

Czech translations of the Gospel of Matthew from the diachronic point of view – plus ça change...



Radek Čech
(Ostravská univerzita)

Ján Mačutek

(Matematický ústav, Slovenská akadémia vied; Univerzita Konštantína Filozofa v Nitre)

Pavel Kosek
(Masarykova univerzita)

MUNI
ARTS



SLOVENSKÁ
AKADÉMIA
VIED



Acknowledgements

- Ján Mačutek: VEGA 2/0096/21
- Pavel Kosek: MUNI/FF-DEAN/1556/2019 The Development of the Czech Pronominal Enclitics "mi", "ho", "mu" in Older Czech Bibles"

Diachronic (non)-stability

- an impact of language changes on text characteristics
- differences in the dynamics of changes
 - lexical plane
 - more or less a continual process
 - if more dramatic changes appear, they are usually caused by extra-linguistic factor
 - phonological and morphological planes
 - very intense changes followed by periods in which these properties of language remain almost constant

Diachronic (non)-stability

¹ Zjěvenie Jezukristovo, ješto jest dal i zjěvil bóh otec, aby je prohlásil sluhám svým o těch věcech, ješto sě mají státi skoro, vzkázav skrzě svého anjela sluzě svému Janovi,

(Bible of Olomouc, 1417)

¹ Zjevení Ježíše Krista, které mu dal Bůh, aby svým otrokům ukázal, co se má brzy stát. On to prostřednictvím svého anděla naznačil svému otroku Janovi.

(Czech Study Translation, 2009)

Diachronic (non)-stability

¹ Zjěvenie Jezukristovo, ješto jest dal i zjěvil bóh otec, aby je prohlásil sluhám svým o těch věcech, ješto se mají státi skoro, vzkázav skrzě svého anjela sluzě svému Janovi,

(Bible of Olomouc, 1417)

¹ Zjevení Ježíše Krista, které mu dal Bůh, aby svým otrokům ukázal, co se má brzy stát. On to prostřednictvím svého anděla naznačil svému otroku Janovi.

(Czech Study Translation, 2009)

Diachronic (non)-stability

¹ Zjěvenie Jezukristovo, ješto jest dal i zjěvil bóh otec, aby je prohlásil sluhám svým o těch věcech, ješto se mají státi skoro, vzkázav skrzě svého anjela sluzě svému Janovi,

29 tokens

(Bible of Olomouc, 1417)

¹ Zjevení Ježíše Krista, které mu dal Bůh, aby svým otrokům ukázal, co se má brzy stát. On to prostřednictvím svého anděla naznačil svému otroku Janovi.

25 tokens

(Czech Study Translation, 2009)

The goal of the study

- modeling of the dynamics of changes
- lexical properties

The goal of the study

- modeling of the dynamics of changes
- lexical properties

- the proportion of identical word-forms (types)
- the average type length (AVL)
- text length (N)
- the proportion of hapax legomena (PHL)
- the moving average type-token ratio (MATTR)
- entropy (H)

Language material

- Gospel of Mathew
- Book of Revelation

Language material

- Bible drážďanská (BiblDrážď), Bible of Dresden (the 70s of the 14th century),
- Bible olomoucká (BiblOl), Bible of Olomouc (1417),
- Bible Melantrichova (BilMel), Melantrich's Bible (1570),
- Bible kralická (BiblKral), Bible of Kralice (1596),
- Bible svatováclavská (BiblSvat), Bible of St. Wenceslas (1677),
- Nový zákon (BiblSýk), New Testament (1909),
- Český studijní překlad Bible (ČSP – Bible), Czech Study Translation (2009).

Language material

- Bible drážďanská (BiblDrážď), Bible of Dresden (the 70s of the 14th century),
- Bible olomoucká (BiblOl), Bible of Olomouc (1417),
- Bible Melantrichova (BibMel), Melantrich's Bible (1570),
- Bible kralická (BiblKral), Bible of Kralice (1596),
- Bible svatováclavská (BibSvat), Bible of St. Wenceslas (1677),
- Nový zákon (BibSýk), New Testament (1909),
- Český ekumenický překlad (BibEkum), Czech Ecumenic Translation (1995)
- Český studijní překlad Bible (ČSP – Bible), Czech Study Translation (2009)
- Bible 21 (Bibl21), (2014).

The proportion of identical word-forms (types) in pairs of texts (PIV)

