

Spatial Language Processing Across Odia, Telugu, and Khasi: Congruency, Axiality, and Distance

Pratik Nanda, Bidisha Som, Cota Navin Gupta
Indian Institute of Technology Guwahati

This study investigates whether spatial language processing differs for categorical relations (e.g., above, on) versus coordinate relations (e.g., near, far), and whether this difference scales with how richly a language lexicalizes each schema [3,4]. We evaluate this in Odia (Indo-Aryan), Telugu (Dravidian), and Khasi (Austroasiatic), three genealogically independent languages, so that shared patterns cannot reflect common inheritance and divergence cannot be a family-level trait [1]. A structured free-production task first established each language's basic spatial adpositions. Axis-based spatial relations accounted for the majority of adpositions in all three languages (Khasi: 71.05%; Odia: 79.62%; Telugu: 86.99%), with cross-linguistic variation in contact lexicalization, front-back and distance encoding [1] that sets up the categorical-coordinate comparison.

Ninety native speakers (30 per language, Age:18-35yrs) participated in a picture-sentence matching task (E-Prime 3.0) in which figure and reference objects were positioned within a 9×9 spatial grid, classified as Congruent, Ambiguous, or Incongruent following the region boundaries in Fig. 2 [2,3]. Reaction Time (RT) was analysed using LME models. Congruency reliably affected RTs across all three languages ($F(2, 188.50) = 28.17, p < .001$), with congruent trials fastest and ambiguous positions, the categorical boundary cases most costly. This effect differed significantly across languages (Language × Congruency: $F(4, 14,463.93) = 6.27, p < .001$): Telugu showed the largest congruency advantage (~170 ms), Odia intermediate (~150 ms), and Khasi the smallest (~66 ms; Table 1, Fig 1). This reflects Telugu's comparatively less and Khasi's richer distance lexicon. A task-specific analysis revealed a vertical-horizontal asymmetry (Task × Congruency: $F(6, 155.22) = 5.20, p < .001$): ambiguous corner positions were costliest for vertical relations, clear mismatches costliest for horizontal relations. This asymmetry was identical across all three languages (Language × Task × Congruency: $F(12, 11,529.68) = 0.89, p = .551$), indicating a language-independent perceptual asymmetry. Finally, categorical and coordinate spatial relations showed a significant three-way Language × Relation Type × Congruency interaction ($F(2, 14,001.61) = 8.33, p < .001$): Odia speakers were more sensitive to categorical (+10.77%) than coordinate mismatches (+4.30%), Khasi showed the reverse pattern (+4.36% vs +9.73%, $p = .051$), and Telugu showed high costs for both (+12.10% vs +10.22%, $p = .711$), each language's cost profile reflects its own lexical richness in that schema. Together, these results indicate that spatial language processing is constrained jointly by a shared perceptual architecture (the vertical-horizontal asymmetry) and language-specific lexicalization (the scaling of congruency costs).

Figures and Tables

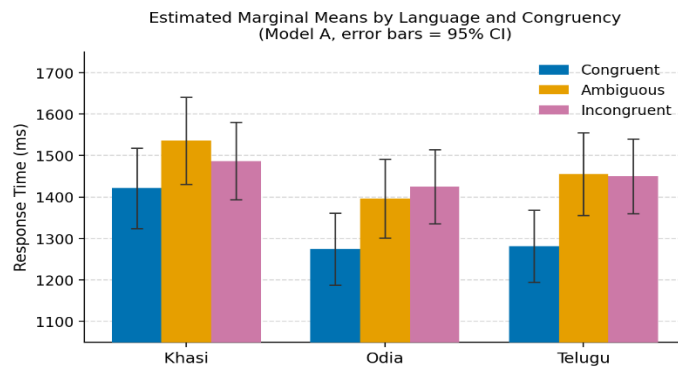


Figure 1. Estimated Marginal Means (EMMs) of response time (ms) by Language and Congruency condition (Model A, linear mixed-effects model). [Bar chart: three grouped bars per language (Khasi, Odia, Telugu), showing Congruent < Incongruent < Ambiguous for Odia/Telugu; and Congruent < Incongruent < Ambiguous for Khasi]

Lan- guage	Congruent (ms)	Ambiguous (ms)	Incongruent (ms)	C–Amb (ms)	p	C–Inc (ms)	p	Amb– Inc	p
Khasi	1,421	1,536	1,487	+115	.001	+66	.015	n.s.	.060
Odia	1,274	1,396	1,425	+122	< .001	+151	< .001	n.s.	.199
Telugu	1,281	1,455	1,450	+174	< .001	+169	< .001	n.s.	.817

Combined model: Congruency $F(2, 188.50) = 28.17, p < .001$ | Language $\times$ Congruency $F(4, 14,463.93) = 6.27, p < .001$

Table 1. Model-adjusted mean response times (ms) and congruency effect sizes per language (Model A). All pairwise contrasts Holm-corrected.

