

Lexical Conceptual Structure and Marathi

Christopher Manning

Jackendoff (1987, 1990) has brought up various problems with the current use of thematic roles (Kiparsky, 1987; Bresnan & Kanerva, 1989 and references cited therein) and suggested a different way of thinking of thematic roles as structural configurations in his semantic Lexical Conceptual Structures (LCSs). Conversely, Joshi (1989) has claimed that Jackendoff's LCSs alone are insufficient, and that an analysis of certain facts in Marathi additionally requires the existence of a level of predicate-argument structure (PAS). Below we will mention a few of Jackendoff's arguments against the current conception of thematic roles. We will then look at Joshi's arguments about the necessity of a level of PAS in addition to LCS and conclude that providing Jackendoff's LCSs are integrated into a suitable syntactic theory, neither of her points are problematic to Jackendoff.¹ From there we will go on to re-examine some of the facts of Marathi, and show that certain facts that have merely been stipulated or left unanalyzed when using thematic roles, receive a rather elegant treatment when described via a combination of their syntax and LCS.

To keep this paper to a squibbish length, we're going to have to take a few things on trust. Firstly, we'll assume that Joshi's (1989, 1990) syntactic analysis of Marathi is essentially correct.² Secondly we're going to assume a grammatical formalism that allows one to wed together LFG-ish syntax with Jackendovian semantics (I do actually have such a thing: Andrews & Manning (forthcoming)). Lastly, I'm not going to detail here the Definitive Linking Theory, but we'll see just enough of it to make the arguments look plausible.

Problems with thematic roles. Jackendoff raises several problems with conventional thematic roles. He points out that many of the roles subcategorized by verbs have no categorization in the standard thematic role lists. His usual argument is to provide unnamed locational arguments such as:

- (1) a. John jumped *the gorge*.
- b. The plane circled *the tower*.

This is a problem for Kiparsky who tries to include various locational roles:

[Source < Goal < Means < Theme < From < To < Via < Verb > > > > > >]

It is less of a problem for Bresnan *et al.* who use just one locational macro-role:

Ag > Ben > Recip(Go)/Exp > Inst > Th/Pt > Loc

However it is not hard to think of other non-locational arguments for which there is no good categorization:

¹The situation is somewhat clearer with the publication of Jackendoff (1990) than at the time that Joshi (1989) was written, though the crucial innovation of the Action Tier appears in Jackendoff (1987). Note in particular that in Chapter 11, Jackendoff (1990) recreates within his LCS something similar to, but interestingly different from, currently touted thematic hierarchies.

²Due to dialect variation among Marathi speaking linguists, it can be rather hard to tell when someone is giving a theory that better explains the same facts as against a theory that is explaining different grammaticality judgements.

(2) John underwent *an operation*.

We all know the joke: ‘Who knows what that NP is—let’s call it a Theme.’ Also, thematic roles should not be merely a theory of NPs: *red* is a Source and *green* a Goal in:

(3) The light changed from *red* to *green*.

An NP may hold multiple θ -roles in a sentence:

(4) a. John (Agent, Theme) is chasing Bill (Theme, Goal).

b. John (Agent, Source, Recipient) paid him (Source, Recipient) \$5 for the book.

Whereas thematic roles were meant to be grounded in clear semantic intuitions, the chosen role in such cases is usually just the one that ‘works’ in the syntax, and thematic roles are thus reduced to a theory-internal artifact.

But even if we could find a satisfactory set of thematic roles, there is a question of the need for them. While some would wish no association with LCSs, for those such as Bature (1991) who use both LCSs and PASs, there is a clear question of parsimony. Given that LCS is necessary anyway as a semantic representation of sentences, aesthetic sense (perhaps as codified in Occam’s Razor) would suggest that we should not additionally postulate a level of PAS unless it is really necessary.

