This work is licenced under a Creative Commons Attribution 4.0 International License
Ovaj rad dostupan je za upotrebu pod licencom Creative Commons Imenovanje 4.0. međunarodna

IZVORNI ZNANSTVENI RAD
UDK: 81'25:81'23:004.8
DOI: https://doi.org/10.22210/strjez/55-1/2
Primljen: 12. 6. 2026.
Prihvaćen: 13. 10. 2026.

Psycholinguistic parameters as markers of (machine-) translationese

Maja Anđel
mandel@ffzg.unizg.hr
Faculty of Humanities and Social Sciences, University of Zagreb

The phenomenon of divergence of machine-translated and human-translated texts from their original is well-documented. Several assessment methods exist that consistently demonstrate this pattern, but they primarily focus on the superficial lexical level of the text. In this paper, we examine the potential of psycholinguistic lexical dimensions of word concreteness and age-of-acquisition to complement the existing metrics with new insights into deeper, semantically based features of translations. In order to do so, we analysed the original text of Kafka's novel "The Trial" and compared its human-made translations into Croatian (by Zlatko Crnković, 1979) and into English (by David Wyllie, 2005) with its machine-translated versions into the two languages. Our results show that psycholinguistic lexical parameters of concreteness and age-of-acquisition do indeed differ from the original text in a systematic way, both for human and machine translations, but in different ways. We therefore conclude that the chosen parameters do show some of the distinctive lexical features of texts, so further investigation of their scope might be of value.

Keywords*: concreteness, age-of-acquisition, translation, measure of textual lexical diversity – MTLD, Lexical frequency profile – LFP*

1 INTRODUCTION

Ever since Mona Baker (1993) first noticed the systematic simplification of human translations in comparison to original texts, extensive research on this phenomenon followed (Baroni, 2005; Volansky et al., 2015; Koppel

& Ordan, 2011; Juola, 2008). Eventually, even the label *translationese* was adopted (Gellerstam, 1986) to denote the inherent features and systematicity of this linguistic phenomenon. In addition to that, it is also known that, along with other imperfections, AI-generated texts are also less lexically diverse than human-produced ones, which is sometimes referred to as The Great Linguistic Flattening (Paschalidis, 2025). So, what happens if we combine the activity of translation with AI text generation, and what can we expect from AI-generated translations with respect to lexical diversity and variety? In this paper, we compared Kafka's novel "The Trial" (Kafka, 2022) with its machine translations into Croatian, an under-resourced language (Alves et al., 2020; Štefanec et al., 2024) and English, and their human-translated versions as controls for machine translation, by Zlatko Crnković (Kafka, 1979) and David Wyllie (Kafka, 2005), respectively.

There have been many attempts and many ways to quantify the simplification of translation as a process-inherent phenomenon (Baker, 1996; Baker et al., 1993; Niu and Jiang, 2024). Some of them look at the lexical richness and diversity (Laufer, 1994; Laufer & Nation, 1995; Brglez & Vintar, 2022), some again look for the syntactic "shining-through" (Teich, 2012). However, these traditional corpus-based approaches suffer from well-documented limitations: lexical diversity and frequency metrics remain at the level of surface word-form distribution and are unable to capture semantic or psycholinguistic properties of the words they count (Brglez & Vintar, 2022, p. 1). In addition to that, as a measure of surface variation, the Type-Token Ratio remains highly sensitive to text length (Niu & Jiang, 2024, p. 5). In this paper, we first examine the lexical diversity of an original text and its translations using the usual measure of textual lexical diversity (MTLD – McCarthy and Jarvis, 2010) and the Lexical Frequency Profile (LFP – Laufer, 1994) methods. After that, we analyse the texts within the framework of psycholinguistic parameters of concreteness and age of acquisition, which show potential to reveal additional semantic depth in the texts. By extending the analytical focus from surface forms to these deep semantic and developmental dimensions, this study represents the first attempt to evaluate literary translations using psycholinguistic databases of human ratings. This approach allows us to map the sensory and developmental expressions of translationese, and to expose some quantitative differences in cognitive dimensions of text production between human authors, professional translators, and NMT systems.