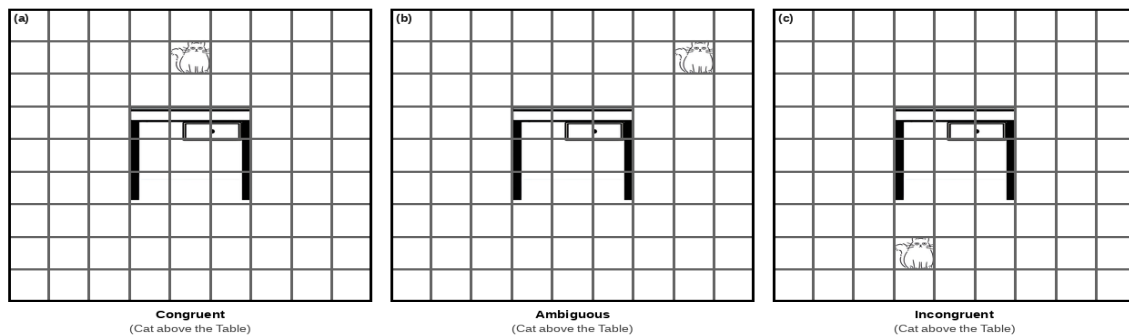


Figure 2. Adapted from Munnich et al. (2001). Stimulus design for the spatial language production task and matching task. A 9×9 grid divides the display into 81 equal cells. The Reference Object (RO; table) is centered in the middle 3×3 region. The Figure Object (FO; cat) appears in one of 40 surrounding cells, classified as: Congruent (6 on-axis cells, e.g., directly above); Ambiguous (8 diagonal/corner cells consistent with two axes); or Incongruent (26 remaining cells). All three panels share the probe "The cat is above the table"; similar congruency structures apply to all spatial axes.

References:

- [1] Levinson, S. C. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press, 2003.
- [2] Munnich, E., Landau, B., & Doshier, B. Spatial language and spatial representation: A cross-linguistic comparison. *Cognition*, 81, 2001.
- [3] Kranjec, A., Lupyan, G., & Chatterjee, A. Categorical biases in perceiving spatial relations. *PLOS ONE*, 9(5), e98604, 2014.
- [4] Soroli, E. How language influences spatial thinking, categorization of motion events, and gaze behavior: A cross-linguistic comparison. *Language and Cognition*, 2024

Optional Supplemental Information

These three languages considered for the study belongs to different language families. While Telugu and Odia are use postpositions following verb-final word order (SOV), Khasi uses prepositions. There is substantial difference in spatial lexicalization patterns and adpositional system.

Khasi

This language belongs to the Austroasiatic (Mon-Khmer branch) family of languages. It is spoken primarily in Meghalaya, northeast India. Typological profile: SOV word order; prepositional; morphologically analytic. Khasi uses a pronominal prefix system and has employs a grammatical gender agreement system. Notably, Khasi uses distinct lexical items for vertical contact (halor 'on-surface') vs. projective above (hajrong).

ka miaw ka don hajrong ka miej

FEM cat FEM be.PRS above FEM table

'The cat is above the table.'

Odia

Odia is an Indo-Aryan language, like most Indo-Aryan languages (Hindi, Bengali, Marathi), Odia is verb final. It uses postpositions. While it has free NP ordering, verb canonically appears at the end of a clause. Remarkably, Odia uses same postposition *upare* for both above and on. The contact feature is encoded by adding the particle *lagiki* (contact/attachment marker)

bilei meza upare acchi

cat table above be.PRS.3P.SG

'The cat is above the table.'

Telugu

Telugu is a Dravidian language and official language of Andhra Pradesh and Telangana. The language follows SOV word order and uses postpositions. It has agglutinative morphology with extensive suffix usage. The language encodes above as 'paina' and top contact as 'meeda.'

pilli balla paina undi

cat table above be.PRS.3P.SG

'The cat is above the table.'