Joshi’s arguments for Predicate Argument Structure. Conversely, Joshi raises two problems for a theory that uses only LCSs. Firstly she mentions the classic problem of differentiating the roles between pairs of verbs such as *buy/sell*, or *like/please* for that matter. However, while these pairs of verbs do have identical roles on the Thematic Tier, they can now be easily differentiated because of their different classification of arguments on the Action Tier. We would have lexical entries such as the following for *buy* and *sell* respectively:³

(5) a.
$$\left[\begin{array}{l} \text{AFF}([\]^{\alpha_i}, \) \\ \text{GO}_{\text{Poss}}([\]_j, \left[\begin{array}{l} \text{FROM} [\]^{\beta_k} \\ \text{TO} [\ \alpha] \end{array} \right]) \\ \left[\text{EXCH} \left[\text{GO}_{\text{Poss}}([\ \text{MONEY}]_m, \left[\begin{array}{l} \text{FROM} [\ \alpha] \\ \text{TO} [\ \beta] \end{array} \right]) \right] \right] \end{array} \right]$$

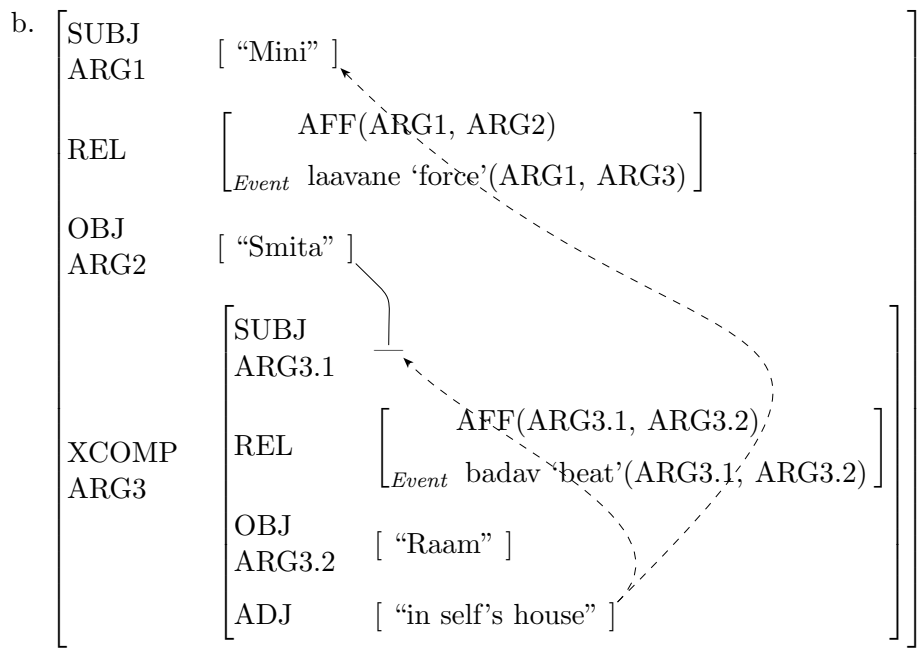
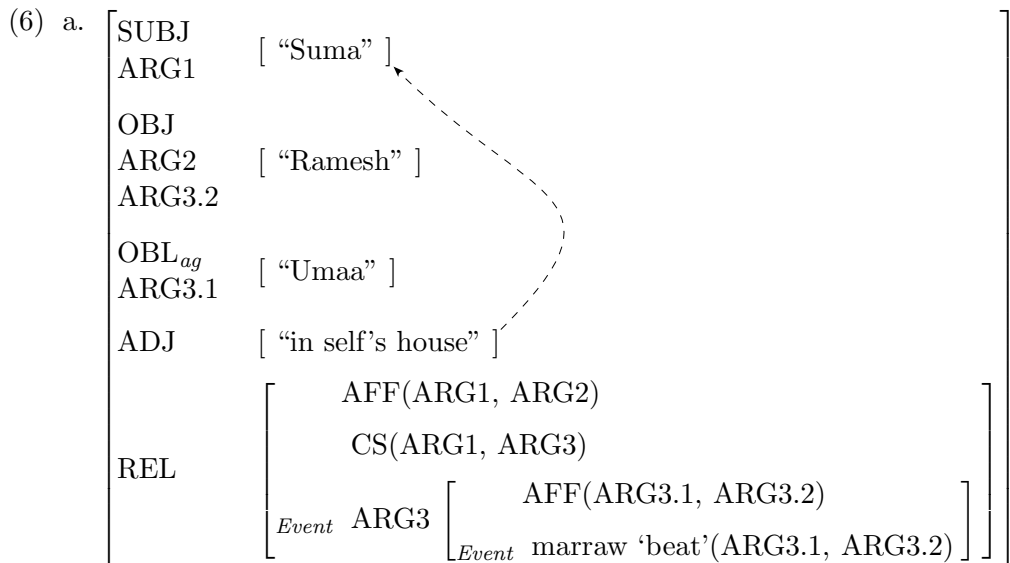
b.
$$\left[\begin{array}{l} \text{AFF}([\]^{\alpha_i}, \) \\ \text{GO}_{\text{Poss}}([\]_j, \left[\begin{array}{l} \text{FROM} [\ \alpha] \\ \text{TO} [\]^{\beta_k} \end{array} \right]) \\ \left[\text{EXCH} \left[\text{GO}_{\text{Poss}}([\ \text{MONEY}]_m, \left[\begin{array}{l} \text{FROM} [\ \beta] \\ \text{TO} [\ \alpha] \end{array} \right]) \right] \right] \end{array} \right]$$