Concreteness is defined as the degree to which a word refers to an entity that can be experienced by the senses (Paivio et al., 1968; Peti-Stantić et

al., 2021). In simplest terms, one can argue that we generally experience material objects, such as a table, as more concrete than abstract concepts, such as loyalty, and that the words we use to denote them inherit their original degree of concreteness. So the original sensory distinction, translated into a linguistic distinction between concreteness and abstractness of meaning, is formalised in Dual Coding Theory, which suggests that concrete words are represented both perceptually and verbally in the mental lexicon, while abstract concepts rely almost entirely on verbal associations (Paivio, 2010; Peti-Stantić et al., 2021). Nevertheless, when writing any text, we have a choice of synonyms which can differ in concreteness value, such as sight (3.21) and vision (3.39, scale 1-5, Brysbaert et al., 2014). If a choice of terms in a translation is systematically more or less concrete than the original, it will reflect on the reading experience. Concrete translation equivalents share a large number of semantic features across languages, whereas abstract words depend more heavily on language-specific contexts and show much lower cross-linguistic semantic overlap (Chaouch-Orozco et al., 2024). By extension, one can expect that the choice of translation equivalents for more concrete words is more or less straightforward, but it leaves the translators, both humans and machines, with a wide spectrum to choose from when it comes to abstract words, and opens the possibility of changing the overall concreteness flavour of the text. (The investigation of exact mechanisms that lead the translators' choices is outside the scope of this paper but is nevertheless a very intriguing question.) Using psycholinguistic databases of concreteness ratings, we measured the overall concreteness scores for lexically full words from the original texts and compared them to those of respective human and machine translations.

The same is true for the parameter of age of acquisition (Kuperman et al., 2012). Age of acquisition (AoA) correlates with concreteness, with early-acquired words tending to be more concrete and later-acquired words more abstract (Peti-Stantić et al., 2021). Therefore, in analogy with the established relationship between concreteness and sensory grounding (Paivio et al., 1968, as cited in Peti-Stantić et al., 2021), one might expect early-acquired vocabulary to reflect this sensory grounding, with later-acquired words carrying more formal or abstract flavour. Again, the choice of words used by the translator can shift the cognitive perspective of the reader – if the chosen translation equivalents are systematically lower on the AoA parameter and abound in words acquired earlier in life, the translation may seem oversimplified; otherwise, the translation can be too complex for the intended audience, irrespective of the translators' awareness when choosing the transla-

tion equivalents (Oshika et al., 2024). When constructing a coherent mental situation model of a real-world situation or text, readers integrate multiple sources of information during comprehension (Traxler, 2011, p. 217). Lexical properties such as frequency and concreteness may contribute to this process; it is therefore plausible that age-related lexical properties affect it as well. To our knowledge, this is the first study applying text-level AoA profiles to comparisons of original literary texts, human translations, and NMT translations. And while human translators navigate synonymy strategically, at least at levels of text genre, formality, target text distribution appropriateness, etc., modern neural machine translation (NMT) systems do rely on distributional frequency, but do not possess the ability to discern fine stylistic properties of words (Vanmassenhove et al., 2021).

2 AIM

The main research question of this paper is whether psycholinguistic parameters can indeed offer a deeper understanding of text structure, which can, subsequently, be applied as a measure of divergence of translated texts from their originals. The parameters in question are concreteness and age-of-acquisition.

3 METHODS

We downloaded the text of the novel "The Trial" by Franz Kafka from the Project Gutenberg page and translated it into Croatian and English, using the Phrase Next GenMT. In accordance with our intention to analyse the raw MT output, there were no interventions into the text in the form of post-editing. The machine-translated texts, as well as their human controls, were lemmatised using the Stanza toolkit for German and English (Qi et al., 2020) and CLASSLA for Croatian (Terčon et al., 2025). The present analysis is focused on lexical choices in human and machine-translated texts (HT and MT), as the psycholinguistic parameters we are examining are bound to lexemes. Even if they are inadequate or even erroneous, they are still lexical choices made by the translator. In the present paper, we will not assess or address the syntactic (in)adequacy of machine translations. The lexical content for the analysis beyond MTLD was restricted to nouns, verbs, adjectives, and adverbs, since they, as lexically full words, are more likely to have ratings for concreteness and the age of acquisition (AoA). Personal nouns were removed from the analysis.

