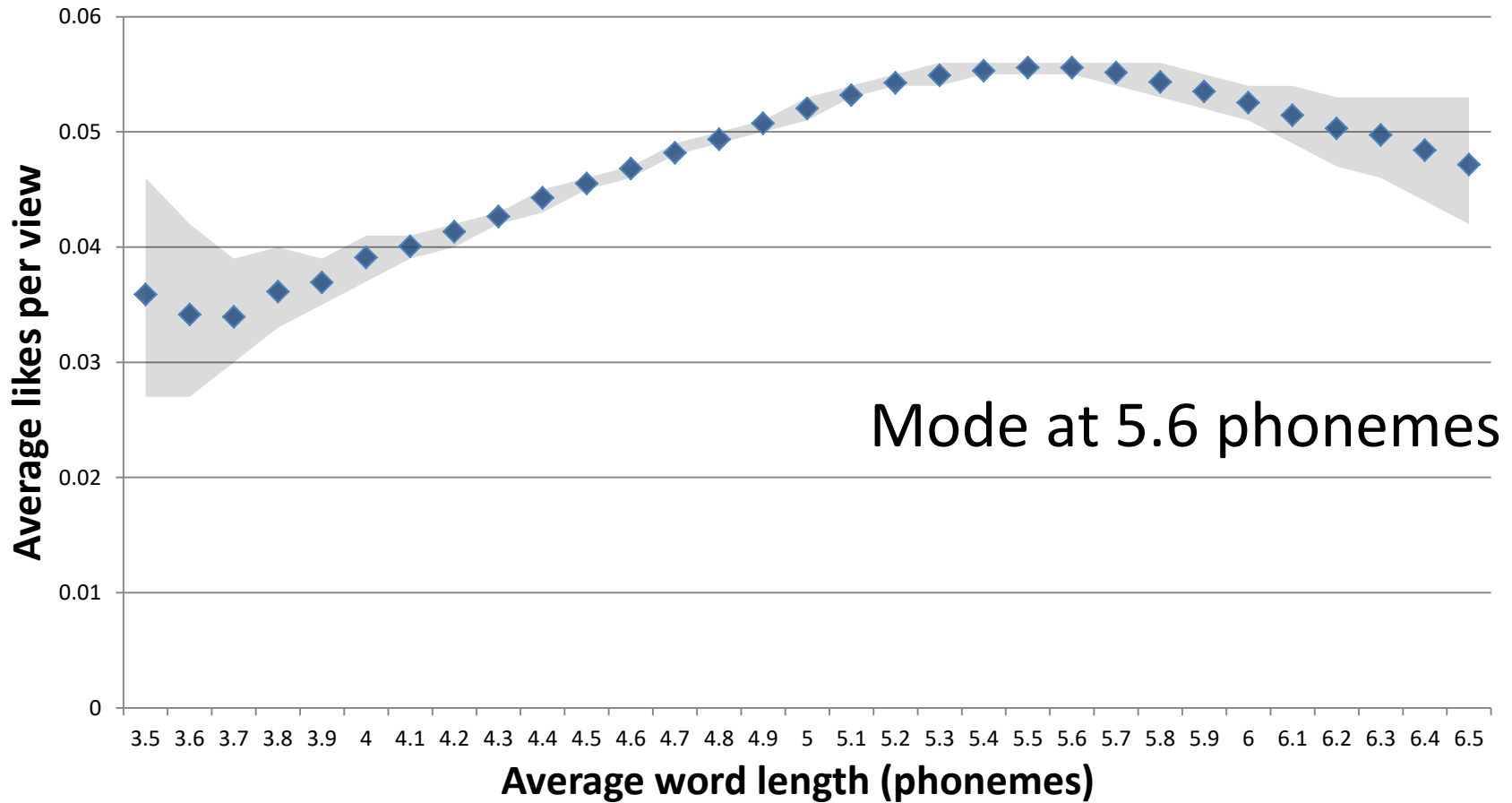
Number of Types (Lemmatized)



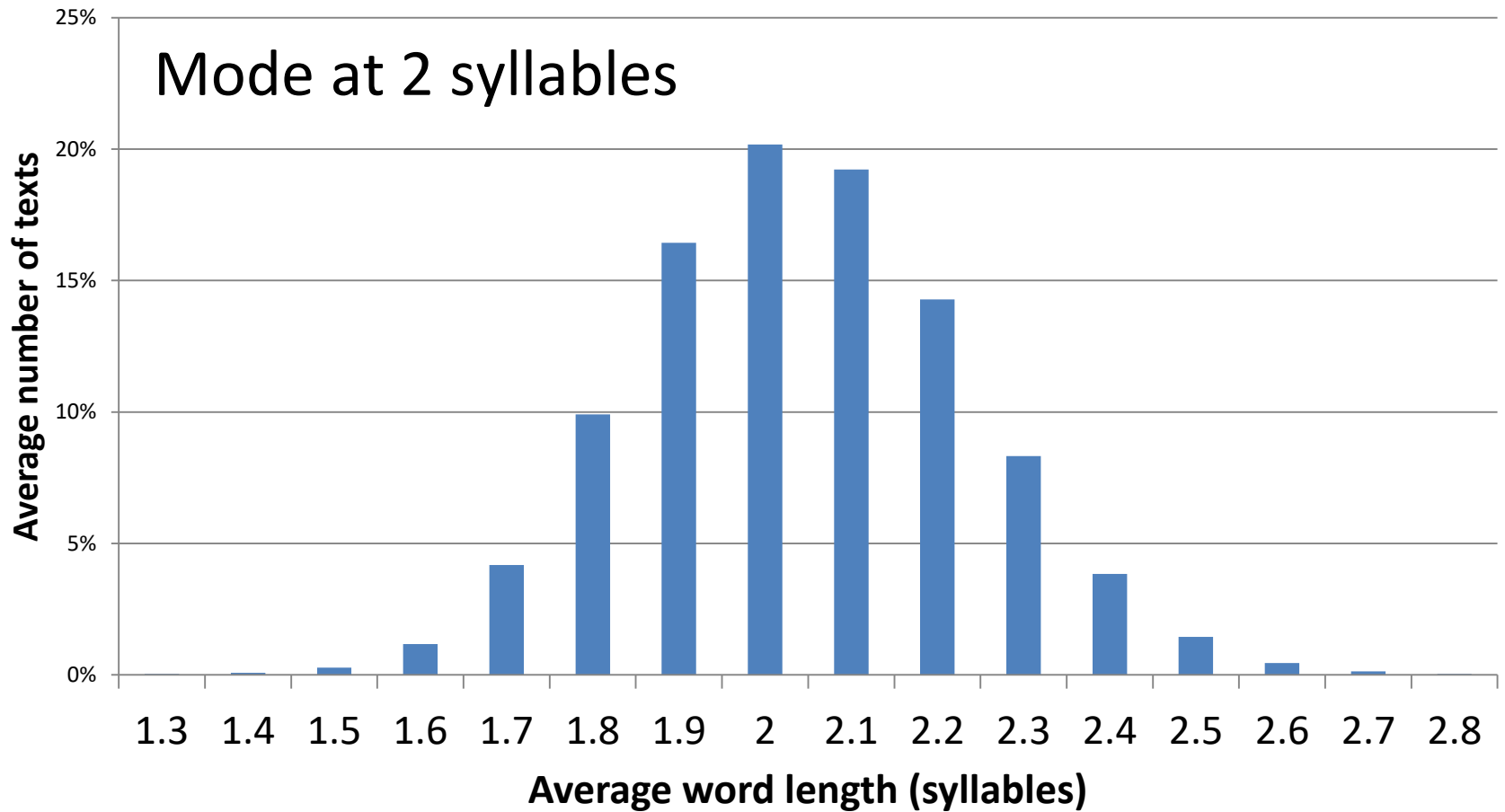
Word Length (in Phonemes)



