



UNIVERSITÄT
DES
SAARLANDES



Information Theory in Translation(ese) Studies

B7: Translation as Rational Communication

Maria Kunilovskaya

LTG research seminar, University of Oslo

December 17, 2024

Outline

- ① B7 context: IDEAL DFG Grant
- ② Translation(ese) Studies
- ③ Enter Information Theory
 - Previous work
- ④ Current research
 - Memory-surprisal trade-off (mono- and cross-lingual)
 - Impact of translation task difficulty
- ⑤ References

Information Density and Linguistic Encoding: Overview

Can language variation and use be explained using Information Theory (Shannon, 1948)?

Information Theory (IT) indices

The information content – or **surprisal** (Tukey, 1962) – of an individual event is calculated from its probability as

$$I(X) = -\log_2(\textit{probability})$$

The smaller the probability of an event, the greater its surprisal.

Average surprisal is **information density**, a measure of complexity used to approximate linguistic effort or processing cost.

Other IT measures: entropy, Kullback-Leibler divergence (KLD)

Surprisal as a proxy for cognitive processing difficulty

① Foundations: surprisal predicts self-paced reading times

- Difficulty of a word is proportional to its surprisal (Hale, 2001).
- Expectation-based account of human language processing (Levy, 2008).

“A word’s predictability, as measured by surprisal, is a major determinant of processing difficulty.” (Hahn et al., 2022)

② GPT2-based surprisal approximates other psycho- and neurolinguistic measures of human language processing such as

- N400 Effect, a specific brain response commonly associated with the brain’s reaction to unexpected words (Huber et al., 2024; Michaelov et al., 2024),
- continuation acceptability judgement task (Wallbridge et al., 2022),
- reading times and eye-tracking metrics (Jaeger, 2010; Hao et al., 2020).

③ Memory-surprisal trade-off (MST), “lossy memory”

- is an attempt to combine expectation-based and memory-based psycholinguistic accounts of human language processing:

“A certain level of average surprisal per word can only be achieved at the cost of storing some amount of information about past context.” (Hahn et al., 2022)

- Natural languages are structured to minimize the effort required to store and retrieve information while keeping surprisal as low as possible (Hahn et al., 2021)

Information Density and Linguistic encoding: Projects

Area-A Situational Context and World Knowledge

- A1: Neurobehavioural Correlates of Surprisal in Online Comprehension
- A5: Language Experience and Surprisal for Learning and Memory
- A8: Adapting Text Generation to Individual Users
- ...

Area-B Discourse and Register

- B1: Information Density in English Scientific Writing: A Diachronic Perspective
- B4: Modelling and Measuring Information Density
- B7: Translation as Rational Communication
- ...

Area-C Variation in Linguistic Encoding

- C1: Information Density and the Predictability of Phonetic Structure
- C4: Mutual Intelligibility and Surprisal in Slavic Intercomprehension
- C7: Cross-linguistic Information-Theoretic Modelling of Communicative Efficiency

(psycholinguistics, phonetics, language typology and contrastive studies, syntax and information structure, discourse analysis, MT and translation studies)

B7: Translation as Rational Communication

Rational communication: effective and efficient encoding is an optimal balance between the available cognitive resources and the amount of transmitted information.

Can information-theoretical account explain translationese deviations?

Can departures from TL norms be viewed as rational information management?
Does translated language exhibit the right amount of wrong?

Key aspects and directions:

cross-lingual modelling

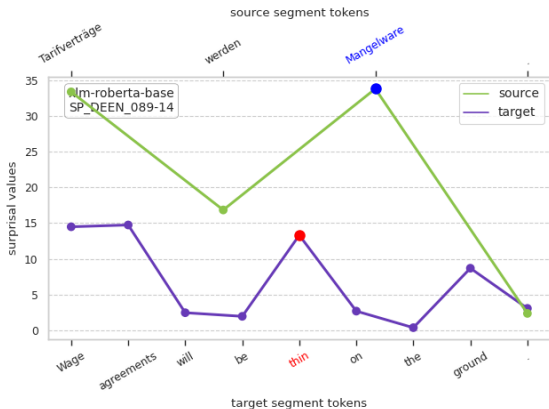
cross-modal comparisons

memory-surprisal tradeoff

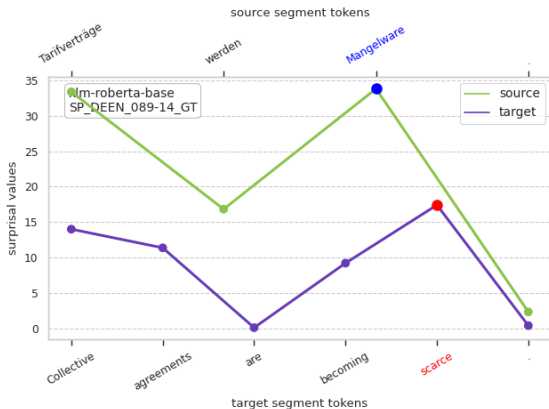
annotated disfluencies and translation latency

Surprisals from monolingual models for aligned segments

Human translation from EPIC-UdS corpus



Google Translate



Outline

- ① B7 context: IDeaL DFG Grant
- ② Translation(ese) Studies
- ③ Enter Information Theory
 - Previous work
- ④ Current research
 - Memory-surprisal trade-off (mono- and cross-lingual)
 - Impact of translation task difficulty
- ⑤ References

Translation (interpreting) and language variation

Mediated language –

- is language produced as a result of translation or (simultaneous) interpreting,
- differs in systematic ways from texts originally authored in the target language (TL),
- can be considered a TL subsystem ([Chesterman, 2017](#)).