That said, finding comparable data for concreteness and AoA estimates in all three languages proved extremely challenging, which will be discussed in more detail in the section on limitations. For German, the concreteness and AoA values were taken from the GPT estimates of concreteness, valence, arousal and AoA for 167K German words (Conde et al., 2026). For Croatian, all estimates stem from the Croatian Psycholinguistic Database (Peti-Stantić et al., 2021). English concreteness values were retrieved from the database of 40,000 English frequency and concreteness estimates (Brysbaert et al., 2014) and the AoA values from Kuperman et al. (2012). The ratings were mapped onto lists of lemmatised words extracted from each text, with concreteness and AoA weighted by the number of tokens. Where necessary, concreteness ratings were normalised on the scale 1–7, in order to match the German database. Due to the nature of the psycholinguistic databases, the methods of their construction, and their different sizes, it was not possible to assign values to all the words from the texts, so the analyses were made taking into consideration only the words which do have respective values. The analyses for concreteness and AoA were performed in JASP (2025) using ANOVA with Brown-Forsythe corrections and Games-Howell Post-hoc comparisons, due to significant Levene's tests.

4 RESULTS AND DISCUSSION

4.1 Measure of Textual Lexical Diversity (MTLD)

First, we looked at MTLD, as one of the classical metrics for assessing lexical diversity. It is calculated as the mean length of word strings that maintain a criterion level of lexical variation. We based our calculations on lemmas, due to the difference in morphological richness between the three languages, maintaining the suggested stabilisation point at 0.720. The analysis was performed on lemmatised text, but before removing parts of speech other than nouns, verbs, adjectives and adverbs. MTLD was calculated on the whole text. Our results reveal the expected lexical poverty of the machine translation, but also the extraordinary similarity of the MTLD for the English HT and the original and an even higher value for the Croatian HT (Table 1). So, while the human translators are able to maintain the lexical richness of the original or even build further upon it, the MT systems exhibit a clear algorithmic bias towards mainstream, frequent expressions, which will be especially visible in the next section on the LFP. At this point, we will not engage in the debate on whether enrichment in

HT should be preferred to maintaining lexical variety; let's just note that an enriched translation is by itself a divergence from the original.

Table 1
Values of MTLD (lexical diversity) for the original text, two human-translated, and two machine-translated texts.

	Tokens	MTLD
DEOrig	70504	63.3564
HRMTrans	64498	53.2950
ENMTrans	80588	59.6834
HRHumTrans	64388	72.8422
ENHumTrans	80680	62.2262

4.2 Lexical Frequency Profile (LFP)

Following the approach suggested in Vanmassenhove et al. (2021) by omitting the 'academic vocabulary' band of frequencies (Laufer, 1994; Laufer & Nation, 1995), we calculated the number of tokens in three frequency bands: the most frequent 1,000 words, the next 1,000 words by frequency (1,000–2,000) and the 'Beyond 2000' band. For our texts, there is also a fourth category, 'Missing', for the words that could not be associated with a frequency value, where the need for better resources for Croatian is especially visible.

Nevertheless, the first three frequency bands offer sufficient insight into the lexical richness of the original compared to all four translations and explain some of the findings from the MTLD analysis. Among all five texts, human- and machine-produced, Kafka uses the fewest words from the first frequency band (Top 1,000), only 9%. Both MTs use substantially more – 16.04% Croatian and 17.03% English. HTs are more similar to the MTs than to the original. From the second band (1,000–2,000), Kafka again draws the fewest words – 6.97%, with all the translations between 11.62% and 14.73%. Croatian fewer than English, and MTs fewer than HTs (Table 2).

Therefore, for the first two bands, we can conclude that translations are indeed somewhat lexically simpler than the original. But Kafka's true lexical diversity becomes particularly clear in the third band, 'Beyond 2,000', where 69% of his lexical types come from, more than in the English translations – 65.43% MT and 60.38% HT. Unfortunately, the Croatian translation could not be evaluated for this group of lemmas, because most of the words ended up in the group of 'Missing' frequencies.