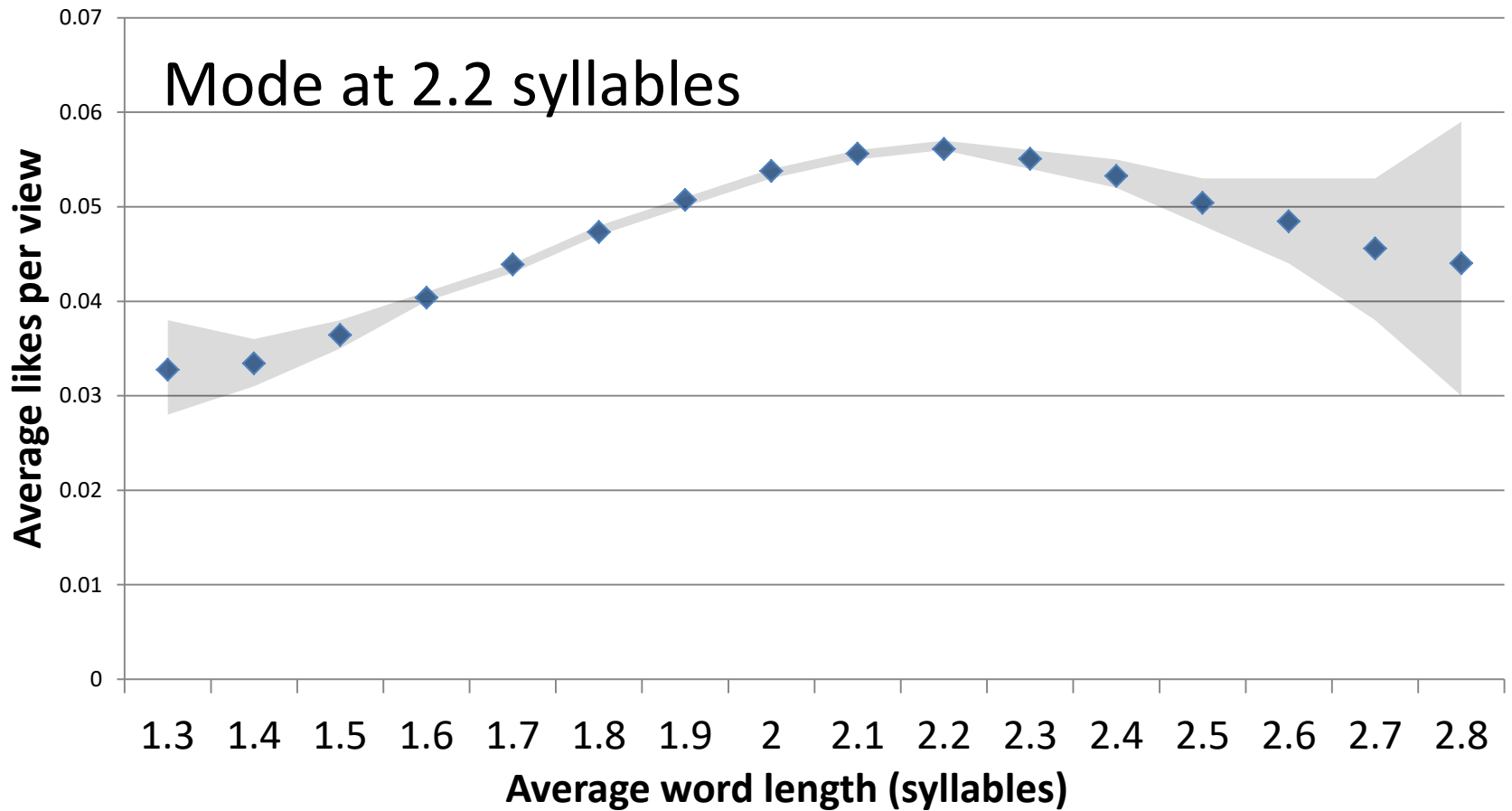
Word Length (in Phonemes)



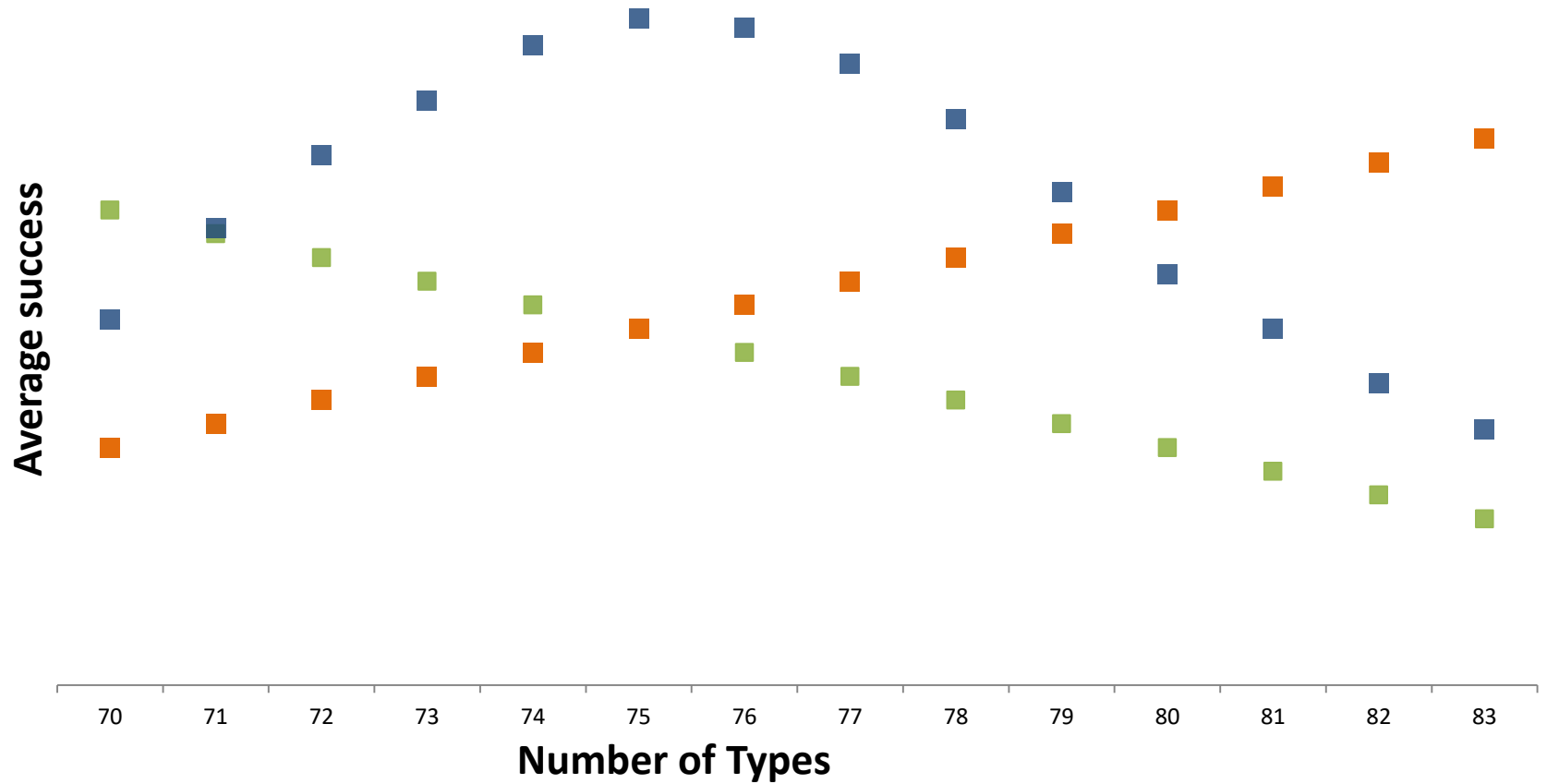
Word Length (in Syllables)