Translationese –

- is systemic deviations of mediated language from the TL norm,
- reflects the specificity of the translation process,
- can be very subtle for an expert human eye ([Wein, 2023](#)) but is easily picked up by a machine, esp. at the document level (chunks of 450 or 2000 tokens), with accuracy over 85% on de-lexicalised features.

Translationese studies

Translation^{ese} studies aims to capture and explain the specificity of translations.

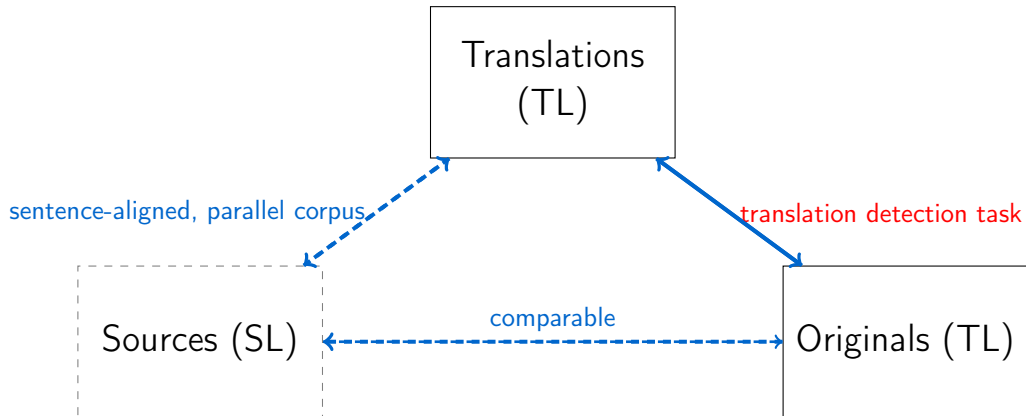
The usual suspects (aka translationese-inducing factors):

- translation direction (distance between SL and TL): e.g. $EN \rightarrow DE \neq DE \rightarrow EN$
- register: e.g. translationese in fiction vs. academic writing
- level of expertise/experience: e.g. professional vs student, professional vs. layman
- mode: translation vs interpreting

Underpinning theories:

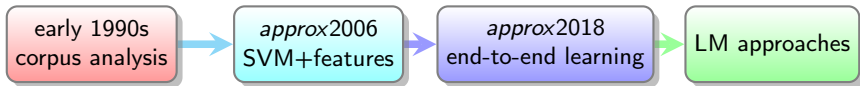
Systemic Functional Grammar, Relevance theory, contact and contrastive linguistics, psycholinguistics, risk management

Typical corpus resources for a translationese study



Methodological paradigms

- ① Statistical tests, single-feature classifiers (univariate analysis)
 - that-complementizer (Olohan, 2001)
 - phrasal verbs (Cappelle and Loock, 2017)
 - conjunctive markers, TTR, contracted forms, loan words (Redelinghuys and Kruger, 2015)
- ② Text classification on feature vectors (multivariate approach)
 - on hand-crafted theory-based features (Evert and Neumann, 2017),
 - on n-grams (Popescu, 2011; Baroni and Bernardini, 2005; Lapshinova-Koltunski and Zampieri, 2018),
- ③ Neural feature-learning classifiers to solve
 - translation detection task (Pylypenko et al., 2021)
 - SL detection task (Sominsky and Wintner, 2019; Dutta Chowdhury et al., 2021)



Described translationese trends (Chesterman, 2004)

- 'shining through' effect
- explicitation
- levelling-out
- lengthening
- simplification
- (over)normalization
- unique items hypothesis

'Shining-through' examples

Innenstadt → inner city
(alternative: city centre, downtown)

foreign accent jokes,
non-native language use

Outline

- ① B7 context: IDEaL DFG Grant
- ② Translation(ese) Studies
- ③ Enter Information Theory
 - Previous work
- ④ Current research
 - Memory-surprisal trade-off (mono- and cross-lingual)
 - Impact of translation task difficulty
- ⑤ References

Typology of research designs

... using information-theoretical (IT) indices such as surprisal, entropy, perplexity, Kullback-Leibler divergence (KLD) in TS

Derivation: IT indices from LM or based on normed frequencies, inc. of translation solution (Martínez and Teich, 2017; Wei et al., 2022)

Indexed item: text category (subcorpus) (Nikolaev et al., 2020; Martínez and Teich, 2017; Bizzoni and Lapshinova-Koltunski, 2021), document (Hu and Kübler, 2021), sentence (Rubino et al., 2016), token, e.g. discourse markers (Hu and Kübler, 2021; Lapshinova-Koltunski et al., 2022)

Usage: direct comparison of indices for items vs. features for a classifier to predict their labels

B7 Data: *German* ↔ *English*, European Parliament debates

EPIC-UdS (spoken): manually transcribed video recordings of parliamentary speeches and their simultaneous interpreting Przybyl et al. (2022),

Europarl-UdS (written): transcripts of speeches (adapted for reading) and their written translation from European Parliament website (extracted using Martinez's pipeline),

- original speeches delivered in L1,
- document- and segment-alignment,
- balanced across modes and directions,
- non-intersecting written and spoken originals,
- lemmatised with Stanza Qi et al. (2020),
- unified punctuation across modes,
- annotated spoken features removed.

		docs	segs	tokens	
				source	target
sp	DE-EN	165	3,247	56,142	49,265
	EN-DE	137	3,435	64,645	46,462
wr	DE-EN	170	2,796	67,726	77,427
	EN-DE	170	2,790	67,965	66,462

(after filtering and pre-processing)

12 K segments, ca. 250 K tokens in each language

Surprisal-based translation detection

preprocessing options: raw tokens, Stanza tokens, Stanza lemmas,

setup: train on originals (3M tokens) separately for each language, test on mediated sets and held-out originals (250K tokens)

models: KenLM 4gram, LSTM 4gram, (pretrained) GPT2 for EN ([Schweter, 2020](#)) and DE(v1.0) ([Radford et al., 2019](#)), and multilingual GPT2