- the number of identical word-forms which occur in two texts

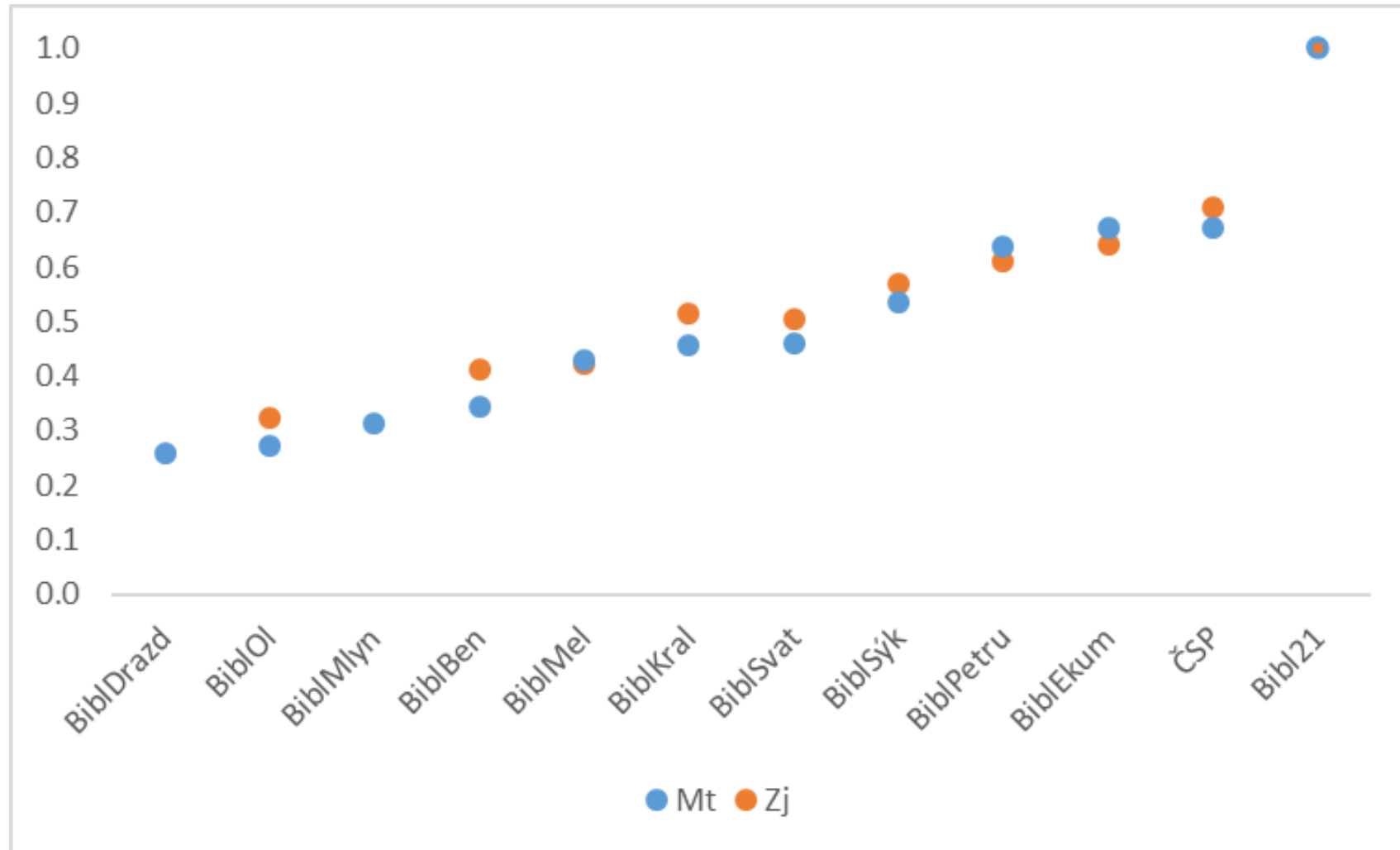
$$PIV = \frac{|V_i \cap V_j|}{|V_i|}$$

V_i ... the list of word-forms (types) of the text i

V_j ... the list of word-forms (types) of the text j

$$PIF = \frac{|V_{Bibl21} \cap V_{BiblDrážď}|}{|V_{Bibl21}|} = \frac{1075}{4150} = 0,259$$