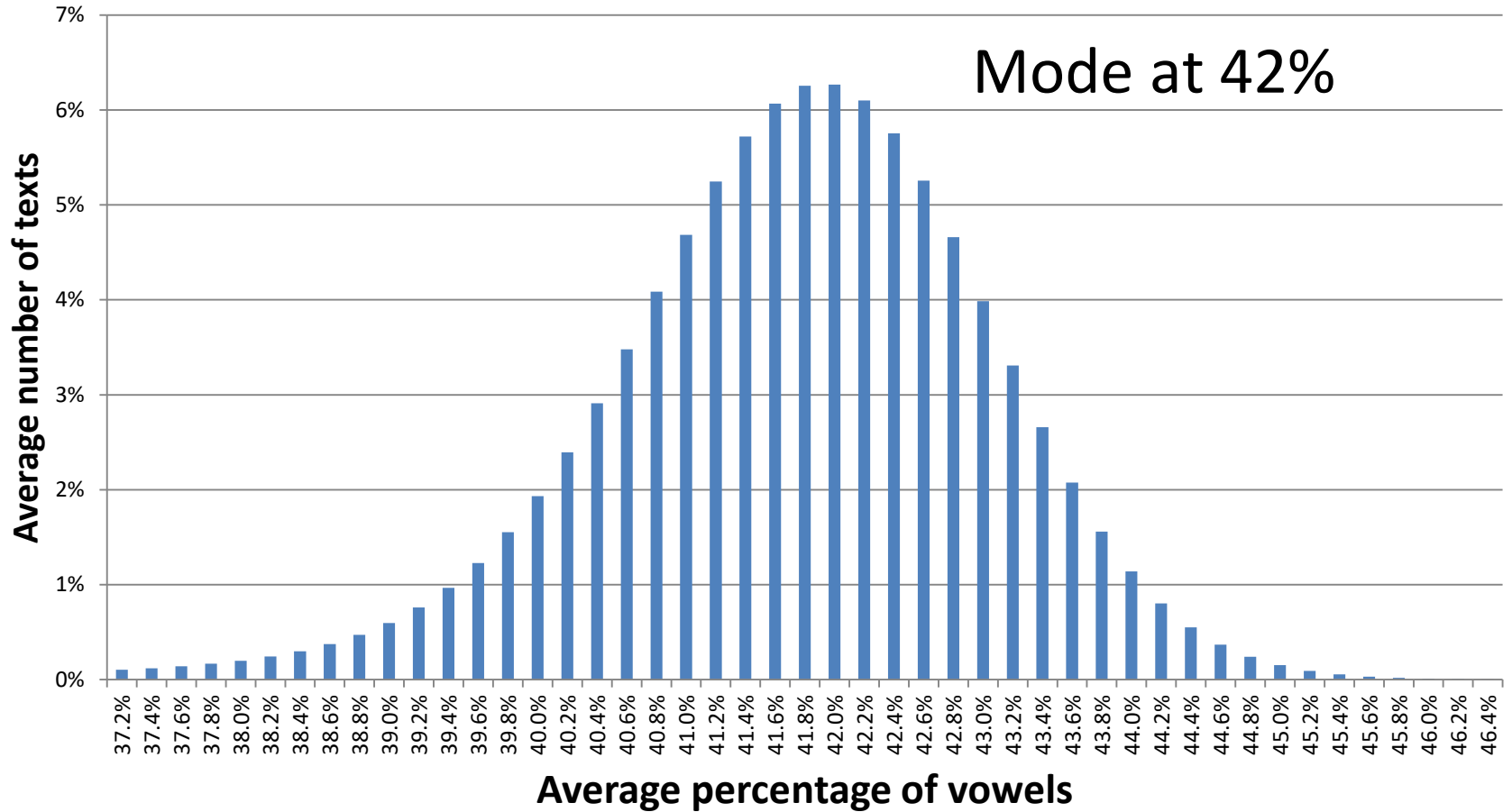
Word Length (in Syllables)