The motivation for these structures and the further predictions are similar to those of Joshi’s (1989) own analysis using PAS.

More seriously, Joshi raises the question of the binding of *aapan*. Joshi shows that a binder of *aapan* in the same clause must be the *logical subject* of the clause. In particular, in a lexical causative

³These differ from those that Jackendoff (1990) proposes: he inserts an extra CS function, but this seems to go against both his own discussion on p. 128–129 and the behavior of these verbs when modified by an adverb like *almost*. At any rate, either way they are differentiable.

aapaṇ cannot be bound by the causee. She then contrasts lexical and periphrastic causatives (where the causee can be the antecedent) and points out that both are biclausal at LCS (by assumption, and as shown by tests using adverbs like *almost*). Hence she suggests that there is no way to characterize the binder of *aapaṇ* at LCS in a way that would distinguish these two cases. This might be true if one is looking just at an LCS, but if we have pieces of LCS integrated with a syntactic formalism, this is not true. Suppose for simplicity that *aapaṇ* can be bound by a logical subject in its own clause or any clause it is embedded in.⁴ We will illustrate the general effect by exhibiting a partial description of the grammatical functions and semantic arguments of a lexical and a periphrastic causative, respectively:



⁴This is obviously a gross oversimplification, as not only can an *aapaṇ* that is a direct argument not be bound inside its own clause, but *aapaṇ* has various logophoric properties as well. However, these complications don’t affect the force of Joshi’s argument.

Then the idea is this: we will specify most of the anaphoric binding path of *aapaṇ* in the functional projection (as in Dalrymple, 1990), but for the final step, we will switch to the semantic projection and select the logical subject (we will copy the notation $\hat{\theta}$, but will mean by this the highest argument in the LCS thematic hierarchy that Jackendoff (1990) develops in Chapter 11). We could then write this restriction as follows:

$$\uparrow_{\sigma} =_c ((\{XCOMP \mid COMP\}^* \uparrow)_{\sigma} \hat{\theta})$$

Moreover, if we return to the real *aapaṇ*, it appears that the biclausal LCS will be part of the key to describing its binding possibilities. Note the following contrast:⁵

- (7) a. *Mini-ne aaplyaa-laa baḍavle
 Mini-ERG self-ACC hit
 *‘Mini_i hit herself_i.’
- b. Mini-ne Shaam-kadun aplyaa-laa have-t ucalavle
 Mini-ERG Shaam-OBL_{ag} self-ACC air-IN lift-CAUSE-PAST
 ‘Mini_i made Shaam lift self_i in the air.’

