Transparency

All idioms were manually annotated for transparency. It was operationalized as the degree of lexical overlap between the literal translation of a Russian idiom and its idiomatic equivalent in the target language (English or German).

Annotations were conducted by comparing translations word-by-word between (i) the literal translation of the Russian idiom and (ii) the established idiomatic expression in the target language, as provided by M. Dubrovin.

Transparency was initially annotated by a single annotator. All annotations were subsequently double-checked at least once by an additional annotator fluent in the respective target language to ensure consistency and linguistic validity.

Transparency was treated as a discrete variable with three values:

- Transparent (1): the literal translation is lexically identical to the idiomatic equivalent, with all core content words preserved across languages (e.g., *белая ворона* "white crow" → *ein weißer Rabe*). In this case, both the color adjective and the noun are directly shared between the literal Russian expression and the German idiom.

- Partially transparent (0.5): the literal and idiomatic translations share the main lexical metaphor but differ by one or two lexical elements, such as the verb or argument structure (e.g., *бросать слова на ветер* "to throw words to the wind" → *in den Wind reden* "to speak into the wind"). Here, the metaphor of *wind* is preserved, while the verbal construction differs.

- Intransparent (0): there is no clear lexical correspondence between the literal and idiomatic translations (e.g., *бабушка надвое сказала* → *das is noch gar nicht heraus*). In this case, neither the lexical items nor the underlying imagery overlap (if exists) between the literal Russian expression and the German idiom.

For some idioms, multiple idiomatic equivalents were attested in the target language. In such cases, the transparency value assigned to the idiom corresponded to the highest transparency level observed among its equivalents. That is, if at least one equivalent exhibited partial or full lexical overlap with the literal translation, the idiom was annotated accordingly, irrespective of the

transparency of other available equivalents.

Across the dataset, English idioms exhibited a larger number of available equivalents than German idioms, and a higher proportion of these English equivalents were fully transparent. This asymmetry likely reflects broader cross-linguistic differences between the two target languages.