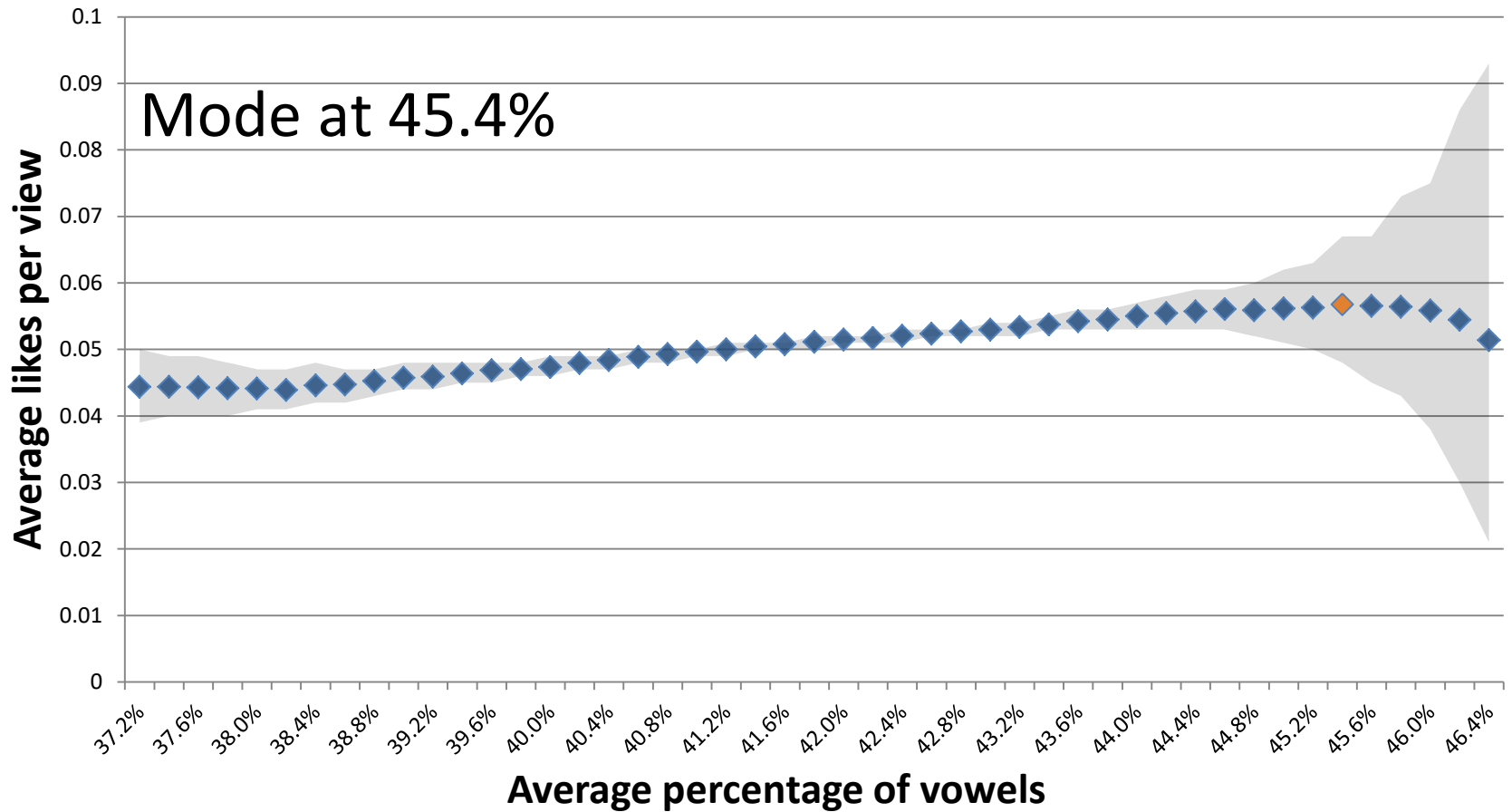
Hypotheses revisited



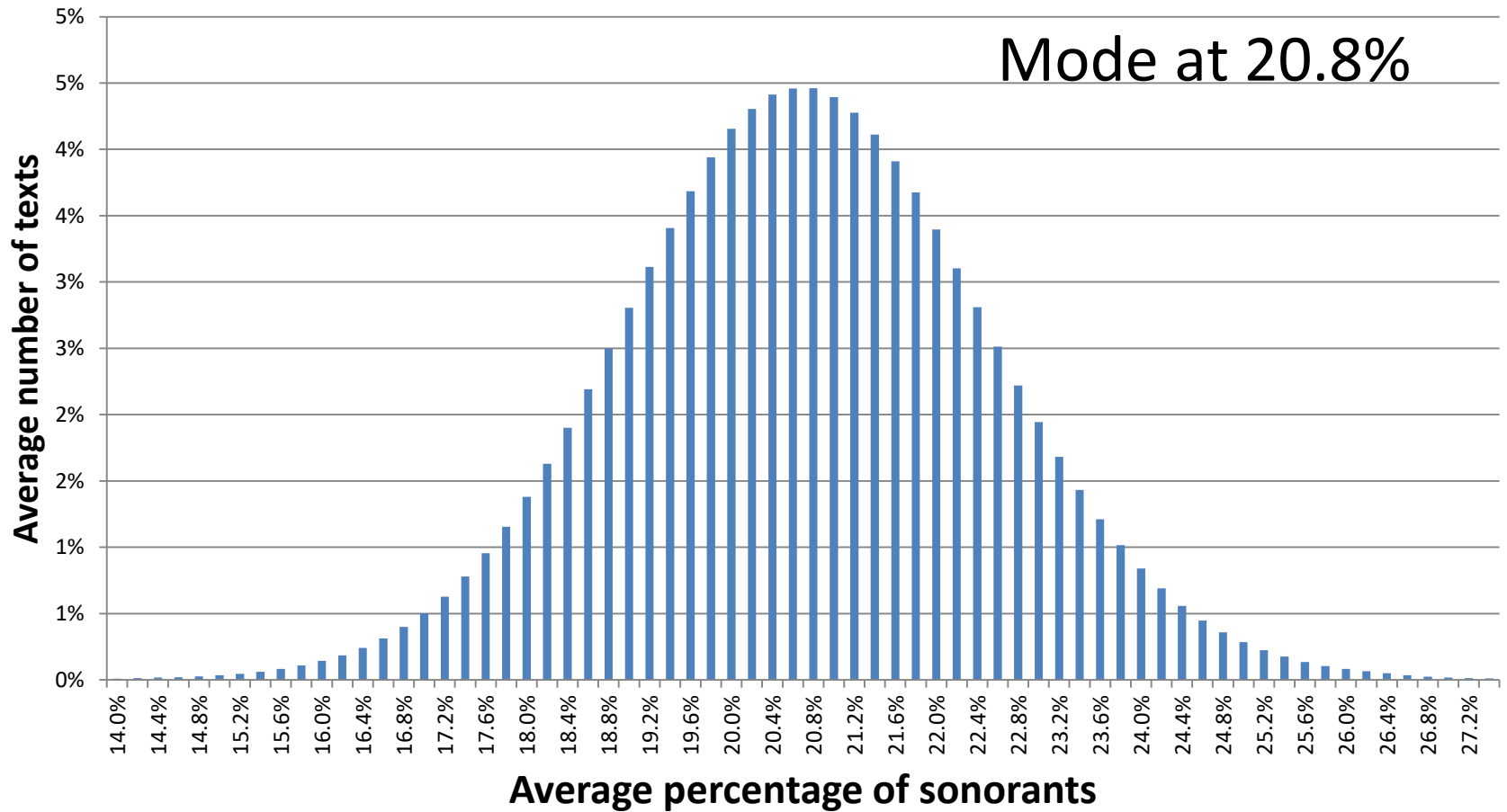
Proportion of Vowels



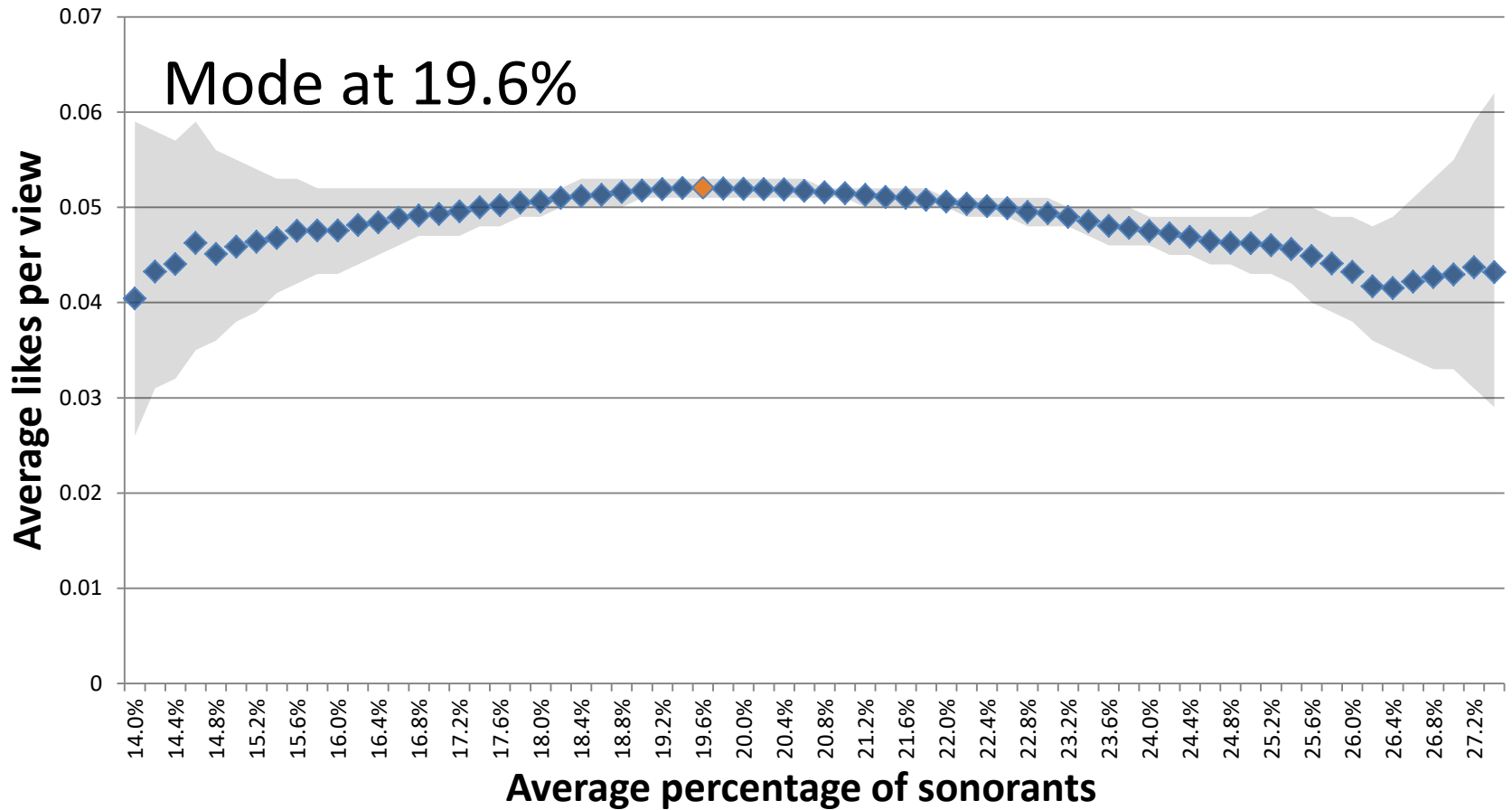
Proportion of Vowels



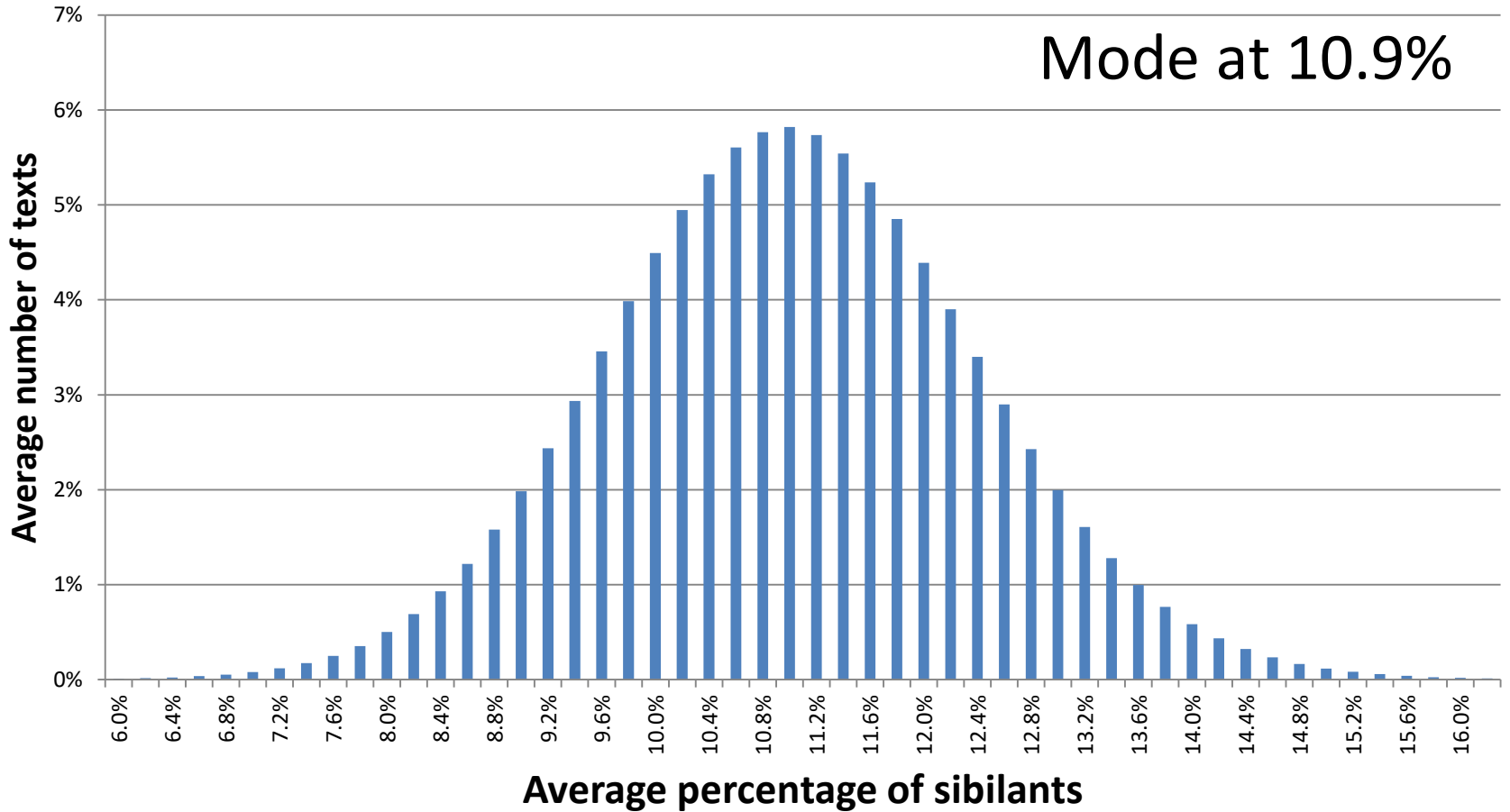
Proportion of Sonorants



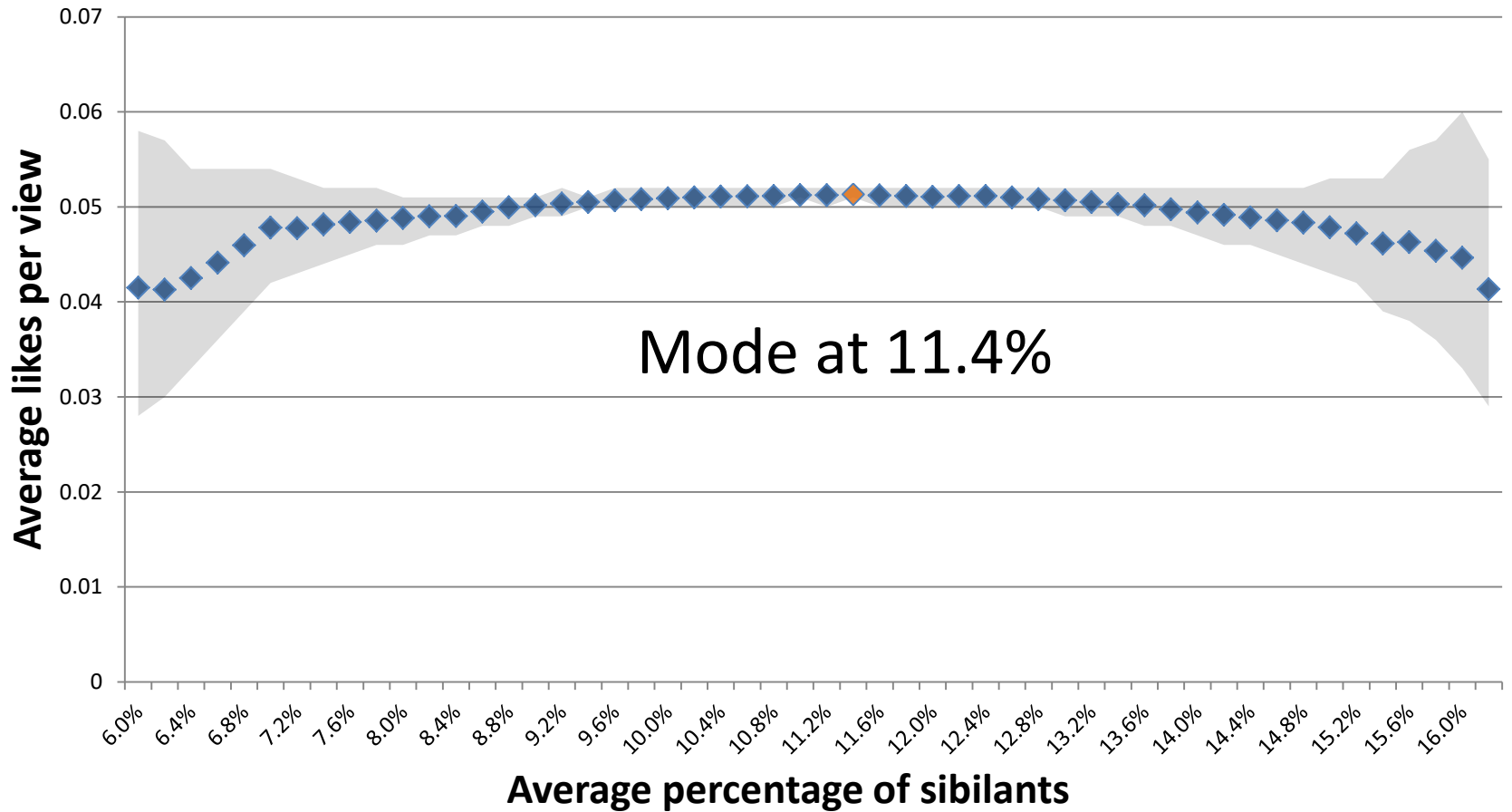
Proportion of Sonorants