In both sentences *aapaṇ* is a Patient and an OBJ, and in both sentences the desired antecedent is an Agent and a SUBJ. As seen in (7a), when *aapaṇ* is a direct argument of a verb, usually it cannot be bound by the SUBJ of that same verb. However, with the lexical causative of a transitive verb, it can be so bound. Intuitively, the difference is that the SUBJ is further away this time, because of the biclausal LCS. We tentatively suggest the following constraint: If *aapaṇ* is an argument (as opposed to an adjunct) it cannot be bound by the potential logical subject immediately superior to it, where by a potential logical subject we mean the highest role on the LCS hierarchy at one level of LCS (it may not be a real logical subject because one LCS might be embedded in another). This correctly predicts that in transitives and the causatives of intransitive verbs that an *aapaṇ* in the OBJ slot cannot refer to the SUBJ or Causer (as it is the immediately superior potential logical subject), but that in the casuative of a transitive verb, *aapaṇ* can refer anaphorically to the Causer, because, even if not overtly expressed, the Causee is an intervening potential logical subject.

Capturing semantic intuitions with LCSs. Pandharipande (1990) does not use complex abstract syntactic analyses and so it might seem the work of a “flat-earth functionalist” (Pullum, 1989), but lately this mindset has been receiving greater respect. Mohanan & Mohanan (1990) admit that K. P. Mohanan “conclude[d] prematurely that the regularities between the dative case and its meaning are at best a tendency, and that the distribution of the dative case... must be stipulated on individual morphemes.” Scanning the references of Bresnan & Kanerva (1989) is also instructive. So let us take seriously Pandharipande’s main intuition, that the dative *-laa* is a marker of location or possession and that dative marked subjects should also be regarded as locationals. In particular, she considers the following paradigm:

- (8) a. nadī-lā pūr yeto
 river-DAT flood comes
 ‘The river gets flooded (The flood comes to the river).’
- b. mī gāvā-lā zāto
 I village-DAT go
 ‘I go to the village.’

⁵ *aaplyaa* is a morphophonemic variant of *aapaṇ*. Also note that, unfortunately, the transcription of examples in this paper is not consistent, but I feared worse results if I had attempted to rectify this.

- c. ma-lā rāg yeto
I-DAT anger comes
'I get angry (Anger comes to me).'
- d. mī tyā-lā p̄ese deto
I he-DAT money give
'I give him money.'
- e. ma-lā ek bhāū āhe
I-DAT one brother is
'I have a brother.'
- f. ma-lā rāg āhe
I-DAT anger is
'I am angry (I have anger).'
- g. zāḍā-lā phula ālet
tree-DAT flowers are
'The tree has flowers.'
- h. phulā-lā manda vās āhe
flower-DAT mild fragrance is
'The flower has a mild fragrance.'

She suggests that all these dative NPs are a unified category covering location and possession. Under Bresnan & Kanerva's thematic role hierarchy, however, they would be distributed between at least the Locative, Experiencer and Recipient thematic roles (and it's not exactly clear what role possessors have on this hierarchy: Joshi (1989) seems to tentatively suggest that Possessor is yet another role that ranks between Agent and Theme). Pandharipande's intuition has been entirely lost. There is no explanation of the case marking, semantic relationship or similar syntactic behavior. But consider the thematic tier representation of these verbs that can be given in Jackendoff's LCSs:

- (9) a. [*Event* GO_{spatial}(flood, [*Path* TO([*Place* River]))]]
 b. [*Event* GO_{spatial}(I, [*Path* TO([*Place* village]))]]
 c. [*Event* GO_{ident}(anger, [*Path* TO([*Place* I]))]]
 d. [*Event* GO_{poss}(money, [*Path* TO([*Place* he]))]]
 FROM([*Place* I])
 e. [*Event* BE_{poss}(brother, [*Path* TO([*Place* I]))]]
 f. [*Event* BE_{ident}(anger, [*Path* TO([*Place* I]))]]
 g. [*Event* BE_{poss}(flowers, [*Path* TO([*Place* tree]))]]
 h. [*Event* BE_{poss}(fragrance, [*Path* TO([*Place* flower]))]]