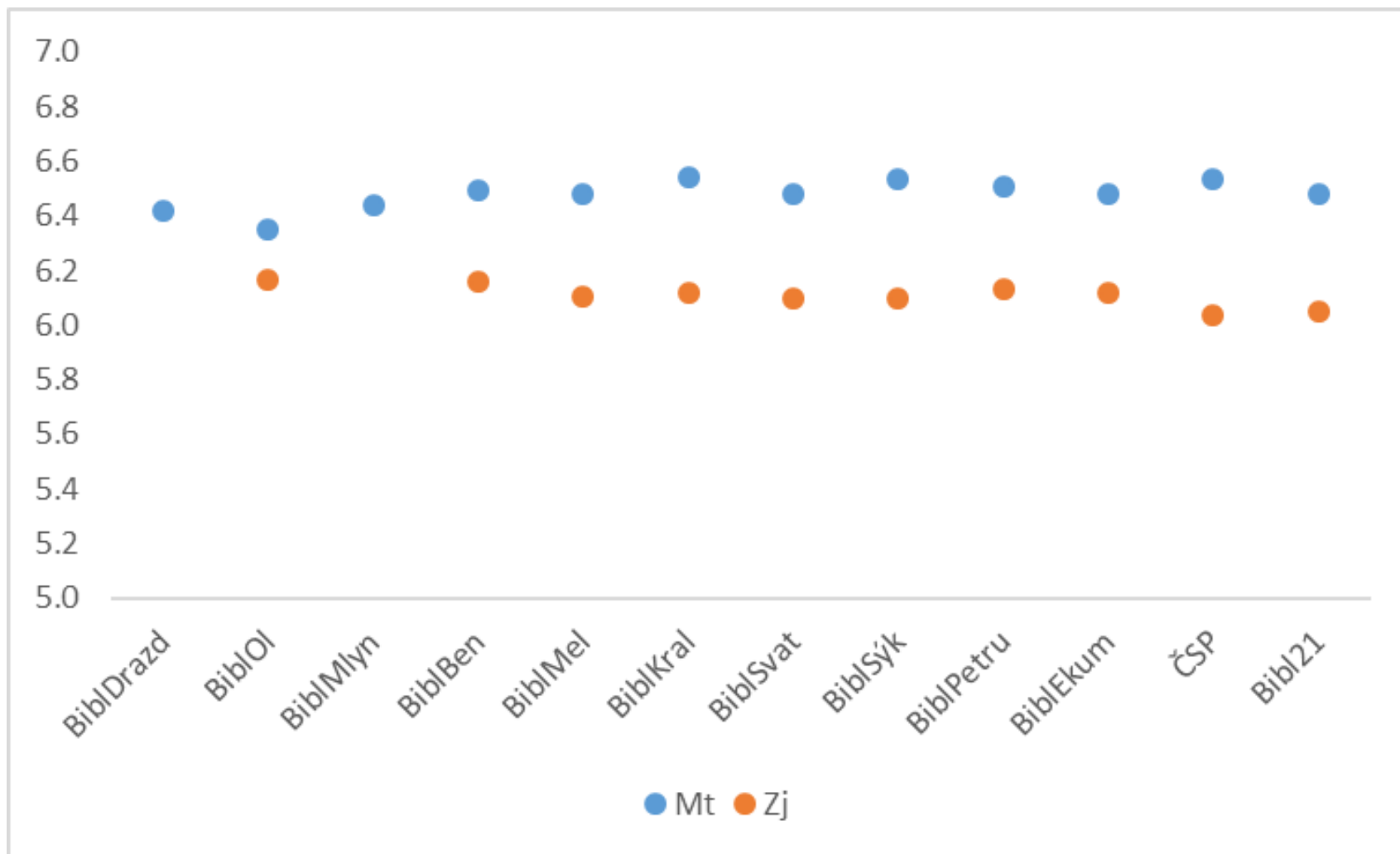
PIV (referential text = Bibl21)



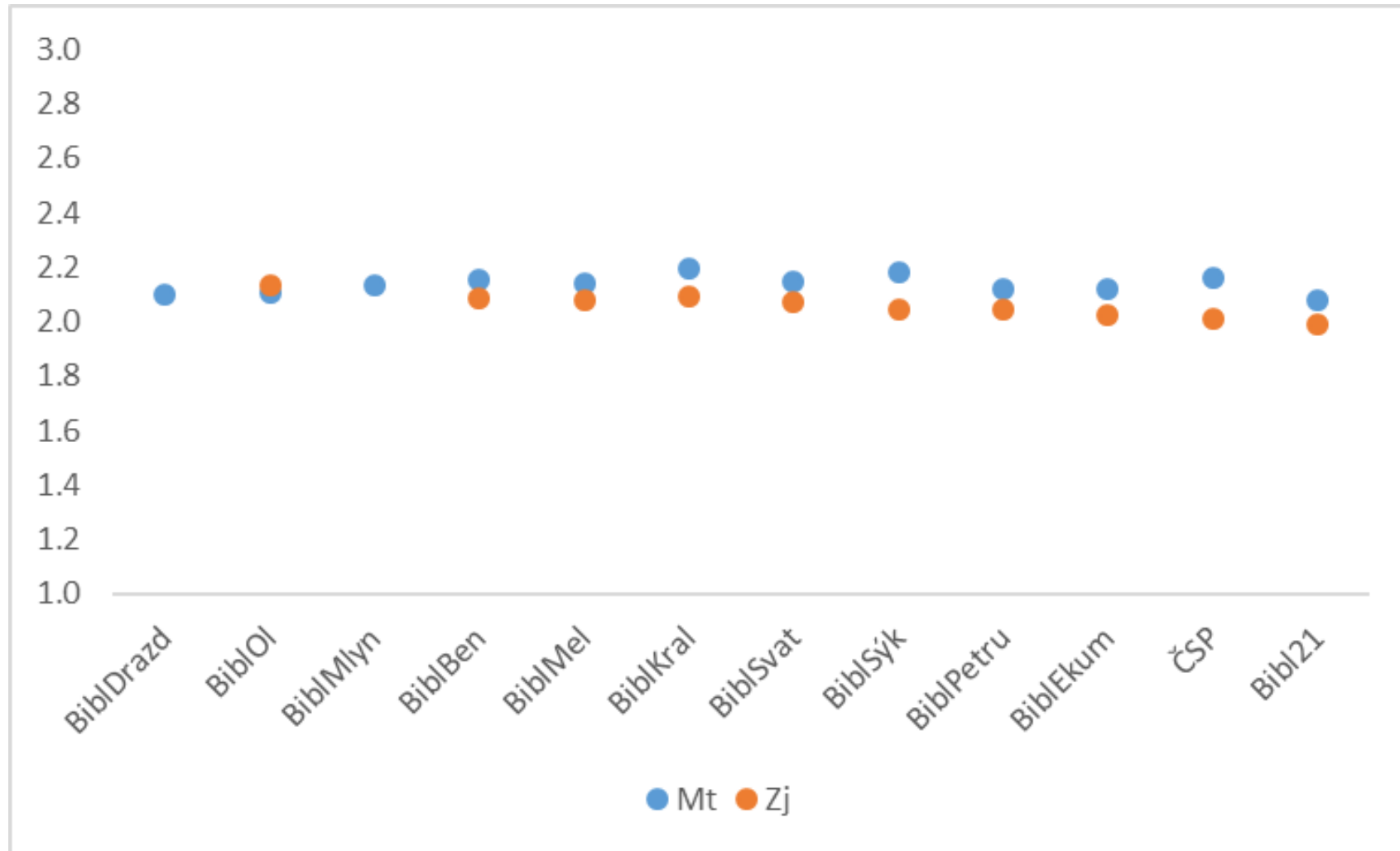
Average type length (AVL)