Table 2
LFP for the German original and the human (HumTrans) and machine translations (MTrans) in Croatian and English

	Types	Tokens	Top 1,000	1,000–2,000	Beyond 2,000	Missing
DEOrig	5121	34172	461 (9.00%)	357 (6.97%)	3534 (69.01%)	769 (15.02%)
HRMTrans	4327	34710	694 (16.04%)	503 (11.62%)	590 (13.64%)	2540 (58.70%)
ENMTrans	3734	37088	636 (17.03%)	509 (13.63%)	2443 (65.43%)	146 (3.91%)
HRHumTrans	5101	33614	718 (14.08%)	520 (10.19%)	773 (15.15%)	3090 (60.58%)
ENHumTrans	3380	36822	644 (19.05%)	498 (14.73%)	2041 (60.38%)	197 (5.83%)

4.3 Concreteness, Age of Acquisition

4.3.1 Concreteness values for the German original text and translations

The values of the mean and the median for lexical concreteness (scale 1–7) show that for all translations, MT and HT, lexical choices are significantly more concrete than the original ($F(4.0,16013) = 2350$, $p < .001$, $\eta p^2 = 0.058$) and strikingly similar to each other. There is no statistical difference between them, which points towards the conclusion that the changes in the sensory profile of the text are not an inherent feature of machine translations.

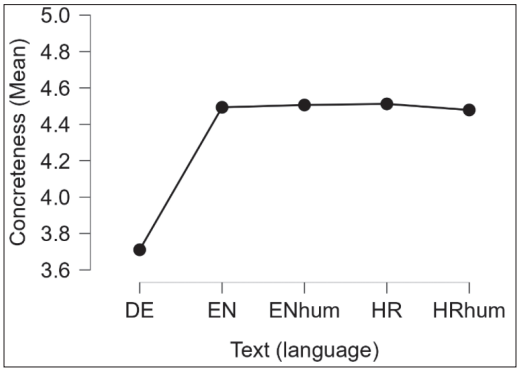
Table 3
Concreteness (C) descriptive statistics (mean, median, *SD*) for the German original and translations.

	Types	Tokens	C(*M*)	C(*Med*)	C(*SD*)
DEOrig	4352	32712	3.711	3.884	1.245
HRMtrans	1787	24897	4.512	4.550	1.349
ENMtrans	3588	36824	4.494	4.240	1.264
HRHumTrans	2011	23522	4.479	4.550	1.389
ENHumTrans	3183	36231	4.506	4.252	1.261

However, Levene's test for equality of variances was violated ($F(4,154181) = 442.6$, $p < .001$; therefore, Games-Howell post-hoc comparisons) and grouping of the standard deviations for concreteness (German original, English MT and English HT) on the one hand and both Croatian translations on the other hand shows that the Croatian translations are more variable in the lexical choice about concreteness. As Brglez and Vintar (2022) note, one should be cautious when interpreting similar results – high lexical variation in experienced human translators is usually controlled, deliberate

and precise, whereas machines use different words erratically, unreliably and inconsistently.

Figure 1
Concreteness values for the German original and Croatian and English human and machine translations in comparison



4.3.2 AoA values for the German original text and its translations

The analysis of the AoA statistical results further confirms that different language varieties (original vs. MT vs. HT) have a different impact on the developmental profile of the text ($F(4,19094) = 1247$, $p < .001$, $\eta p^2 = 0.030$), albeit with a small effect.

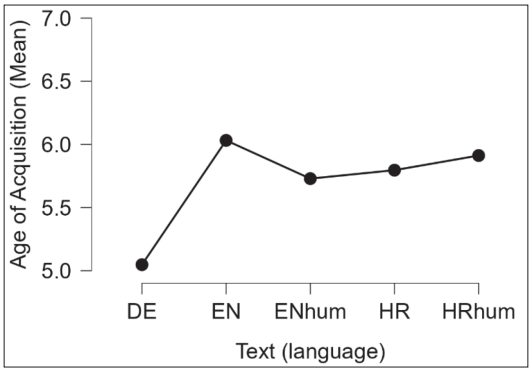
Table 4
AoA descriptive statistics (mean, median, *SD*) for the German original and translations.