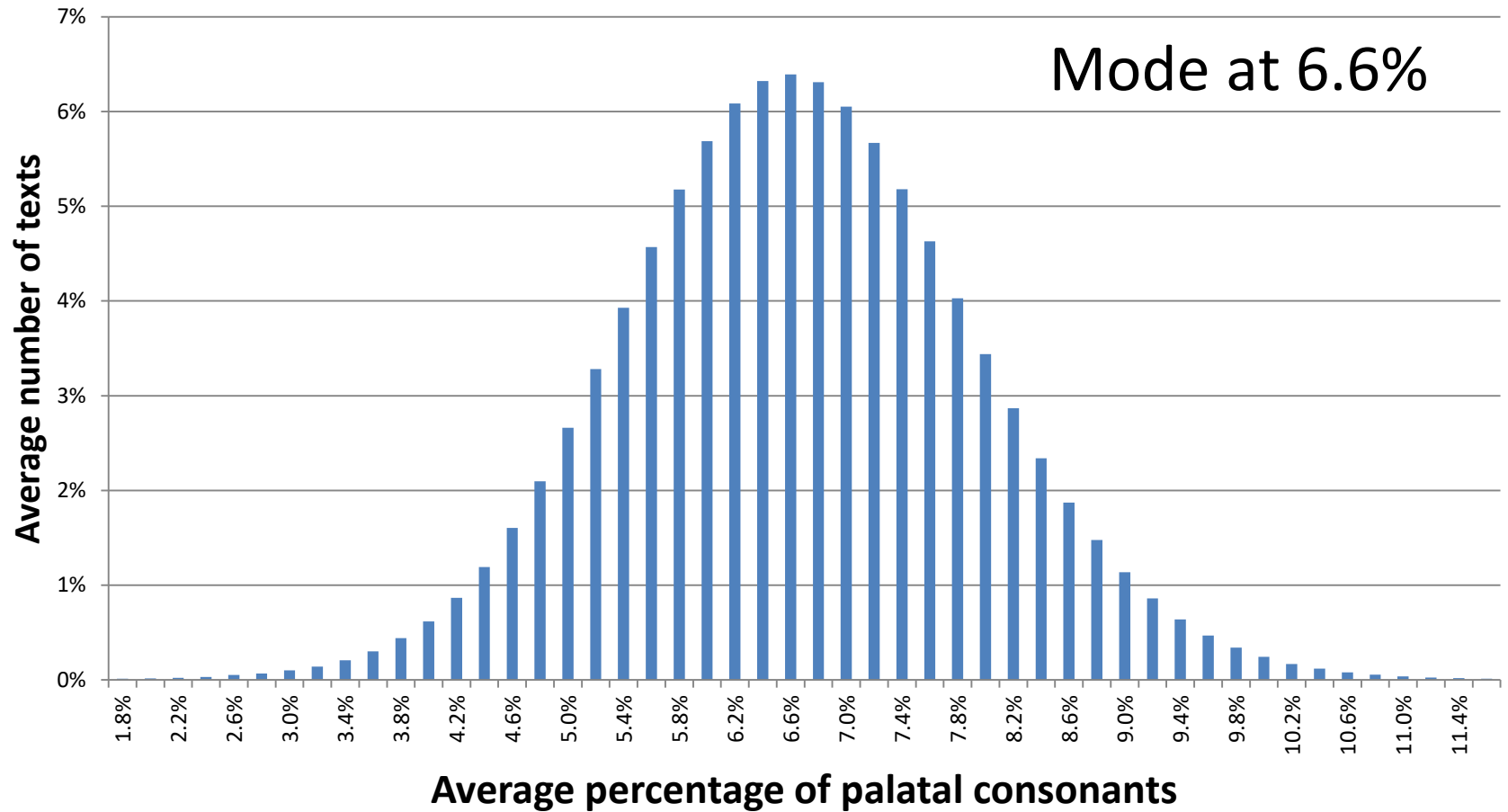
Proportion of Sibilants



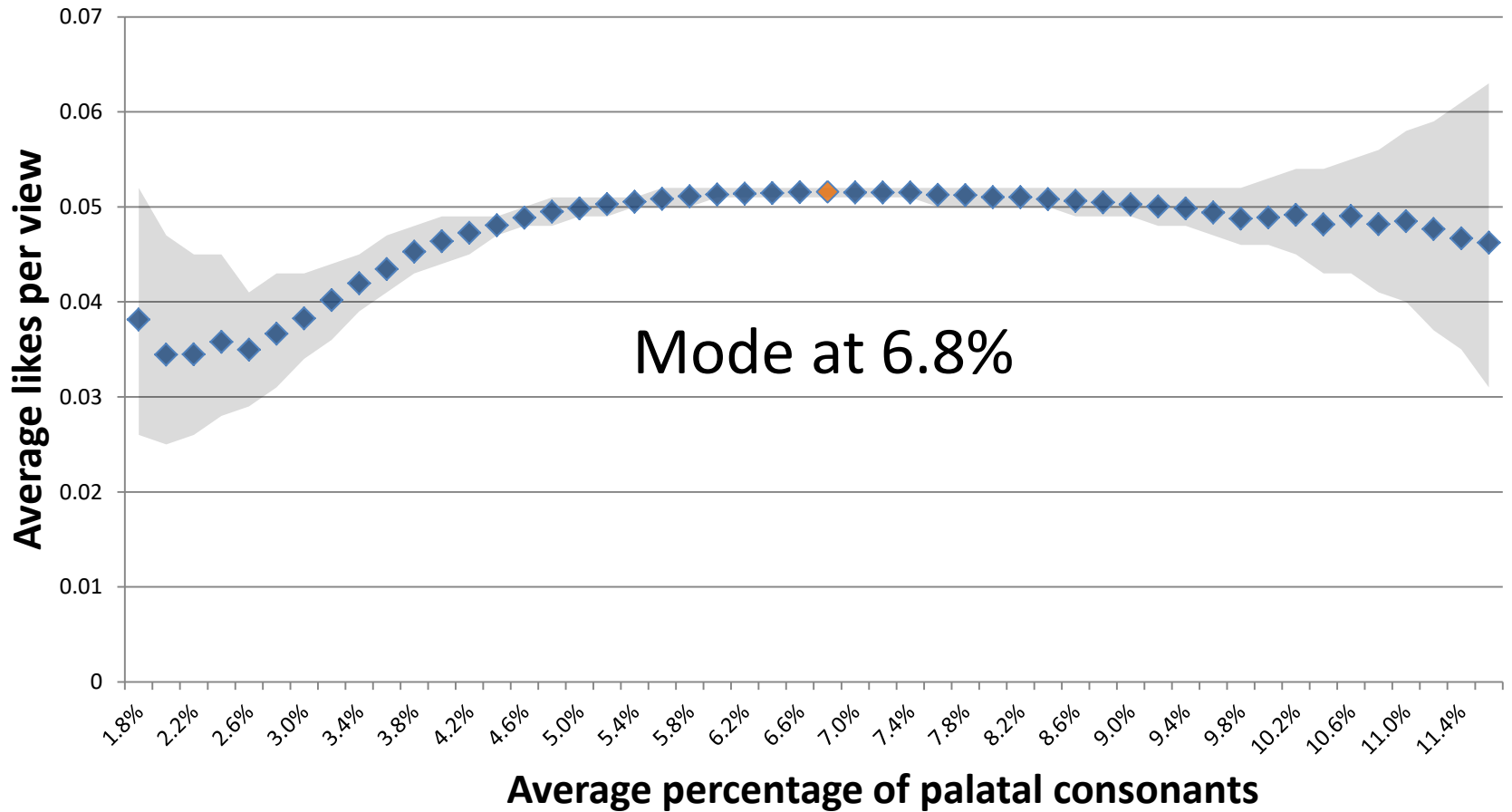
Proportion of Sibilants



Proportion of Palatal Consonants

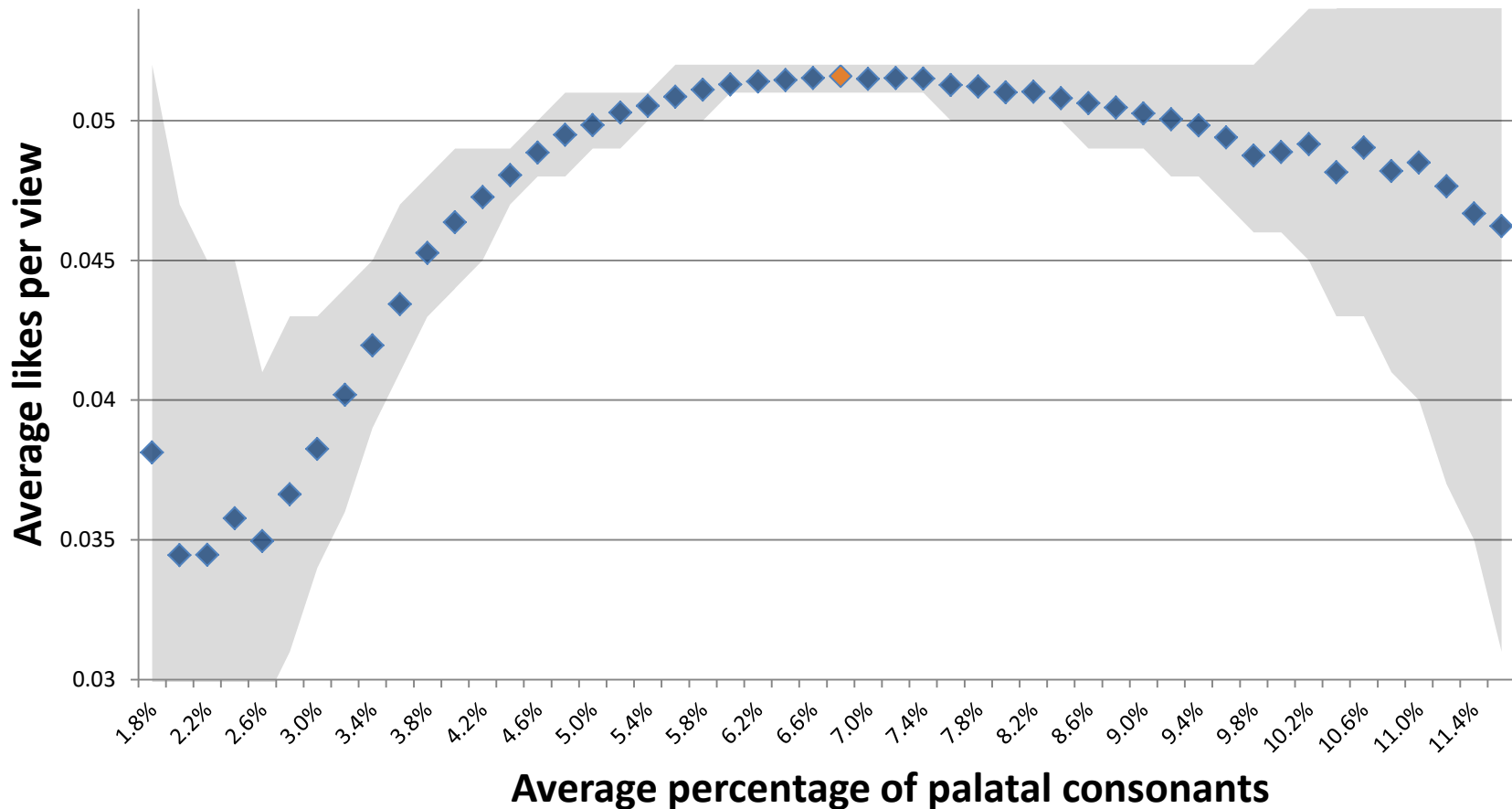


Proportion of Palatal Consonants

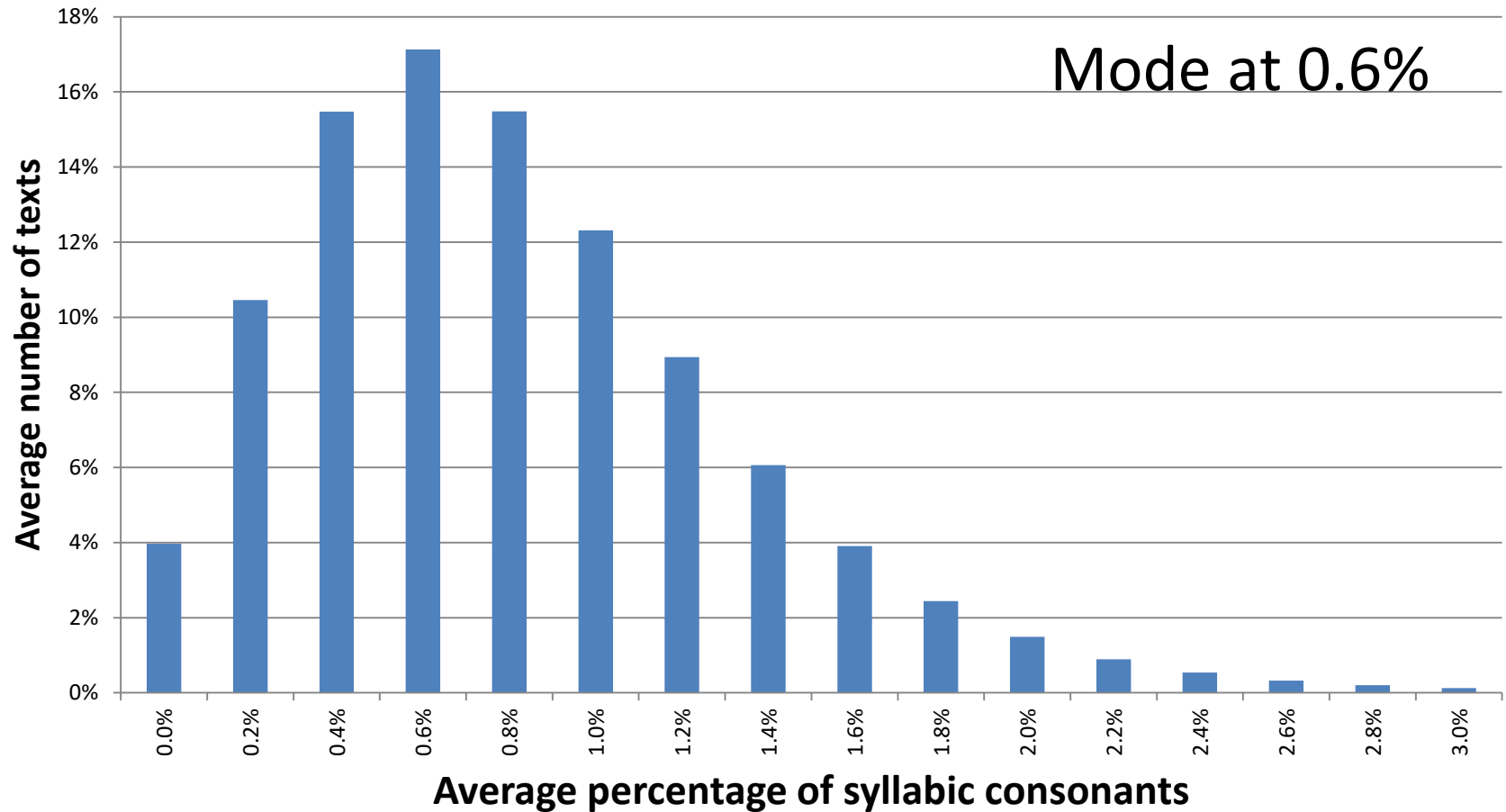


Proportion of Palatal Consonants

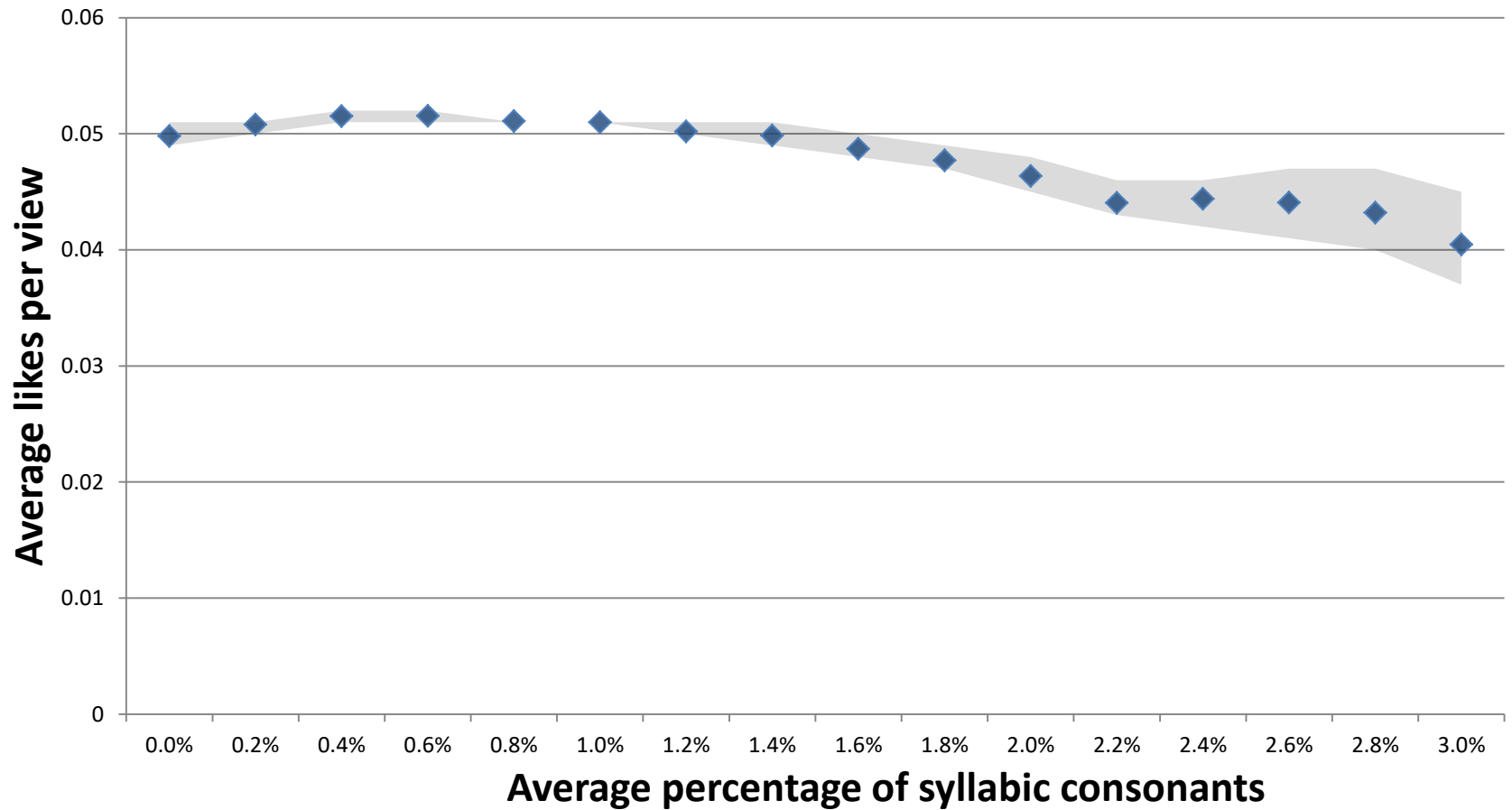