- SVM with linear kernel
- Baseline: 4 known translationese indicators (TTR, lexical density, mean dependency length, mean hierarchical length)
- Surprisal-based features: mean segment surprisal (AvS), min and max, standard deviation
- Granularity: segment and document levels

Evaluation results: binary SVM classification

Best-performing approaches vs baseline (min seg. length: 5)

Segment level F1-score

		baseline	surprisal	best setup
de	sp	55.24	54.19	KenLM(tok)
	wr	53.41	56.72	KenLM(tok)
en	sp	52.17	53.63	LSTM(lem)
	wr	50.76	50.7	GPT2(detok)

Document level F1-score

		baseline	surprisal	best setup
de	sp	75.43	67.91	KenLM(tok)
	wr	66.32	69.27	KenLM(tok)
en	sp	67.3	62.85	LSTM(lem)
	wr	63.07	55.03	KenLM(tok)

Translated German is more detectable than translated English.

For baseline traditional features spoken is more detectable than written,
for surprisal-based features - no consistency.

How much information gets through (Kunilovskaya et al., 2023b)

(Source-Target surprisal correlations from parallel data)

Do the conditions of the mediation influence the information density (ID) of the message?

– based on average segment surprisal (AvS) from n-gram LM –

Mediation conditions

- written, spoken
- spoken: read out vs impromptu sources
- spoken: delivery of source (words/min)

Comparison methods

- regression,
- correlation analyses,
- statistical tests

Expectations:

- Positive association between sources and targets
- Lower less-correlated AvS in interpreting than in translation due to fluent simplified production.
- Higher more source-correlated AvS in intuitively easier conditions (impromptu, low speaker's speech rate)

Is mediated language simpler or more complex than comparable non-mediated target language?

Experimental Setup

Modelling decisions

- Each LM (de, en) uses all available data (originals, translation)
- Point-wise probabilities are based on the counts from the entire corpus, **excluding the current document**,
- OOV: backoff to lower-order n-grams,
- *Hapax legomena* → *UKN*,
- n-grams respect sentence boundaries.

4-gram statistical LMs
for English and German



Average surprisal (AvS)
across segment tokens

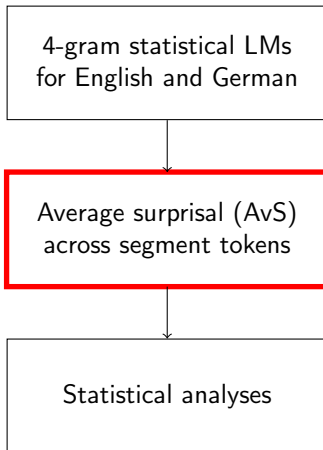


Statistical analyses

Experimental Setup

Surprisal as a measure of information density.

$$S(w_i) = -\log_2(P(w_i|w_{i-3}, w_{i-2}, w_{i-1}))$$



Experimental Setup

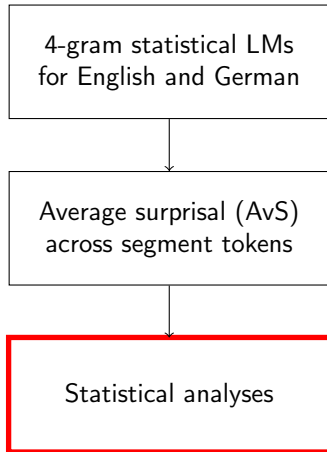
Hypotheses Estimation

cross-lingual analyses of AvS for aligned segs

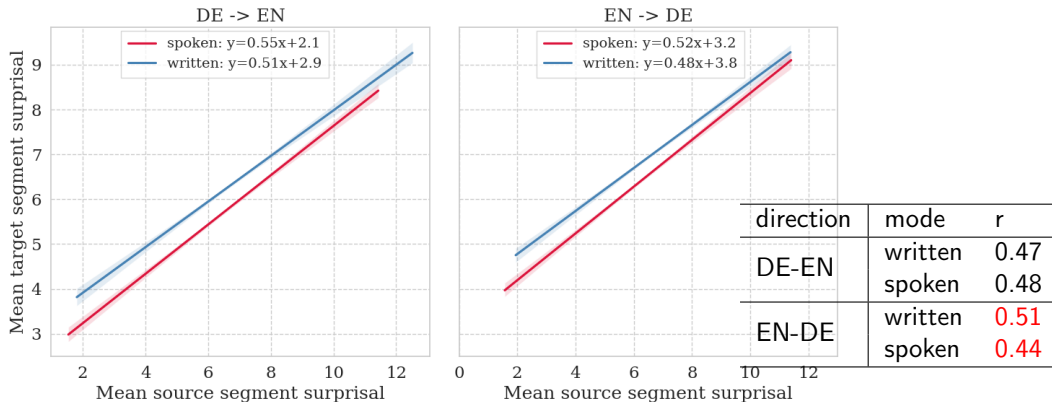
- linear regression,
- correlation analyses (Spearman r)

monolingual comparisons –

- statistical significance tests



Association sources – targets

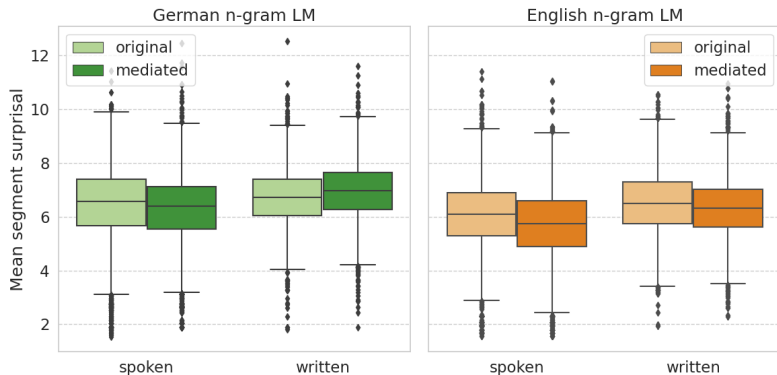


There is a reasonably strong association between sources and targets.