The dative-marked arguments have regained their conceptual unity. They all appear as the Place in the second argument of the conceptual functions BE and GO:

- (10) [*Path* TO([*Place* —])]

As Jackendoff (1983, 1990) has argued, postulating the conceptual functions BE and GO allows us to explain the pervasive use of notions of motion in both real and metaphorical domains. BE and GO generalize across conceptual domains, though we can index the function with the particular domain being used (GO_{spatial} , GO_{poss} , etc.).

Of course, the whole reason why Experiencer is carefully placed above Theme in most thematic hierarchies is to explain the more “logical-subject-like” properties of an Experiencer. But with the addition of the Action Tier to Jackendoff’s LCSs, properties can be factored between thematic and ‘action’ properties. Although Experiencers are clearly metaphorical locations in Indian languages, they differ from normal locations by being entities that react. This is expressed in Jackendoff’s system by their being the first argument of the REACT function on the Action Tier. So the LCS for the verbs in (8b–c) will look like the following:

- (11) b. $\left[\begin{array}{l} \text{AFF}(\text{---}_A, \text{---}) \\ \text{Event } GO_{\text{spatial}}(\text{---}, [\text{Path } TO([\text{Place } \text{---}])]) \end{array} \right]$
- c. $\left[\begin{array}{l} \text{REACT}(\text{---}_A, \text{---}_A) \\ \text{Event } GO_{\text{ident}}(\text{---}, [\text{Path } TO([\text{Place } \text{---}])]) \end{array} \right]$

Thus in (8b) the *village* is merely a location, while in (8c), the speaker is both a location and an entity that reacts to the arrival of *anger*.

Similarly, we will see the possessor in *I have a brother* as reacting to this event. It gets somewhat more metaphorical with nonsentient beings, but for them also, we will view the difference between:

- (12) a. The tinsel is on the tree.
 b. The tree has flowers.

as the *tree* in (a) being a simple location, but also appearing on the Action Tier in (b). Even inanimate possessors become the logical subject in Marathi, as can be shown with the *-un* test (only logical subjects can control *-un* participial clauses: Joshi, 1990):

- (13) a. gharaalaa don darvaaje aahet
 house-DAT two doors-NOM are
 ‘The house has two doors.’
- b. vaadaḷaat dashaa hounahi gharaalaa adzun don darvaaje aahet.
 ‘In spite of [the house] being ruined in the storm, the house still has two doors.’

This linking of arguments to the first argument of REACT on the action tier forces them to appear higher on Jackendoff’s thematic hierarchy, indeed, as the logical subject.

This recognition of experiencers as locations seems essential to an adequate analysis of Marathi. We have already seen that *-laa* can mark either a vanilla location (8b) or an ‘experiencer subject’ (8c). But, in fact, it is not only *-laa* that has these properties.⁶ For example, consider another locative suffix *-t* ‘in’:

- (14) to śāḷe-t gelaa
 he school-IN went
 ‘He went to school.’

⁶I thank Smita Joshi for the data crucial to the explication of this point, and in general for assistance with other questions.

It, also, can be used with inanimate qualities and human locations:

- (15) tyaačyaa-t taakad aali
 he-IN strength came
 ‘He became strong (He got strength).’

and in these sentences, just like the ‘Dative Subject’ construction, we find that *he*, which is both a Place and the first argument of REACT on the Action Tier, becomes the logical subject, again demonstrated by using the *-un* test:⁷

- (16) jew-un tyaačaa-t taakad aali
 Upon dining he-IN strength came
 ‘Upon dining, he felt strong.’

If we acknowledge that this usage of ‘Experiencer Subjects’ is semantically determined but fail to explicate the connection between ‘Experiencers’ and locations, then we are clearly missing a generalization.

