

# Imposing Coherence: AI and the Construction of Meaning in Greek–Arabic Texts

Doaa Elalfy

PhD Candidate, Hebrew University of Jerusalem

Generative AI has rapidly transformed access to ancient texts, including complex bilingual corpora such as Greek–Arabic medical writings. Yet its apparent success in translating, summarising, or aligning such material often rests on a critical assumption: that textual units are stable, segmentable, and internally coherent. In the context of translation, this assumption becomes particularly problematic. Greek–Arabic textual transmission is shaped by ambiguity, fragmentation, and the coexistence of multiple possible interpretations, rather than by direct equivalence between source and target texts. Work on Greek–Arabic lexical data shows that meaning is not confined to clearly bounded segments. Greek frequently encodes reasoning in syntactically compressed forms—through participial constructions and implicit inferential links—while Arabic translations may redistribute these relations across more explicit and sequential structures. As a result, correspondence between the two languages cannot be reduced to straightforward alignment or segmentation, but requires repeated interpretative decisions about where a unit begins and ends, what constitutes a coherent passage, and how inferential relations are distributed across fragmented textual evidence. AI systems, however, tend to treat such material as if it consisted of stable and well-defined units. In doing so, they flatten ambiguity, resolve uncertainty prematurely, and generate outputs that appear coherent and authoritative, while obscuring the instability, fragmentation, and plurality inherent in the data. These outputs conceal the interpretative work required in translation, presenting results as if they were derived from self-evident units rather than from competing interpretative possibilities. This paper argues that this mismatch stems from a more fundamental issue: the instability observed in such material lies not only in the texts themselves, but in the units through which they are constructed and interpreted. In Greek–Arabic corpora, such units do not exist as clearly bounded entities, but must be actively produced through interpretative decisions concerning scope, coherence, and the distribution of meaning across linguistically and conceptually heterogeneous material. Drawing on work within the Greek–Arabic Lexicon (GALex), the paper examines the *Belegstelle*—the unit linking Arabic textual segments to specific Greek references—as a case study. While such units may appear to correspond to discrete passages, in practice they are defined through segmentation and alignment shaped by how meaning is distributed across the two languages. Alignment is not a mechanical procedure but an interpretative act that negotiates partial correspondences, asymmetries, and competing possibilities of meaning in translation. A single Greek passage may correspond to multiple Arabic segments, or vice versa, and the boundaries of a *Belegstelle* must therefore be determined through interpretative decisions about coherence, scope, and context. In practice, segmentation

involves navigating partial correspondences and non-one-to-one relationships between Greek and Arabic. The identification of a Belegstelle thus depends not only on textual proximity but on interpretative judgments about how reasoning is structured and where coherence can be established within a plurality of possible meanings. Preliminary pilot testing with AI-assisted segmentation further highlights this mismatch. While such systems assume stable and clearly defined textual units, they impose consistency on data that is inherently variable, fragmented, and context-dependent, resolving ambiguity prematurely and producing outputs that fail to reflect the interpretative processes through which meaning is constructed. AI systems thus risk imposing coherence where none exists, obscuring the uncertainty that fundamentally structures the data. This is not merely a technical limitation but a methodological concern: by treating textual units as stable and self-evident, AI systems bypass the interpretative work through which meaning is produced in translation, replacing scholarly judgment with an illusion of clarity and reducing the plurality inherent in the process.