- the arithmetic mean of type lengths in a text
- the type length is calculated as the number of characters

Average type length (AVL)



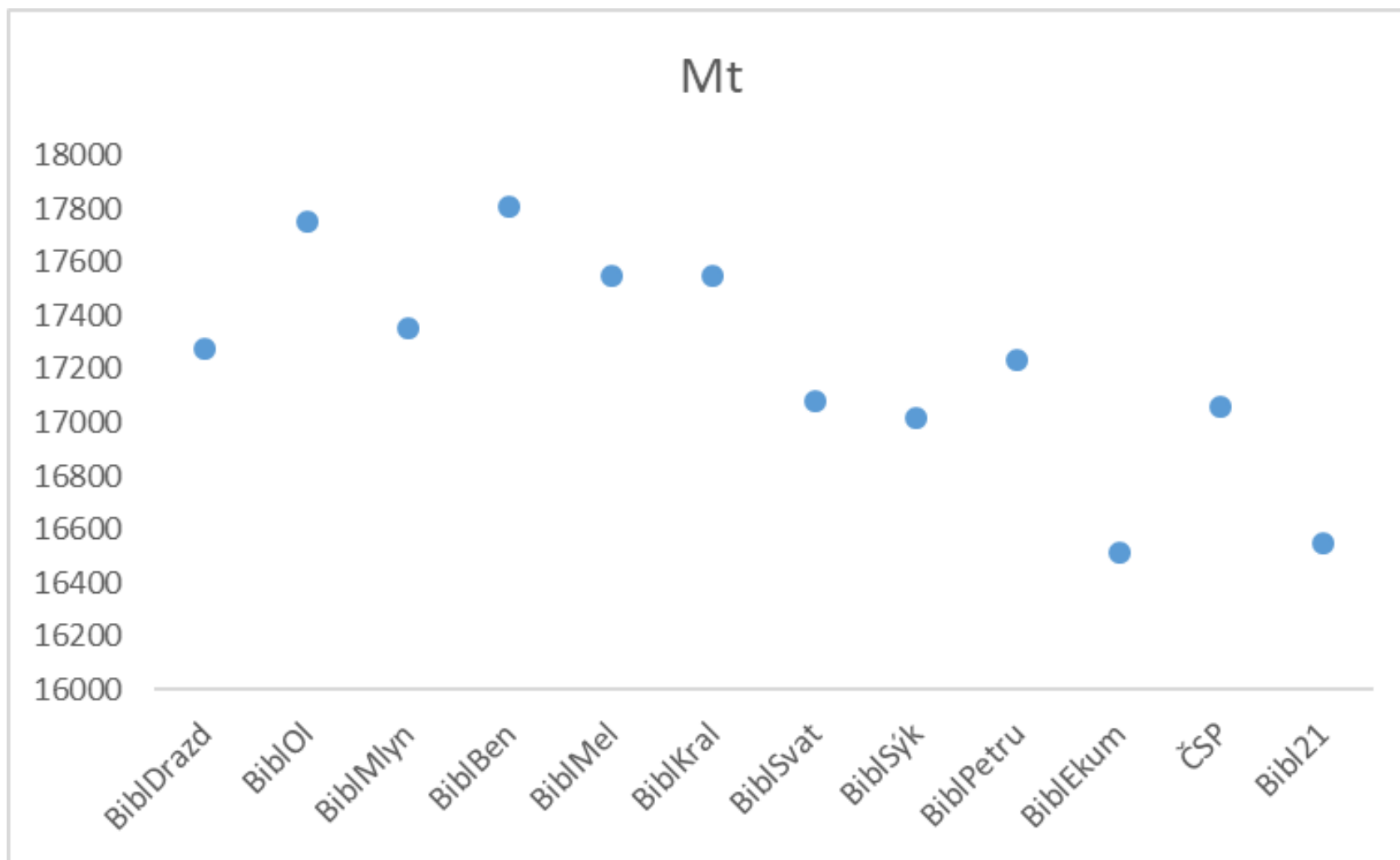
Average type length (AVL) – standard deviations