	Types	**Tokens**	**A(*M*)**	**A(*Med*)**	**A(*SD*)**
DEOrig	4352	32712	5.048	4.500	2.066
HRMTrans	1787	24897	5.796	5.471	1.585
ENMTrans	3588	36824	6.032	5.390	2.257
HRHumTrans	2011	23522	5.912	5.533	1.724
ENHumTrans	3183	36231	5.729	5.188	2.116

The mean rating values for all the translations reveal that the lexical choices for the translations all cluster around an age almost a whole year older than that used by Kafka, even if the median results are roughly equivalent. For the translations, this indicates that at least a part of Kafka's use of early-acquired, yet abstract vocabulary is undergoing a clear shift towards the lexical material acquired later in life, altering the cognitive signature of the translations. This should not be taken as a direct measure of readability, but more as a developmental descriptor of the lexical material used in the original vs. translations. The standard deviations reveal, once again, the wider

variety of acquisition age for the German original and both English translations on the one hand, than for both Croatian translations on the other hand (Levene's test $p < .001$). Interestingly, the Croatian MT with the SD = 1.58 and the English MT with the SD = 2.258 sit at the extremes of this AoA continuum. Expectedly, Croatian is lagging behind the values of the original text (SD = 2.066) due to the lack of available resources, exhibiting the hallmarks of lexical levelling, and English, with its vast selection of models, offering more variability than required from the original and engaging in misguided 'creativity' (Brglez & Vintar, 2022).

Image 2
AoA values for the German original and Croatian and English human and machine translations in comparison



4.3.3 Examples of word patterns for concreteness and AoA

Since the present analysis explores systematic patterns in the texts, individual words can only be offered as illustrations of the observed tendencies and cannot be treated as direct examples of translation equivalents. To illustrate the findings, we selected the most frequent words within the lowest concreteness quartile, that is, words with concreteness values below the 25th percentile of the respective corpus, and the most frequent words with AoA values above the 75th percentile. Tables 5 and 6 show the most frequent items from these two groups for the German original and the four translations. When interpreting the examples, it should be kept in mind that Kafka's text is systematically less concrete, that is, more abstract, and uses vocabulary associated with an earlier age of acquisition. The differences are relatively small at the level of the means, but they are consistent across the corpus.

Table 5
Most frequent words below the 25th percentile (concreteness)

DE	HR	EN	HRhum	ENhum
tun	misliti	kind	ponešto	important
vergessen	ponekad	recognise	učiniti	progress
zuerst	kazati	easy	kriv	along

Table 6
Most frequent words above the 75th percentile (AoA)

DE	HR	EN	HRhum	ENhum
kaufmann	sudac	director	ured	defence
richter	radni	approach	pogledati	moment
vollständig	dokument	avoid	stajati	simply

5 CONCLUSION

As an answer to our original research question, the systematic nature of the results obtained in this analysis shows that psycholinguistic parameters, in this case concreteness and age-of-acquisition, can be considered additional markers when evaluating discrepancies between texts and their translations. In the present paper, we have shown that the translations deviate from the original in that they systematically shift the original values of concreteness and age of acquisition, and in doing so, exhibit different extents of variability, either higher or lower than the original. The results obtained on MTLD and the Lexical Frequency Profile measures are generally consistent with previous observations of reduced lexical richness and a stronger tendency towards frequent lexical items in machine translation (Vanmassenhove et al., 2019, 2021; Niu & Jiang, 2024). At the same time, they confirm that these tendencies are not necessarily uniform across languages, translations or metrics. In this process, all translations, both human- and machine-made, tend to be more concrete and use more mature vocabulary. However, the fact that this tendency is found in both human and machine translations suggests that it should not be regarded as an exclusive feature of machine translation.

Nevertheless, the nature of lexical variability in human translators and machines is not the same. For machines, it depends on the number of resources available for their training and results in erratic use of different words, whereas for humans, it is deliberate and stems from the skilful and thought-through use of synonyms. The present results support a distinction between these two patterns, but they do not imply the translators'

intentions or allow the observed differences to be attributed exclusively to training-resource size. Rather, they indicate that human and machine translations may differ in the distribution and variability of their lexical choices, even when their average psycholinguistic values are similar.

The concreteness results are compatible with psycholinguistic accounts that associate concrete vocabulary with sensory and perceptual grounding (Paivio, 2010; Peti-Stantić et al., 2021), as well as with research indicating greater semantic alignment among concrete than abstract translation equivalents (Chaouch-Orozco et al., 2024). In the present study, however, concreteness is used as a text-level indicator of the sensory-semantic profile of the lexical material, not as a direct measure of readers' processing or reception. The results for age of acquisition similarly confirm that different language varieties in the present corpus have a different impact on the developmental profile of the text. The higher mean AoA values found in all translations indicate a shift towards vocabulary estimated to have been acquired later in life. Since the median values are more similar, this shift may not represent a uniform replacement of early-acquired words, but may partly result from the distribution of later-acquired lexical items.