Given the same information input, translators produce more informative output.

Slope: +0.5 bit in output per 1 bit of input. → Mediated language is denser than sources!

Are mediated texts simpler than comparable TL originals?



Mediated language has lower levels of surprisal than TL originals (except for written into German).

All results are statistically significant.

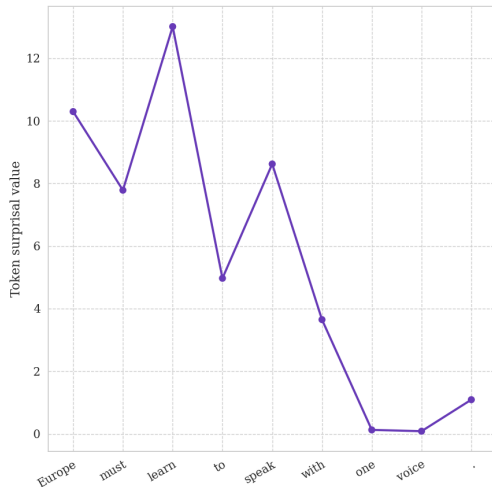
Examples: low-surprisal spoken targets – simplification

(1)

source: Europa muss lernen, mit einer Stimme zu sprechen und dann auch mit einer Position zu handeln.

translation: Europe must learn to speak with one voice and to take united action.

interpreting: Europe must learn to speak with one voice $\emptyset$. (AvS = 5.52)



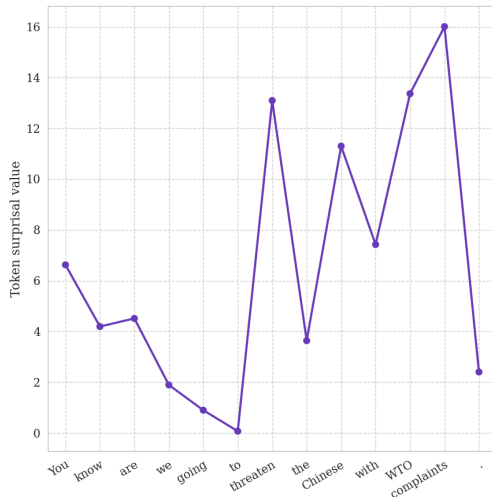
Discussion: high-surprisal spoken targets – interference

(2)

source: Wollen wir den Chinesen
mit WTO **Klagen** drohen.

translation: Do we want to threaten the
Chinese with World Trade
Organisation (WTO)
sanctions?

interpreting: You know are we going to
threaten the Chinese with
WTO **complaints** (AvS =
6.57).



Findings from previous work

- ① IT indices yield additional evidence for translationese trends, especially shining-through ([Nikolaev et al., 2020](#); [Hu and Kübler, 2021](#)) and normalization, esp for EN>DE [Bizzoni et al. \(2020\)](#), and for the impact of professional expertise on the properties of translations.
- ② Surprisal from attempted modelling approaches was too crude to capture the differences between translations and non-translations ([Rubino et al., 2016](#); [Hu and Kübler, 2021](#)) and own.
- ③ Information density of source and target are reasonably well correlated ([Kunilovskaya et al., 2023b](#)).
- ④ Interpreting has lower information density ($\approx$ less complex, requires less effort to process) than translation (for the same input).
- ⑤ Mediated language has lower information density than comparable TL non-translations (except in EN-DE).
- ⑥ Any mediation generates denser output than aligned sources.

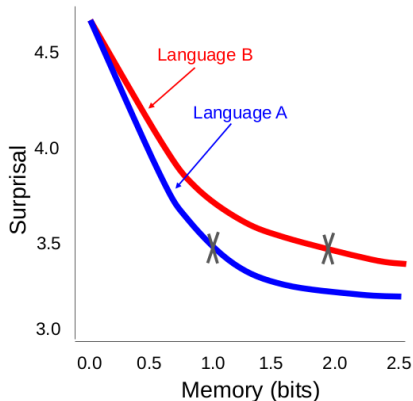
NB! 5 and 6 contradict other findings, inc. in [Kunilovskaya et al. \(2023a\)](#)

Outline

- ① B7 context: IDEAL DFG Grant
- ② Translation(ese) Studies
- ③ Enter Information Theory
 - Previous work
- ④ Current research
 - Memory-surprisal trade-off (mono- and cross-lingual)
 - Impact of translation task difficulty
- ⑤ References

Rationality-based approach to translationese

MST: A more optimal system (Language A) requires less memory to generate the same level of surprisal.



(Hahn et al., 2021)

Applications:

- 1 comparing language varieties (languages exhibiting stronger information locality have more efficient tradeoff)
- 2 prediction of human processing patterns (e.g. RT) for fine-grained cases of nested structures with various conditions for semantic compatibility or embedding bias

The report that the doctor annoyed the patient was surprising.

The report that the doctor **cured** the patient was surprising (incompatible verb).

The **fact that** the doctor cured the patient was surprising (high embedding bias)

Corpus-level comparisons after Hahn et al. (2021)

Memory-Surprisal Tradeoff (MST) based on monolingual n-gram models.

Goal: assess the (relative) overall efficiency of translation/interpreting compared to original language use.

Important setup decisions

train content: written_orig+tran or written_orig?

preprocessing: raw vs. lemma vs. lemmos(lowercase, tokenised MWE, xxxx for NUMs);

type of instance: segment vs. document.