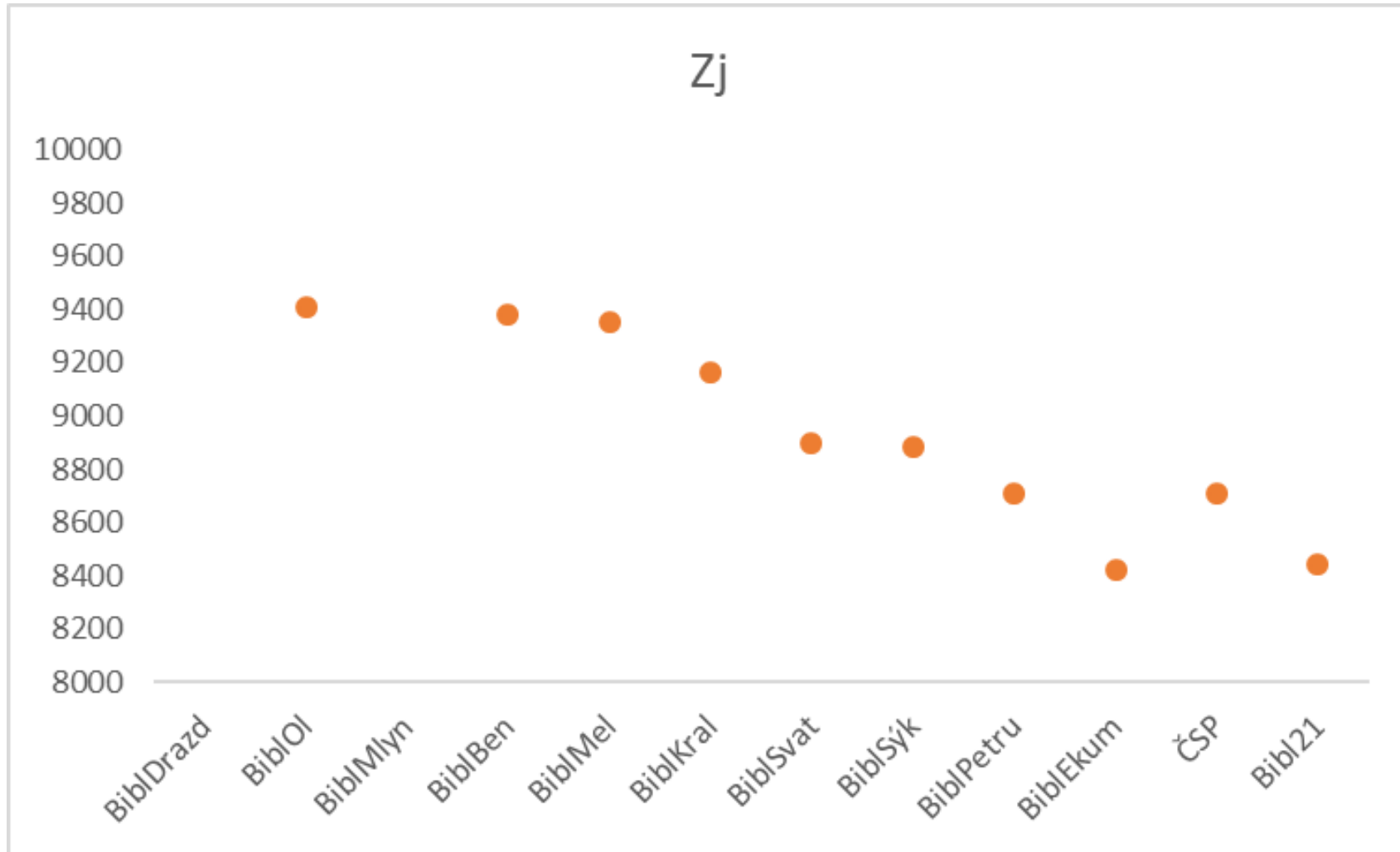
Text length (N)