As such, concreteness and AoA comparisons between the original and the translation offer a deeper insight into the sensory and developmental dimensions of texts, complementing the more surface-oriented metrics such as MTLD and LFP. The comparison of variability is particularly informative, since similar mean values may conceal different patterns of lexical distribution. The present findings should merely be regarded as exploratory, but they indicate that psycholinguistic parameters can complement traditional corpus-based measures and provide a useful basis for further research into psycholinguistically based dimensions of literary translationese.

6 LIMITATIONS

The biggest constraint for the present study is the mismatch between rating studies, both within individual languages and between different languages. All the studies cover different parameters, so the ratings needed to be combined, which necessarily brings up the question of their comparability. However, the results for parameters that do overlap in some studies exhibit strong correlations, so the discrepancies should be minimal. In addition, most databases with human ratings cover only a few thousand words. Had we used them, the number of overlapping words would have been very low, so we decided to use databases with extrapolated values for many more

words (Conde et al., 2026), again, with high correlation with human estimates. Second, the analysis was performed on the averaged estimates of words in isolation. These can differ from the values in context, but collecting the ratings for words in context and assigning them to the same words in the same context would be even more complex than the present situation. Third, in contrast to German and especially English, Croatian only has one applicable psycholinguistic database, so that some analyses (notably LFP) could only be performed partially. Regardless of constraints, which, at present, are considerable, taking the route of psycholinguistic semantic dimensions offers promising new insights into the nature of translations, both human and machine-made.

REFERENCES

Alves, D., Thakkar, G., & Tadić, M. (2020). Evaluating language tools for fifteen EU-official under-resourced languages. In N. Calzolari, F. Bechet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 1866–1873). European Language Resources Association. https://aclanthology.org/2020.lrec-1.230/

Baker, M. (1996). Corpus-based translation studies: The challenges that lie ahead. In *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager* (pp. 175–188). John Benjamins Publishing Company.

Baker, M., Francis, G., & Tognini-Bonelli, E. (1993). Corpus linguistics and translation studies: Implications and applications. In *Text and Technology: In Honour of John Sinclair*. John Benjamins Publishing Company.

Baroni, M. (2005). A new approach to the study of translationese: Machine-learning the difference between original and translated text. *Literary and Linguistic Computing*, *21*(3), 259–274. https://doi.org/10.1093/llc/fqi039

Brglez, M., & Vintar, S. (2022). Lexical diversity in statistical and neural machine translation. *Information*, *13*(2). https://doi.org/10.3390/info13020093

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Conde, J., Martínez, G., Grandury, M., Arriaga, C., Haro, J., Schroeder, S., Hintz, F., Reviriego, P., & Brysbaert, M. (2026). Updating the German psycholinguistic word toolbox with AI-generated estimates of concreteness, valence, arousal, age of acquisition, and familiarity. *Journal of Cognition*, *9*(1), Article 9. https://doi.org/10.5334/joc.482

Gellerstam, M. (1986). Translationese in Swedish novels translated from English. In L. Wollin & H. Lindquist (Eds.), *Translation studies in Scandinavia: Proceedings from the Scandinavian Symposium on Translation Theory (SSOTT) II* (pp. 88–95). CWK Gleerup.

Juola, P. (2008). Assessing linguistic complexity. In M. Miestamo, K. Sinnemaki, & F. Karlsson (Eds.). John Benjamins Publishing Company. https://doi.org/10. 1075/slcs.94.07juo

Kafka, F. (1979). *Proces* (Z. Crnković, Trans.). Sveučilišna Naklada Liber.