Memory calculation for each n-gram model of order N ($N \in [2, 6]$):

e.g if $N=3$:

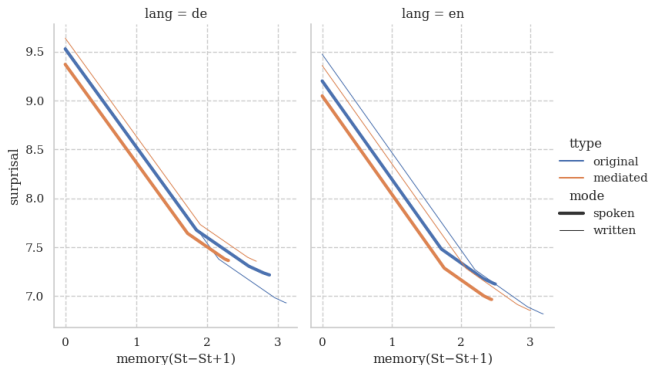
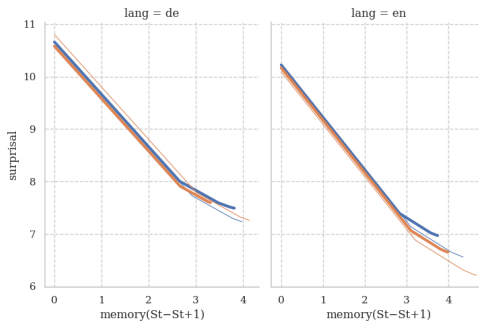
$3\text{gram_ngram_memory} = 2\text{gram_ngram_memory} + 3\text{gram_memory_rate}$,

where $3\text{gram_memory_rate} = 2 * (2\text{gram_AvS} - 3\text{gram_AvS})$

MST relational plots

(Best) training setting: orig, **lempos**, doc

Training setting: orig+tran, **raw**, seg



Interpretation:

Spoken mediated language is the most optimal language variety.

Towards a cross-lingual approach to MST

To model a cross-lingual process we need to take into account sources, e.g. using a MT model.

Processing stages in translation =

source comprehension

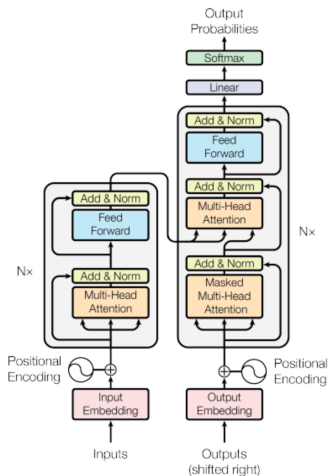
+ transfer difficulty

+ difficulty of TT generation

How to factor in source language input to represent production difficulty in translation more realistically?

Goal: Get fine-grained word-level indicators / triggers of translationese and compare varieties (translated and non-translated language).

A sketch of MT model morphology



Vaswani et al. (2017)

`MarianMTModel.from_pretrained`
("Helsinki-NLP/opus-mt-de-en")

3 modules by 6 layers by 8 heads

3 X X 6



`MarianForCausalLM.from_pretrained`
("Helsinki-NLP/opus-mt-en-de")

Memory Definition

Memory (I_{mem}) is the amount of information (in bits) stored in the N context words.
 $N_{\text{context tokens}}$ is a proxy for memory limits (capacity).

source memory bits: from encoder self-attention weights

target memory bits: from cross-attention weights + decoder self-attention weights

Higher attention weights mean higher probabilities assigned to the context words in the output probability distribution, easier prediction and lower surprisal, i.e Pearson $r(I_{\text{mem}}, \text{Srp}) < 0$.

$$I_{\text{mem}} = -1 * \sum \log_2(\text{prob}(w_n))$$

where N can be:

- ① highlights: N most attended tokens (“lossy memory”),
- ② distance-weighted highlights: N most attended tokens * `index_from_EoS`,
- ③ recency: (for self-attentions only) N most recent tokens.

How do we evaluate the models sanity?

Evaluation setups:

(based on expectations from Information Theory, Translationese Studies)

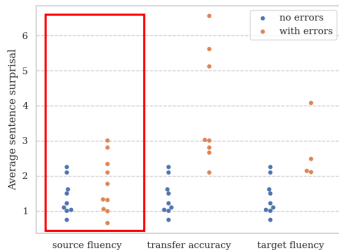
- ① MST: Pearson $r(I_{\text{mem}}, Srp) < 0$,
- ② Interpreting has steeper (=more optimal) MST than translation (based on the same sources),
- ③ Interpreting has more optimal MST than original speeches,
- ④ Spoken has more optimal MST than written (against comparable non-translations),
- ⑤ 'More-translated' (e.g. literal) text is more surprising than 'less-translated' (fluent), and has lower I_{mem} (as bigger memory capacity is needed to store the context for an effective rendition).

Is the source fluency captured?

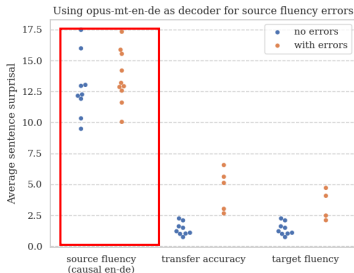
Each dot represents AvS (y-axis) for the target based on a pair of sentences with three types of problems (x-axis).

Colours help see which type of errors the MT model indices capture.