- total count of all tokens in a text

Text length (N) – Gospel of Mathew



Text length (N) – Book of Revelation



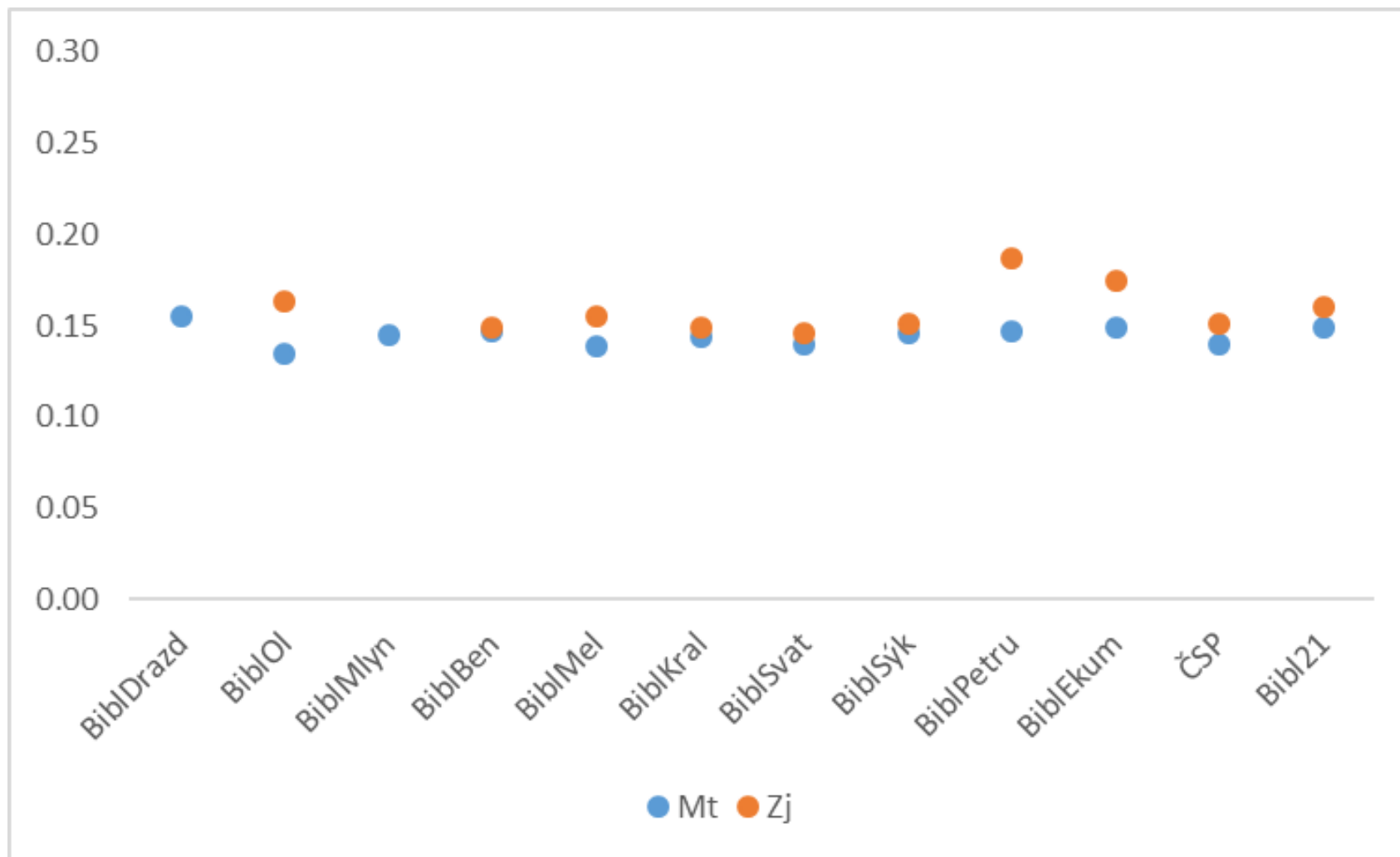
Proportion of hapax legomena (PHL)

$$PHL = \frac{N_{HL}}{N}$$

N_{HL} ... number of hapax legomena

N ... number of tokens

Proportion of hapax legomena (PHL)



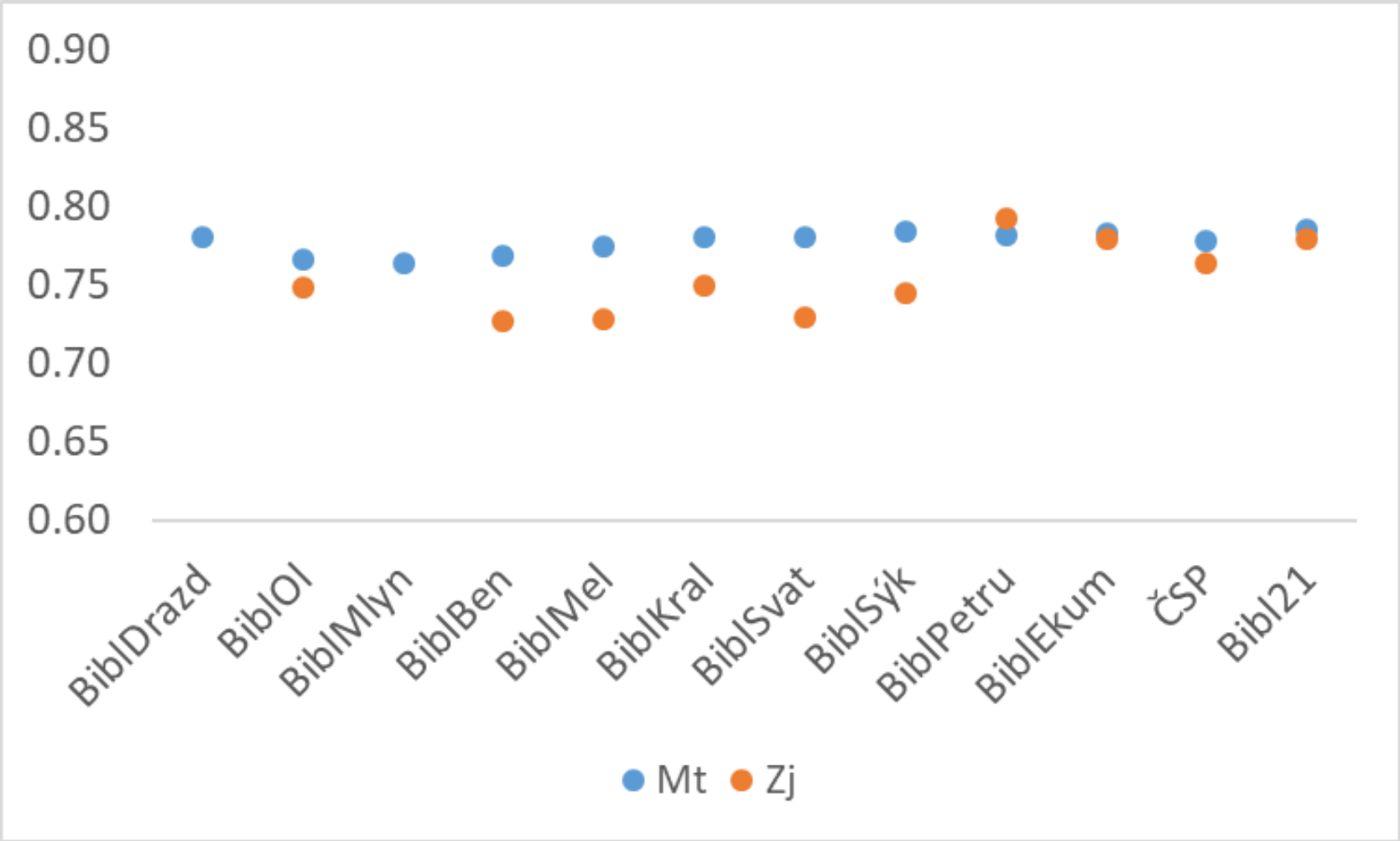
Moving average type-token ratio (MATTR)