Yet another linking theory. Now that PAS has been abandoned, we need to explain how we will link directly from LCS to grammatical functions (GFs).⁸ Jackendoff (1990, Chapter 11) has one stab at such a theory. However, it does not seem that his approach (of lining up a grammatical and semantic hierarchy and then pairing them off) is completely suitable for handling the sort of Reversal Constructions (Joshi, 1990) which are found in Marathi.⁹ We could just import LFG’s Lexical Mapping Theory wholesale, but it seems a little unnatural. For although we have decided that ‘Experiencers’ are a special type of Locative, we would have to say that while things that are only Locatives are intrinsically [−o], things that are also linked to REACT on the Action Tier don’t have this intrinsic feature, but instead are intrinsically [−r]. For if ‘Experiencers’ were also intrinsically [−o], they would not be able to participate in Reversal Constructions.

These ideas are still a bit up in the air, and space is running out, but let me just briefly outline how another LCS linking theory might work. We’ll inherit the distinction between arguments and adjuncts from Jackendoff and I’m assuming knowledge of the facts in Joshi (1990). Let’s name some LCS structural configurations we will use:

Actor:	First argument of AFF	Theme:	First arg. of BE or GO
Reacter:	First argument of REACT	Locative:	Second arg. of BE or GO
Patient:	Second arg. of AFF [−] or REACT	Beneficiary:	Second argument of AFF ⁺

The LCS thematic hierarchy will be:

⁷I’m not an expert on RG, but these facts would also seem to be rather damaging to Rosen & Wali’s (1989) analysis. The first argument here must start off as an Initial-1 and then strangely demote into a locative non-term.

⁸Marathi would seem quite problematic to Kiparsky’s (1987) Direct Linking Theory because arguments with identical case, position and agreement possibilities can bear different GFs. Whatever faults their article may have, Rosen & Wali (1989) make this point quite strongly.

Also, Bature (1991) suggests that “A direct link between LC structure and grammatical functions predicts that all lexical meanings of predicates are accessible to syntax.” But this is not necessarily the case. The marking of argument positions (subscripted with an ‘A’ in Chapter 11 of Jackendoff (1990)), limits what information the linking theory is allowed to make use of. Besides, the logic is faulty. For next we could argue that because a direct link between PAS and LCS predicts that all lexical meanings of predicates are capable of altering the ordering of the thematic hierarchy, we should postulate another level between LCS and PAS. The infinite regress that follows should be obvious to the reader.

⁹These are constructions where for two semantic arguments, either can appear as the surface SUBJ and the other then appears as the surface OBJ.

Ag > Reacter > Pt/Th/Ben > Loc

Intransitive verbs make their one argument the SUBJ. Normal transitive verbs will have an Agent and a Patient/Theme and link to a SUBJ and an OBJ in the obvious way. For ‘Experiencer Subject’ verbs, although the Reacter is the logical subject, let’s presume the Reacter and the Patient are sufficiently ‘near’ in ranking that either can be the SUBJ and the other is the OBJ. For ditransitives, the Agent becomes the SUBJ and the other two arguments will be a Theme and a Beneficiary¹⁰ or similar and they both become OBJs (one or other of which might be ‘restricted’). Transitive verbs passivize by making the Actor unavailable to argument linking (linkable only as an adjunct). The Patient, being the only grammatically linked argument will become the SUBJ. ‘Experiencer subject’ verbs don’t passivize, lacking an Actor, or causativize. When ditransitive verbs passivize, the Actor is again restricted, and the remaining two arguments (Theme and Beneficiary) can again engage in a Reversal Construction. In Marathi one can only have causatives of verbs with Patients, because in Marathi, the Patient of AFF on the Action Tier of the causative morpheme must unify with a Patient-like argument in the lower verb. Unergative verbs do not have morphological causatives. In the main class of (monoclausal at f-structure, morphological) causative verbs, the (outer) Actor will link to the SUBJ and the (only) Patient will link to the OBJ. If the base verb is transitive, the Causee (inner Actor) will be left unlinked, but it can be introduced via an Adjunct rule that introduces unlinked Actors, which in Marathi involves the suffix *-kadun*. The small class of ‘ingestive verbs’ (Joshi, 1990) form a Causative with two OBJs. The semantics of these verbs seems to suggest that on the Action Tier these verbs are specified for both AFF and REACT. Here, normally the Patient of the cause morpheme will unify with the Reacter (with most ingestive verbs this is mandatory, but a very small number allow the Patient of the cause morpheme to unify with either the lower Reacter or the lower Patient). The active form will thus have two Patients, and both become OBJs. When such a verb is passivized, the Actor is restricted from grammatical linking and we will again have a Reversal Construction: either of these Patients can become the SUBJ and the other becomes the OBJ.