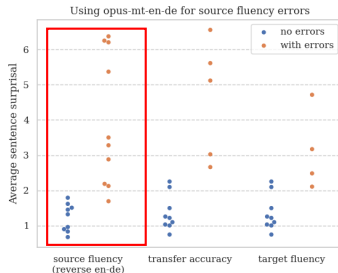
Default de-en MT



Reverse en-de MT as causal



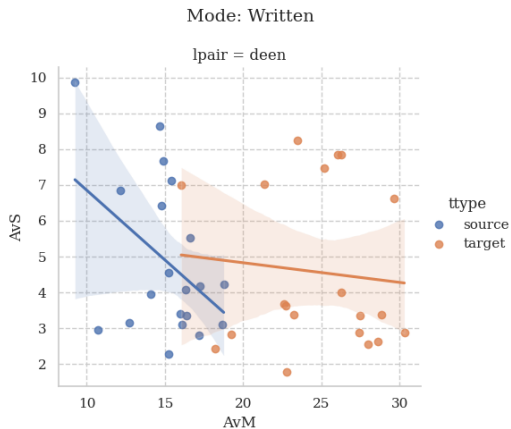
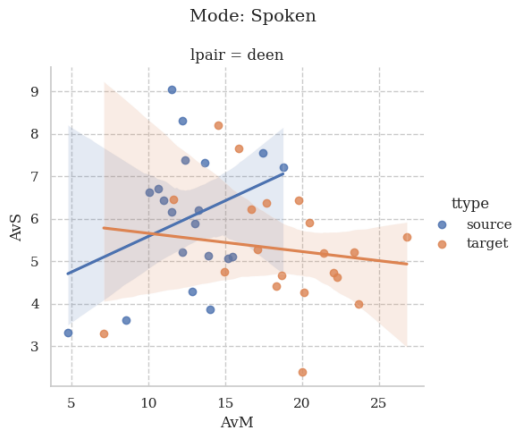
Reverse en-de MT



Solution: derive surprisal for the source from the output of the reverse MT model (Marian MT de-en and en-de have $\approx$ the same vocabulary).

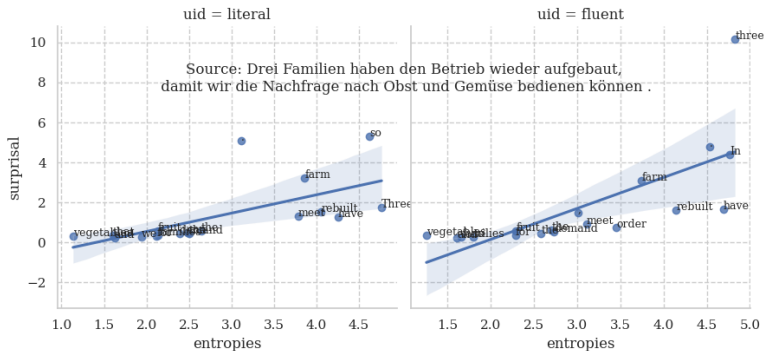
Results: No consistency in pilot on spoken-written modes

Settings: sentence-level, max-pooled+softmaxed layers, memory limit: highlights-3, SentencePiece subwords aggregated to get raw tokens indices.



Results on fluent and literal targets: Work in progress

Association between model surprisal and entropy for words in literal and fluent targets



Literal: Three families have rebuilt the farm so that we can meet the demand for fruit and vegetables.

Fluent: In order to meet the demand for fruit and vegetables, three families have rebuilt the farm.

① B7 context: IDEaL DFG Grant

② Translation(ese) Studies

③ Enter Information Theory

- Previous work

④ Current research

- Memory-surprisal trade-off (mono- and cross-lingual)

- Impact of translation task difficulty

⑤ References

Is translationese linked to ST difficulty?

Aim –

explore the local processing difficulty of a translation task as a predictor for the translational properties of the target documents.

Contributions/Novelty –

- cross-lingual approach to translationese (not multilingual typical in empirical TS),
- focused on operational units in translation (aligned phrases, not AvS or isolated words),
- indexing two subprocesses in translation: ST comprehension and SL-TL transfer (in line with MT approaches, not commonly used in empirical TS).

Methodology: Step by step

- 1 Represent each ST-TT document pair as a sequence of subsentential translation units.
 - translation units are aligned dependency-based subtrees and/or content words

surprisal, alignment:
lemposrel for entropy:

We believe sufficient –
PRON believe ccomp –

Wir sind der Auffassung ausreichen
PRON sein DET Auffassung ccomp

- 2 Index aligned ST items with translation difficulty:
 - SL comprehension: surprisal from SL-specific GPT-2 and
 - transfer difficulty: entropy of translation solutions from and alignment score.
- 3 Capture amount of TT translationese as “probability of being a translation” from a translation detection task.
 - use a contrastive sample of ‘more-translated/less-translated’ documents.
- 4 Use ST representations to predict TT probability of being a translation.
 - learning setups:
SVM on document averages vs. NN (encoder+regressor) on 3D variable-size vectors for document items
- 5 Evaluate: Is processing (translation) difficulty related to translationese properties of documents?

Target scores: Neural vs. SVM translationese classifiers

Data size: 3000 docs for each language pair (length min 450 words)

	model	accuracy	f1_score
de en	deberta multi ¹	96.67	96.66
		90.22	90.22
de en	distilbert multi ²	94.22	94.22
		90.66	90.66
de	distilbert de ³	96.89	96.89
en	distilbert en ⁴	92.89	92.89
de	SVM (best 34)	88.8	88.79
en	SVM (best 51)	79.73	79.72

BTW: Is out of English easier? See Bugliarello (2020)

¹<https://huggingface.co/microsoft/mdeberta-v3-base>, (He et al., 2021)

²<https://huggingface.co/distilbert/distilbert-base-multilingual-cased>

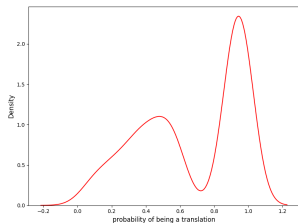
³<https://huggingface.co/distilbert/distilbert-base-german-cased>

