ISSN:2828-6235
Impact factor:13.77

International Journal of Literature and Language Studies(IJLLS)

DISCOURSE BOUNDARIES IN ENGLISH–UZBEK AI PROMPTS: DISTINGUISHING INSTRUCTIONS FROM QUOTED MATERIAL

Abduraximova Gulnoza Xolmirzayevna
Qo‘qon universiteti Andijon filiali,
Tibbiyotda tillarni o‘qitish kafedrasi o‘qituvchisi
ORCID: 0009-0001-5974-7825
E-mail: Gulnozaabdurahimova396@gmail.com
doi:https://doi.org/10.5281/zenodo.22911154

Abstract. This paper examines how instruction-bearing expressions function when they occur inside material submitted to a generative AI system for translation, editing or linguistic analysis. Drawing on selected examples from an English–Uzbek prompt dissertation, the analysis distinguishes grammatical form from the authority to direct the current task. Particular attention is given to the interpretation of quoted operations, participant roles and output constraints. The discussion argues that translation must preserve both the meaning of an embedded instruction and its relationship to the enclosing request. A framework based on discourse function, addressee assignment and constraint scope is proposed for subsequent bilingual evaluation. The findings are qualitative and do not constitute evidence of measured model performance.
Keywords: generative artificial intelligence, English–Uzbek prompts, quotation, discourse boundaries, directive meaning, constraint scope, translation.

An instruction does not acquire operational authority merely because it contains an imperative verb. In scholarly work on generative artificial intelligence, expressions such as translate, rewrite and tahlil qiling may occur as examples, quotations or objects of grammatical analysis. Their interpretation therefore requires attention to the activity established by the enclosing request. A system asked to explain a prompt must describe the instruction it contains; a system asked to follow that prompt is assigned a different task.

Speech-act theory provides a basis for distinguishing linguistic expressions according to their communicative function. John R. Searle’s classification identifies directives as acts oriented towards getting an addressee to perform an action. In the present context, however, the analyst must distinguish the directive represented inside an example from the communicative act performed by reproducing that example. Correctly identifying directive meaning does not resolve which interaction the instruction belongs to. [1]

The distinction can be examined through a sentence reproduced in the supplied dissertation manuscript: “Translate the following text into academic Uzbek.” Within its envisaged use context, the sentence specifies a translation operation, an object and a target language with a register expectation. Within the dissertation, it illustrates the organisation of a prompt. Its reproduction does not instruct the dissertation’s reader to translate the paragraph that happens to follow it.

Consider the analytical formulation “Translate this sentence into Uzbek: ‘Translate the following text into academic Uzbek.’” An Uzbek equivalent of the enclosing request is “Shu gapni o‘zbekchaga tarjima qiling: ‘Translate the following text into academic Uzbek.’” These

formulations are constructed to expose the distinction between two discourse levels. The outer request establishes the translation task, while the quoted sentence supplies its object. Rendering the inner sentence as “Quyidagi matnni ilmiy uslubda o‘zbek tiliga tarjima qiling” preserves its directive meaning without authorising another translation operation.

The expression the following text, corresponding to quyidagi matn, introduces a further interpretive problem. It refers to material positioned relative to the represented instruction. When the original accompanying text is absent, the analyst cannot assume that the expression refers to the next paragraph of the enclosing document. Such an assumption would supply a referent that the documented example does not establish. Translation must preserve this referential limitation rather than silently resolve it through textual proximity.

Quotation also affects participant assignment. Another manuscript example reads: “Assume that you are an academic editor. Rewrite the article in a formal scientific style.” Here, you identifies the addressee within the represented task. When the passage is submitted for linguistic explanation, that participant role does not automatically transfer to the system providing the explanation. Moreover, assume that presents a role to be adopted for an activity; it does not establish an actual professional qualification.

François Recanati’s treatment of quotation supports examining represented wording in relation to its surrounding use. Applied to prompt analysis, this perspective directs attention beyond punctuation to the framing expressions that establish whether the quoted passage is being demonstrated, translated or adopted as a task. Quotation marks alone cannot resolve the distinction: “Explain this instruction” and “Follow this instruction” establish different relations to the same quoted wording. [2]

Output constraints require comparable attention. If a quoted prompt contains “Limit the text to 1000 words,” the requirement normally governs the output introduced within that prompt. It does not automatically impose a thousand-word limit on an enclosing explanation. Likewise, a quoted requirement to write in English need not determine the language of an Uzbek discussion of that requirement. The analytical question is not simply whether a constraint exists, but which output it governs.

The computational significance of these distinctions is evident in research on indirect prompt injection. Kai Greshake and colleagues investigated how instructions introduced through external content could redirect LLM-integrated applications. The scholarly examples discussed here are legitimate research material, and their imperative form is not evidence of adversarial intent. Nevertheless, executing them instead of processing them under the user’s request would involve a failure to maintain the distinction between supplied content and task-governing instructions. [3]

For English–Uzbek evaluation, task completion and preservation of embedded content should therefore be assessed separately. A response may avoid executing a quoted command but incorrectly omit it from a translation. Conversely, it may reproduce the command accurately while failing to provide the requested explanation. These outcomes concern different properties of the response and should receive distinct descriptions.