- based on the segmentation of text into overlapping chunks (so-called windows), where the type-token ratio is calculated for each window
- the arithmetic mean of the type-token ratios in all windows

$$MATTR = \frac{\sum_{i=1}^{N-L} V_i}{L(N-L+1)}$$

- N ... the number of tokens in the text
- L ... the length of the window
- V_i ... the number of types in the window
- $L = 100$

Moving average type-token ratio (MATTR)



Entropy (H)

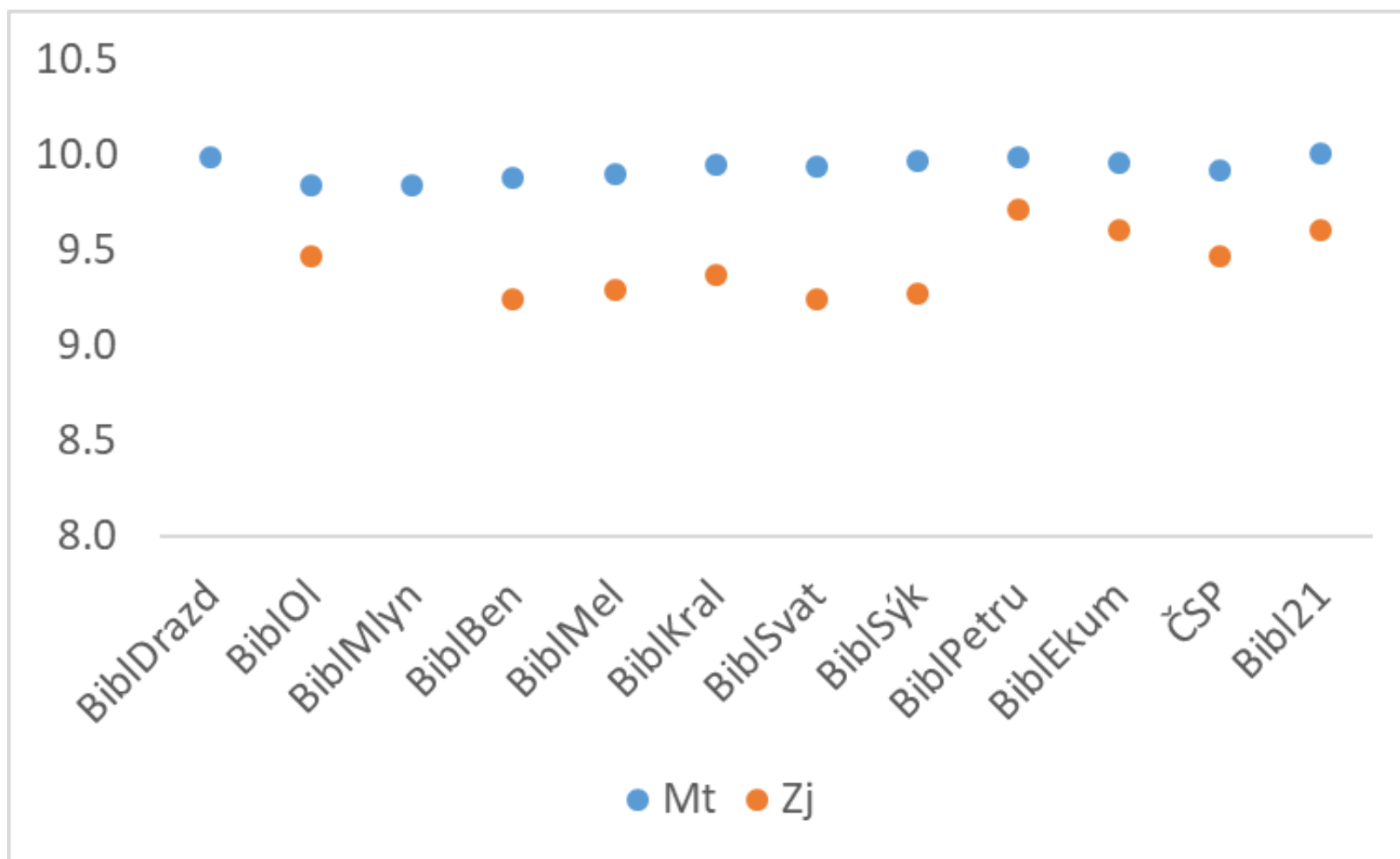
- measure of system uncertainty
- the degree of lexical and word form diversity
- the greater the value of entropy, the more diversified the vocabulary is