⁴<https://huggingface.co/distilbert/distilbert-base-cased>

Distributions of probabilities from SVM and NN+LLM

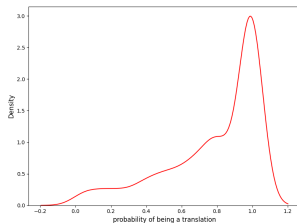
DE→EN Distributions

Morphsyntactic features + SVM

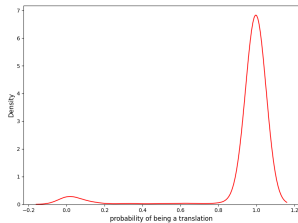


EN→DE Distributions

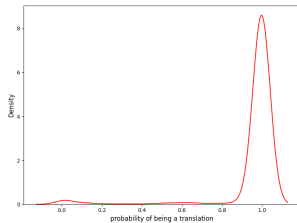
Morphsyntactic features + SVM



DistilBERT



DistilBERT



SVM on doc-averaged IT features for source items

Information-theoretical (IT) indices

- surprisal
- alignment score
- translation entropy

lpair	item	Pearson	MAE	features
deen	trees	0.17 ± 0.07	0.26 ± 0.01	IT
ende		-0.08 ± 0.07	0.21 ± 0.02	
deen	words	0.16 ± 0.09	0.26 ± 0.01	
ende		0.02 ± 0.12	0.21 ± 0.02	
deen	trees+words	0.15 ± 0.09	0.27 ± 0.01	
ende		0.02 ± 0.04	0.21 ± 0.02	

All feature values are close to normally distributed.

Observation: In DE→EN, there is a weak but consistent relation between ST complexity (regardless of item) and TT probability of being a translation.

SVM results: syntactic vs. IT complexity features

IT indices and/or syntax features

- surprisal
- translation entropy
- alignment score
- branching factor
- max tree depth
- mdd
- mhd

lpair	item	Pearson	Spearman	MAE	features
deen	trees	0.24±0.07	0.27±0.06	0.25±0.01	IT+syntax
ende		0.16±0.07	0.17±0.05	0.21±0.01	
deen	words	0.29±0.08	0.30±0.07	0.25±0.01	
ende		0.18±0.11	0.21±0.09	0.21±0.01	
deen	trees+words	0.27±0.10	0.28±0.08	0.25±0.02	
ende		0.18±0.11	0.18±0.10	0.21±0.01	
deen	—	0.30±0.10	0.31±0.09	0.24±0.01	syntax
ende		0.18±0.08	0.18±0.07	0.20±0.01	

Syntactic features perform better but the probabilities of being a translation are derived from a classifier based on morphosyntactic feature.

Summary

- ① Surprisal/entropy features **from word-based LMs are inferior** to other features in translation detection classification tasks.
- ② More **coarse-grained** setups based on entropy/perplexity measures for entire languages/documents are more promising.
- ③ Cross-lingual models were not tortured enough.
- ④ Non-LM approaches (entropy from frequency lists), including defining the space of translation solutions, are promising (also based on ([Kunilovskaya et al., 2025](#))).
- ⑤ There is a **lack of psychometric experiments** to shed light on what info-theoretical metrics actually capture in the context of cross-lingual communication, where production and comprehension costs need to be distinguished.

Thank you!

Information Theory in Translation(ese) Studies

SFB 1102 – Information Density and Linguistic Encoding (IDeaL)
funded by the Deutsche Forschungsgemeinschaft, Project ID 232722074

Questions?

Maria Kunilovskaya
maria.kunilovskaya@uni-saarland.de

Outline

- ① B7 context: IDeaL DFG Grant
- ② Translation(ese) Studies
- ③ Enter Information Theory
 - Previous work
- ④ Current research
 - Memory-surprisal trade-off (mono- and cross-lingual)
 - Impact of translation task difficulty
- ⑤ References

References I

- Baroni, M. and Bernardini, S. (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.
- Bizzoni, Y., Juzuk, T., Espã na Bonet, C., Dutta Chowdhury, K., van Genabith, J., and Teich, E. (2020). How Human is Machine Translationese? Comparing Human and Machine Translations of Text and Speech. In Proceedings of the 17th International Conference on Spoken Language Translation (IWSLT), pages 280–290. Association for Computational Linguistics.
- Bizzoni, Y. and Lapshinova-Koltunski, E. (2021). Measuring Translationese across Levels of Expertise: Are Professionals more Surprising than Students? In Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa), pages 53–63, Reykjavik. Linköping University Electronic Press.
- Cappelle, B. and Loock, R. (2017). Typological differences shining through : The case of phrasal verbs in translated English. In De Sutter, G., Lefer, M.-A., and Delaere, I., editors, Empirical Translation Studies. New Theoretical and Methodological Traditions, pages 235–264. Mouton de Gruyter.

References II

- Chesterman, A. (2004). Hypotheses about translation universals. Claims, Changes and Challenges in Translation Studies, pages 1–14.
- Chesterman, A. (2017). Reflections on Translation Theory: Selected papers 1993 – 2014.
- Dutta Chowdhury, K., España-Bonet, C., and van Genabith, J. (2021). Tracing source language interference in translation with graph-isomorphism measures. In Mitkov, R. and Angelova, G., editors, Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021), pages 375–385, Held Online. INCOMA Ltd.
- Evert, S. and Neumann, S. (2017). The impact of translation direction on characteristics of translated texts: A multivariate analysis for English and German. Empirical Translation Studies: New Methodological and Theoretical Traditions, 300:47–80.
- Hahn, M., Degen, J., and Futrell, R. (2021). Modeling word and morpheme order in natural language as an efficient trade-off of memory and surprisal. Psychological Review, 128(4):726.
- Hahn, M., Futrell, R., Levy, R., and Gibson, E. (2022). A resource-rational model of human processing of recursive linguistic structure. Proceedings of the National Academy of Sciences, 119(43):e2122602119.