Kafka, F. (2005). *The Trial* (D. Wyllie, Trans.). Project Gutenberg. https://www.gutenberg.org/cache/epub/7849/pg7849.txt

Kafka, F. (2022). *Der Prozess*. Project Gutenberg. https://www.gutenberg.org/cache/epub/69327/pg69327.txt

Koppel, M., & Ordan, N. (2011). *Translationese and Its Dialects*. https://www.semanticscholar.org/paper/Translationeseand-Its-Dialects-Koppel-Ordan/461ecb09a03b261a21b97465691f31924281aa4c

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, *25*(2), 21–33. https:// doi.org/10.1177 / 003368829402500202

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*(3), 307–322. https://doi.org/10.1093/applin/16.3.307

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

Niu, J., & Jiang, Y. (2024). Does simplification hold true for machine translations? A corpus-based analysis of lexical diversity in text varieties across genres. *Humanities and Social Sciences Communications*, *11*(1), 480. https://doi.org/10.1057/s41599-024-02986-7

Oshika, M., Morishita, M., Hirao, T., Sasano, R., & Takeda, K. (2024). Simplifying translations for children: Iterative simplification considering age of acquisition with LLMs. arXiv. In *Findings of the Association for Computational Linguistics: ACL 2024* (pp. 8567–8577). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-acl.506

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1 pt 2), 1–25. https://doi.org/10.1037/h0025327

Paschalidis, A. I. (2025). AI and the great linguistic flattening. In *Unesco Ideaslab*. https://philarchive.org/rec/PASAAT-7

Peti-Stantić, A., Anđel, M., Gnjidić, V., Keresteš, G., Ljubešić, N., Masnikosa, I., Tonković, M., Tušek, J., Willer-Gold, J., & Stanojević, M.-M. (2021). The Croatian psycholinguistic database: Estimates for 6000 nouns, verbs, adjectives and adverbs. *Behavior Research Methods*, *53*(4), 1799–1816. https://doi.org/10.3758/s13428-020-01533-x

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020, April). Stanza: A Python natural language processing toolkit for many human languages. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations* (pp. 101–108). Association for Computational Linguistics. https://doi.org/10.48550/arXiv.2003.07082

Team, J. (2025). JASP.

Teich, E. (2012). *Cross-linguistic variation in system and text: A methodology for the investigation of translations and comparable texts*. De Gruyter Mouton. https://www.perlego.com/book/653525/crosslinguistic-variation-in-system-and-text-pdf (Original work published 2012)

Terčon, L., Dobrovoljc, K., & Ljubešić, N. (2025). CLASSLA-Stanza: The next step for linguistic processing of South Slavic languages. *Contributions to Contemporary History*, *65*. https://doi.org/10.51663/pnz.65.3.05

Vanmassenhove, E., Shterionov, D., & Gwilliam, M. (2021). Machine translationese: Effects of algorithmic bias on linguistic complexity in machine translation. arXiv. https://doi.org/10.48550/arXiv.2102.00287

Volansky, V., Ordan, N., & Wintner, S. (2015). On the features of translationese. *Digital Scholarship in the Humanities*, *30*(1), 98–118. https://doi.org/10.1093/llc/fqt031

Psiholingvistički parametri kao specifična obilježja (strojnoga) prevoditeljskoga jezika

Maja Anđel
Filozofski fakultet Sveučilišta u Zagrebu

Činjenica da se strojno prevedeni tekstovi i tekstovi koje prevode ljudi po različitim parametrima razlikuju od izvornika dobro je dokumentirana. Postoji nekoliko metoda za procjenu toga fenomena i one pouzdano pokazuju njegovu sustavnost, ali sve su te metode primarno usredotočene na površinsku leksičku razinu. U ovome radu istražit ćemo potencijal psiholingvističkih leksičkih dimenzija konkretnosti i dobi usvajanja da obogate postojeće metode mjerenja uvidom u dublja obilježja prijevoda temeljena na semantici. U tu svrhu analizirali smo izvornik Kafkina romana „Proces" i usporedili ga s humanim prijevodima na hrvatski (Zlatka Crnkovića iz 1979.) i engleski (Davida Wyllieja iz 2005.) te sa strojnim prijevodima na hrvatski i engleski jezik. Rezultati pokazuju da se psiholingvistički leksički parametri konkretnosti i dobi usvajanja doista sustavno razlikuju od izvornika i za ljudske i za strojne prijevode, ali na različite načine. Stoga zaključujemo da odabrani parametri dosljedno ukazuju na individualna leksička obilježja tekstova te da bi vrijedilo dalje istražiti koliko nam novih informacija mogu pružiti.

Ključne riječi: *konkretnost, dob usvajanja, prevođenje, prijevod, mjera tekstualne leksičke različitosti – MTLD, profil leksičke frekvencije – LFP*