The analysis supports three practical questions: what operation does the enclosing request authorise, who is addressed inside the represented instruction, and which output does each constraint govern? These questions offer an inspectable basis for bilingual annotation. Their

reliability requires independent assessment using matched English and Uzbek materials. Within the limits of the present qualitative analysis, the central conclusion is that faithful prompt processing requires preserving the relation between an instruction-bearing passage and the task under which that passage is supplied.

A practical request might therefore read: “Edit the paragraph, but keep the quoted examples unchanged.” Its Uzbek counterpart is “Abzasni tahrir qiling, lekin iqtibosdagi misollarni o‘zgartirmang.” These are analytical formulations proposed here rather than observed corpus examples. Their value lies in identifying the permitted operation and its boundary. The paragraph is available for editing, while the quoted examples retain their documentary form. Where a quotation contains a suspected error, the editor can identify it separately rather than silently replace the evidence.

This situation also reveals a distinction between the original speaker’s responsibility and the present writer’s responsibility. A researcher who quotes an instruction is responsible for reproducing and interpreting it accurately, but does not necessarily endorse its requirements. Similarly, translating a role assignment does not confirm that the assigned role is factually held. Preserving attribution allows a reader to distinguish what the source says from what the researcher claims about that source. The relevance of quotation thus extends beyond typographical separation to the organisation of represented voices within a text. [2]

Negation provides a useful test of these boundaries. Consider the instruction “Do not use bullet points” within a quoted article-writing prompt. When the current task is translation, the negative meaning must remain in the translated sentence. When the current task is explanation, the response should identify a prohibition governing the represented article. In neither case does the negative form alone establish a prohibition governing the present response. The analyst must determine both the content of the restriction and the activity to which it applies.

An Uzbek rendering such as “Bandlar ko‘rinishidagi ro‘yxatdan foydalanmang” preserves the prohibitive force of the sentence. However, placing that rendering outside the quotation or omitting the phrase introducing it as an example could change how a reader interprets its status. The translation would retain the negative imperative while weakening the boundary between represented instruction and current instruction. This illustrates why grammatical accuracy and discourse accuracy should receive separate attention.

Ambiguous framing requires particular care. If a user supplies a quotation without indicating whether it should be translated, explained or followed, the quotation may not contain enough information to identify the intended operation. Previous conversational turns may resolve the uncertainty. Where they do not, a brief clarification can identify the missing relation: “Should I translate this instruction or carry it out?” The corresponding Uzbek question is “Bu ko‘rsatmani tarjima qilaymi yoki undagi topshiriqni bajaraymi?” The purpose of clarification is to establish the operation, not to demand that the user reformulate the entire request.

This does not imply that every quoted passage requires a clarification question. An explicit enclosing request can already supply the necessary relation. Asking again after “Translate the following instruction into Uzbek” would add little interpretive value. The relevant distinction is between a short request whose meaning is recoverable from context and a request that leaves materially different actions unresolved. A useful evaluation should recognise both unnecessary

clarification and unsupported task selection as possible departures from the user's communicative purpose.

Evaluation could then distinguish several observable properties of an answer. One concerns whether the requested operation was performed. Another concerns whether the embedded instruction's meaning was preserved where relevant. A third concerns whether roles, references and constraints remained attached to their appropriate discourse levels. These properties may diverge: an answer can perform the correct operation while mistranslating a restriction, or translate individual sentences accurately while treating their contents as additional instructions.

The connection with prompt-injection research should remain carefully delimited. Greshake and colleagues examined adversarial instructions introduced through external material. The present discussion concerns how legitimate quotations can be mishandled even without an adversarial purpose. The shared issue is the possibility that supplied content acquires an inappropriate task-governing role. This connection supports a comparison of boundary failures, but it does not justify classifying ordinary academic quotations as attacks. [3]

The theoretical implication is that directive interpretation requires more than recognising an action-oriented expression. The analyst must identify the communicative situation in which that expression operates and the larger activity within which it is represented. Speech-act classification contributes to the first task; analysis of quotation and discourse framing contributes to the second. Bringing these perspectives together makes it possible to explain how a sentence retains directive meaning while serving as evidence, translation material or an object of scholarly commentary. [1] [2]

References:

1. Searle, J. R. (1976). A classification of illocutionary acts. Language in Society, 5(1), 1–23. https://doi.org/10.1017/S0047404500006837
2. Recanati, F. (2001). Open quotation. Mind, 110(439), 637–687. https://doi.org/10.1093/mind/110.439.637
3. Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., & Fritz, M. (2023). Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection. arXiv:2302.12173. https://doi.org/10.48550/arXiv.2302.12173