References III

- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics, pages 159–166, Stroudsburg, PA. Association for Computational Linguistics.
- Hao, Y., Mendelsohn, S., Sterneck, R., Martinez, R., and Frank, R. (2020). Probabilistic predictions of people perusing: Evaluating metrics of language model performance for psycholinguistic modeling. In Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics, pages 75–86.
- He, P., Gao, J., and Chen, W. (2021). Debertav3: Improving deberta using electra-style pre-training with gradient-disentangled embedding sharing.
- Hu, H. and Kübler, S. (2021). Investigating Translated Chinese and Its Variants Using Machine Learning. Natural Language Engineering, 27(3):339–372.
- Huber, E., Sauppe, S., Isasi-Isasmendi, A., Bornkessel-Schlesewsky, I., Merlo, P., and Bickel, B. (2024). Surprisal from language models can predict erps in processing predicate-argument structures only if enriched by an agent preference principle. Neurobiology of Language, 5(1):167–200.

References IV

- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. Cognitive psychology, 61(1):23–62.
- Kunilovskaya, M., Mitkov, R., and Wandl-Vogt, E. (2023a). Cross-lingual Mediation: Readability Effects. In Štajner, S., Saggio, H., Shardlow, M., and Alva-Manchego, F., editors, Proceedings of the Workshop on Text Simplification, Accessibility, and Readability (TSAR-2023), pages 33–43. INCOMA Ltd., Shoumen, Bulgaria.
- Kunilovskaya, M., Przybyl, H., Teich, E., and Lapshinova-Koltunski, E. (2023b). Simultaneous Interpreting as a Noisy Channel: How Much Information Gets Through. In Proceedings of the International Conference on Recent Advances in Natural Language Processing, pages 608–618. INCOMA Ltd.
- Kunilovskaya, M., Zaitova, I., Xue, W., Stenger, I., and Avgustinova, T. (2025). Predictability of microsyntactic units across slavic languages: A translation-based study. In The Joint 25th Nordic Conference on Computational Linguistics and 11th Baltic Conference on Human Language Technologies.

References V

- Lapshinova-Koltunski, E., Polkläsener, C., and Przybyl, H. (2022). Exploring explicitation and implicitation in parallel interpreting and translation corpora. The Prague Bulletin of Mathematical Linguistics, 119:5–22.
- Lapshinova-Koltunski, E. and Zampieri, M. (2018). Linguistic features of genre and method variation in translation: A computational perspective. The grammar of genres and styles: from discrete to non-discrete units, 320:92–117.
- Levy, R. (2008). Expectation-based syntactic comprehension. Cognition, 106(3):1126–1177.
- Martínez, J. M. M. and Teich, E. (2017). Modeling routine in translation with entropy and surprisal: A comparison of learner and professional translations. In Cercel, L., Agnetta, M., and Lozano, M. T. A., editors, Kreativität und Hermeneutik in der Translation, pages 403–427. Narr Francke Attempto Verlag.
- Michaelov, J. A., Bardolph, M. D., Van Petten, C. K., Bergen, B. K., and Coulson, S. (2024). Strong prediction: Language model surprisal explains multiple n400 effects. Neurobiology of Language, 5(1):107–135.
- Nikolaev, D., Karidi, T., Kenneth, N., Mitnik, V., Saeboe, L., and Abend, O. (2020). Morphosyntactic predictability of translationese. Linguistics Vanguard, 6(1):1–12.

References VI

- Olohan, M. (2001). Spelling out the optionals in translation : a corpus study. UCREL Technical Papers, (13):423–432.
- Popescu, M. (2011). Studying translationese at the character level. In Proceedings of the International Conference Recent Advances in Natural Language Processing (RANLP 2011), pages 634–639.
- Przybyl, H., Lapshinova-Koltunski, E., Menzel, K., Fischer, S., and Teich, E. (2022). EPIC-UdS - creation and applications of a simultaneous interpreting corpus. In Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022), pages 1193–1200, Marseille, France. ELDA.
- Pylypenko, D., Amponsah-Kaakyire, K., Dutta Chowdhury, K., van Genabith, J., and España-Bonet, C. (2021). Comparing feature-engineering and feature-learning approaches for multilingual translationese classification. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pages 8596–8611, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

References VII

- Qi, P., Zhang, Y., Zhang, Y., Bolton, J., and Manning, C. D. (2020). 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.
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8):9.
- Redelinghuys, K. and Kruger, H. (2015). Using the features of translated language to investigate translation expertise. a corpus-based study. International Journal of Corpus Linguistics.
- Rubino, R., Lapshinova-Koltunski, E., and van Genabith, J. (2016). Information density and quality estimation features as translationese indicators for human translation classification. In Proceedings of NAACL HT 2006, pages 960–970, San Diego, California.
- Schweter, S. (2020). German gpt-2 model.
- Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27:379–423, 623–656.

References VIII

- Sominsky, I. and Wintner, S. (2019). Automatic Detection of Translation Direction. In Mitkov, R. and Angelova, G., editors, Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019), pages 1131–1140, Varna, Bulgaria. INCOMA Ltd.
- Tukey, J. W. (1962). The future of data analysis. In Breakthroughs in Statistics: Methodology and Distribution, pages 408–452. Springer.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. (2017). Attention is all you need. Advances in neural information processing systems, 30.
- Wallbridge, S. C., Lai, C., and Bell, P. (2022). Investigating perception of spoken dialogue acceptability through surprisal. In INTERSPEECH, pages 4506–4510.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., Zhou, D., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. Advances in Neural Information Processing Systems, 35:24824–24837.

Wein, S. (2023). Human raters cannot distinguish English translations from original English texts. In Bouamor, H., Pino, J., and Bali, K., editors, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, pages 12266–12272, Singapore. Association for Computational Linguistics.