A Final Word. It could be suggested that the reason that PAS is no longer needed is that Jackendoff (1990) has incorporated all the desirable aspects of PAS hierarchies into his LCS, and there is some truth in that. Also, Bresnan & Kanerva suggest that the labels on their thematic hierarchy are taken to be abstractions over the roles of a more fine-grained semantic analysis, and we have seen (above and Jackendoff, Chapter 11), that we can indeed define such abstractions over LCS structural positions. So there is maybe not that much to argue about. But the use of LCS does give one a more expressive and less problematic vocabulary in which to state facts that vary in their scope and degree of generality. With thematic roles, one is restricted to the single level of generality decided by the person that wrote the thematic role hierarchy. More fine-grained semantic distinctions are uncapturable, while less fine-grained generalizations can be described only via stipulated ‘hyper-roles’.

¹⁰The meaning of Jackendoff’s Beneficiary is really much closer to LMT’s Recipient (sometimes also referred to as Goal).

Bibliography

- Andrews, A. D. & Manning, C. D. (forthcoming). Untitled manuscript, Stanford University & The Australian National University.
- Bature, Abdullahi. (1991). Thematic Arguments and Implicit Semantic Roles. Draft Dissertation Proposal, Stanford University.
- Bresnan, J. & Kanerva, J. (1989). Locative Inversion in Chicheŵa: a Case Study of Factorization in Grammar. *Linguistic Inquiry*, 20, 1–50.
- Dalrymple, M. E. (1990). *Syntactic Constraints on Anaphoric Binding*. Unpublished PhD Thesis, Stanford University.
- Jackendoff, R. S. (1983). *Semantics and Cognition*. Cambridge, MA: The MIT Press.
- Jackendoff, R. S. (1987). The status of thematic relations in linguistic theory. *Linguistic Inquiry*, 17, 369–411.
- Jackendoff, R. S. (1990). *Semantic Structures*. Cambridge, MA: The MIT Press.
- Joshi, Smita. (1989). Logical Subject in Marathi Grammar and the Predicate Argument Structure. *Proceedings of the Eighth West Coast Conference on Formal Linguistics*, 207–219.
- Joshi, Smita. (1990). Multistratalism in Marathi. Unpublished manuscript, Stanford University.
- Kiparsky, Paul. (1987). Morphology and grammatical relations. Unpublished manuscript, Stanford University.
- Mohanan, K. P. & Mohanan, Tara. (1990). Dative Subjects in Malayalam: Semantic Information in Syntax. In Verma & Mohanan (1990), 43–57.
- Pandharipande, Rajeshwari. (1990). Experiencer (Dative) NPs in Marathi. In Verma & Mohanan (1990), 161–179.
- Pullum, G. K. (1989) Topic . . . Comment: Formal Linguistics meets the Boojum. *Natural Language & Linguistic Theory*, 7, 137–143.
- Rosen, Carol & Wali, Kashi. (1989). Twin Passives, Inversion and Multistratalism in Marathi. *Natural Language & Linguistic Theory*, 7, 1–50.
- Verma, Manindra & Mohanan, K. P. (Eds). *Experiencer Subjects in South Asian languages*. Stanford, CA: Stanford Linguistics Association.
- Wali, Kashi. (1979). Two Marathi Reflexives and the Causative Structure. *Studies in Language*, 3, 405–438.