Zoom in



Proportion of Syllabic Consonants

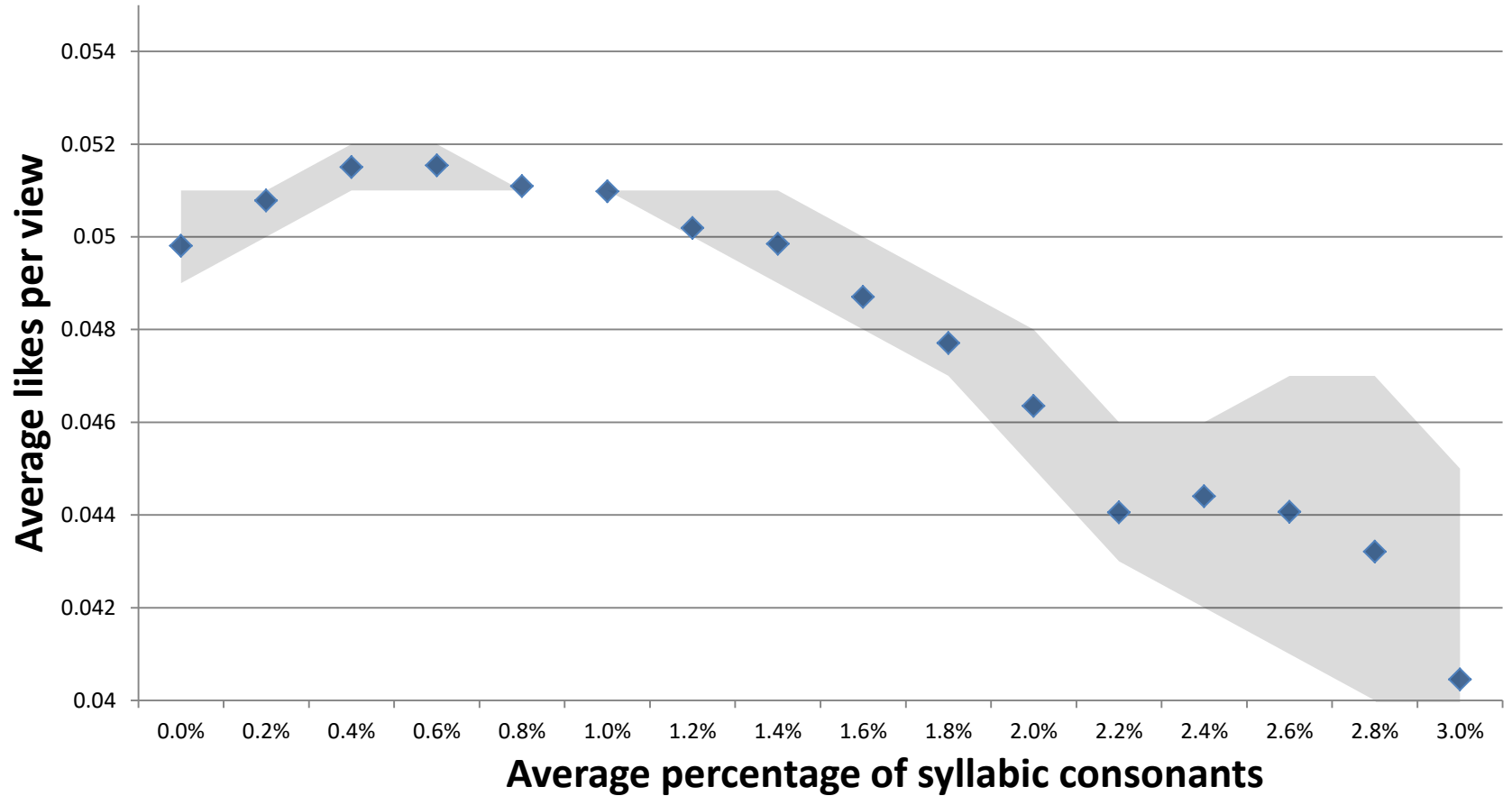


Proportion of Syllabic Consonants



Proportion of Syllabic Consonants

Zoom in



Discussion

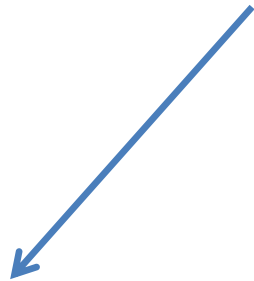
Average value of the given variable



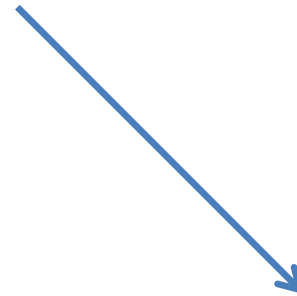
Text success

Discussion

Well written text



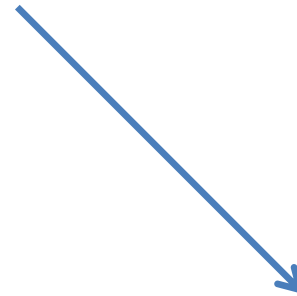
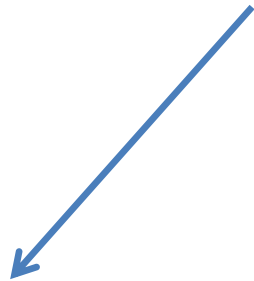
Average value
of the given
variable



Text success

Discussion

Well written text



Average value
of the given
variable



Text success

Discussion

Average value
of the given
variable



Text success

Conclusions (Theoretical)

- The readers' demand mostly meets the writers' supply.
- For some variables, only few texts achieved the values that the readers want.
- For some variables the most of the texts have almost the “ideal” values.

Conclusions (Practical)

- L1/L2 acquisition
- Write just... normally :-)

Thank you!

jiri@milicka.cz

ruzickovaalzbeta@seznam.cz