Entropy (H)

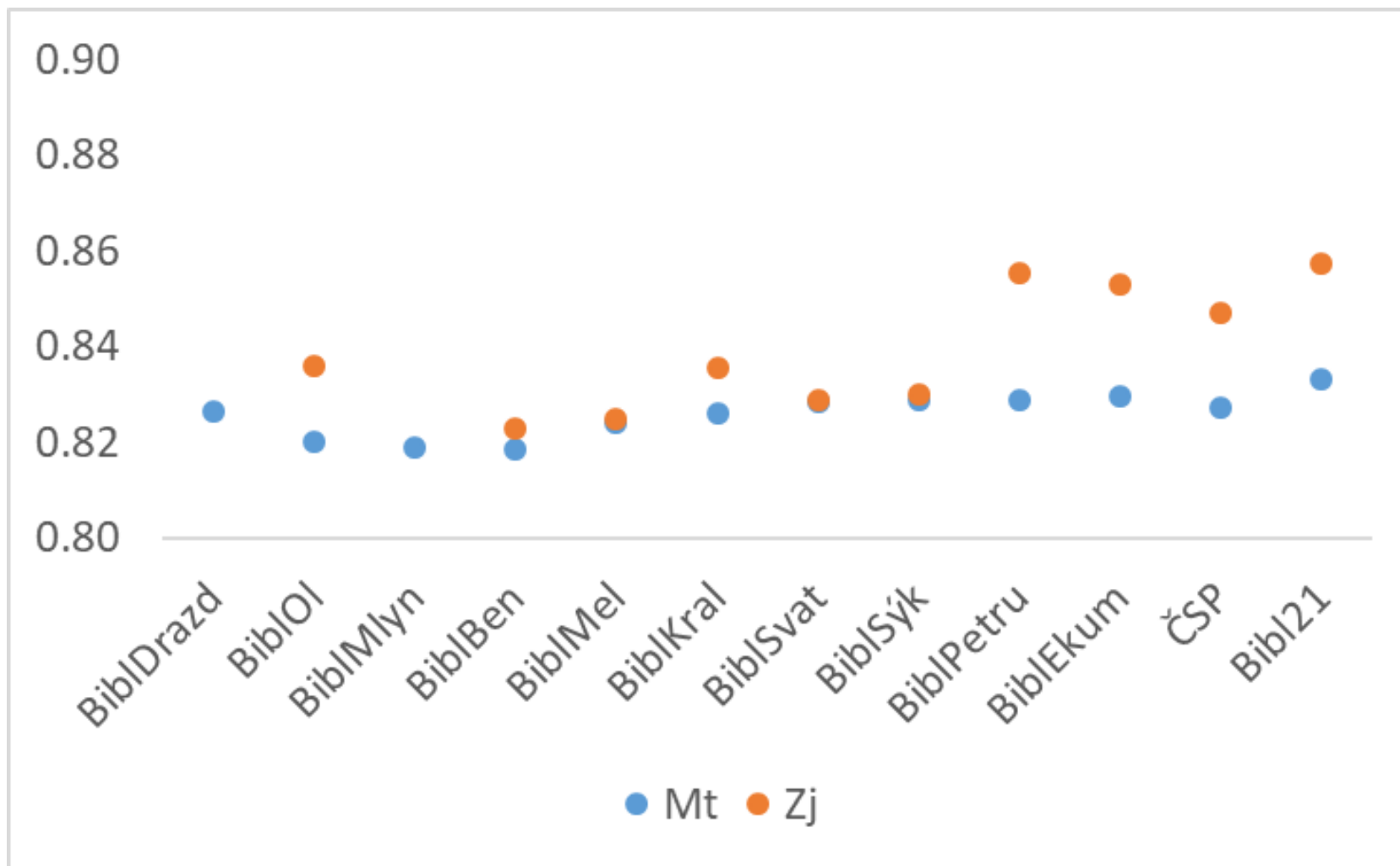
$$H = \log_2 N - \frac{1}{N} \sum_{r=1}^v f_r \log_2 f_r$$

- N ... the frequency of tokens in the text (text length)
- f_r ... the frequency of a given token

Entropy (H)



Relative entropy (H_r)



Conclusions

- non-stability (diachronic development)
 - proportion of different word forms (PIV)
 - text length (N)
- stability
 - word-form length
 - lexical diversity (Gospel of Mathew)
 - proportion of hapax legomena (PHL)
 - moving average type-token ratio (MATTR)
 - entropy (H)

Thank you.