When Is an Answer Enough? Q-Alternative Construction and Probing Behavior Across Korean, English, and Thai

Author: Axel Altmann (Seoul National University) – axelaltmann87@snu.ac.kr

1. Background and research gap

Roberts's (2012, 2018) Question Under Discussion (QUD) framework models discourse as the progressive resolution of questions through alternative elimination. A complete answer eliminates all but one member of the q-alternative set (q-alt); a partial answer eliminates a proper subset, leaving the question unresolved on the QUD stack. A fundamental gap in the framework, however, concerns q-alt construction: the model specifies what alternatives do once they are in play but does not specify how many alternatives are introduced or at what level of granularity. This paper argues that q-alt construction is governed by conventionalized standards of precision in the common ground (Lewis 1979), and that these standards vary systematically across speech communities. The same underspecified response will therefore eliminate a higher proportion of alternatives in a coarser-grained set than in a finer-grained one, yielding cross-linguistically divergent completeness judgments for identical propositional content.

A second gap concerns what happens when a question is not answered but rendered pragmatically inaccessible through motivated non-disclosure. The standard QUD stack mechanics have no mechanism for this outcome. We propose stack suspension as a theoretically distinct discourse result in which the unresolved question is neither eliminated nor retained as an active pursuit target, but becomes interactionally unavailable due to privacy-driven suppression. Stack suspension extends QUD mechanics into the domain of motivated withholding and provides a formal home for privacy performance within discourse theory.

2. Experimental study

An experimental study tested whether cross-linguistic differences in standards of precision produce observable differences in completeness judgments and whether such differences predict probing behavior. Native speakers of Korean (N = 77), English (N = 32), and Thai (N = 26) completed six discourse scenarios in a mixed design (language between-subjects, scenario within-subjects). Four primary scenarios (Group A) contained underspecified responses (e.g., *A: "Why couldn't we meet?" B: "I had something to do."*), while two control scenarios (Group B) contained referentially complete responses exhausting the q-alt set at any level of granularity. Participants rated answerhood completeness and probing likelihood on seven-point Likert scales and provided free-text probe content responses. Because the data were ordinal and the language groups differed in size, all analyses employed non-parametric statistical tests (Wilcoxon signed-rank, Kruskal-Wallis, and Mann-Whitney U).

3. Results

Three principal findings emerged. First, English speakers rated underspecified responses as significantly less complete than Korean and Thai speakers (Kruskal-Wallis $H = 6.980$, $p = .031$; English $M = 3.42$ vs. Korean $M = 4.08$ and Thai $M = 4.12$), while all three groups converged on near-ceiling ratings for referentially complete control responses. This convergence rules out a general scale-use bias and is consistent with the interpretation that English speakers operate with a finer-grained q-alt set as a conventionalized baseline.

Second, completeness ratings did not predict probing likelihood in any language group (Spearman correlations non-significant across all groups and scenarios, rs ranging from −.313 to .330). Perceived q-alt incompleteness and probing behavior are orthogonal constructs: detecting unresolved alternatives does not automatically generate a probing response.

Third, the three languages exhibited structurally distinct probing profiles. English speakers

showed significantly higher probing after complete than after underspecified responses ($W = 94.0$, $p = .004$), a pattern incompatible with straightforward QUD-incompleteness-driven probing. Korean speakers probed at equivalent rates regardless of completeness, indicating that probing in Korean is sustained by interactional pressures independent of q-alt resolution. Thai speakers showed the lowest probing likelihood after underspecified responses overall (Thai vs. English: $U = 281.0$, $p = .012$, $r = -.385$). Qualitative probe content analysis revealed that Korean probes were almost exclusively slot-targeting, while English probes competed between slot-targeting and welfare-inference-driven questions not predicted by the QUD structure.

4. Theoretical implications

The dissociation between completeness and probing indicates that QUD-incompleteness is a necessary but not sufficient condition for probing. Competing pressures including progressivity, social-normative constraints, and privacy-motivated suppression shape whether an unresolved q-alt actually generates a follow-up. Stack suspension formalizes the most extreme case: the unresolved question does not merely go unanswered but becomes pragmatically inaccessible, removed from active pursuit without resolution. This outcome requires an extension of standard QUD stack mechanics and situates privacy performance as a theoretically tractable object within formal discourse theory. The cross-linguistic results further suggest that q-alt granularity is not a universal constant but a culturally variable parameter of the common ground, with consequences for both completeness norms and interactional follow-up behavior.

References

Lewis, D. 1979. *Scorekeeping in a language game*. *Journal of Philosophical Logic* 8. 339–359.

Roberts, C. 2012. *Information structure in discourse*. *Semantics and Pragmatics* 5. 1–69.

Roberts, C. 2018. *Speech acts in discourse context*. In R. Fogal et al. (eds.), *New Work on Speech Acts*. Oxford: OUP.

Simons, M., J. Tonhauser, D. Beaver & C. Roberts. 2010. *What projects and why*. *SALT* 20. 309–327.