1 Introduction

Optimality theory (OT; Prince and Smolensky 1993) has in a short time proven influential in theoretical phonology, and to a lesser degree in theoretical syntax. One of the main attractions of the theory is its strong commitment to typology. The basic hypothesis is that there is a set of universal constraints and languages differ only in how they rank the constraints. Since constraints are violable, this will result in the same input having a different optimal expression in different languages. The fact that constraint reranking is the only difference between languages is a very important assumption in OT and is necessary for its explanatory scope (Smolensky, 1996b; Bresnan, 2000), its learnability proofs (Tesar and Smolensky, 1998), and most importantly for typology. Once we assume “richness of the base” — as this assumption is often referred to — a natural typology arises. If there are \( n \) constraints in a strict ordering then there are \( n! \) different ways of ranking them. Since languages differ only in these rankings, it follows then that we predict \( n! \) languages. This is called “factorial typology”.

Factorial typology is an important, automatic consequence of optimality-theoretic grammars. It is this property that truly sets OT grammars apart from ones in other generative frameworks. However, expressing something easily in a framework often comes at the cost of complicating the formalization of certain other facts. Frameworks in which constraints are inviolable can deal with ambiguity and optionality fairly well. If there are multiple ways of satisfying constraints, then there is optionality (in production) or ambiguity (in comprehension). Similarly, these frameworks have no significant problems with ineffability. This latter term refers to the fact that certain inputs to a grammar have no expression. In frameworks with inviolable constraints, these inputs simply violate one or more constraints and are therefore rejected by the grammar. However, it is precisely the properties of ineffability, ambiguity, and optionality that standard Optimality Theory (Prince and Smolensky, 1993) has trouble expressing.

In this paper, I take steps to accommodate ambiguity and optionality in OT syntax. In particular I examine optionality in linking arguments to grammatical functions in Marathi and give an OT-LFG account of the relevant facts. In doing so, I will illustrate four points. The first is that linking can be achieved with a small set of cross-linguistically plausible, violable constraints. Second, optionality can be captured in OT by modifying the architecture of the theory only slightly. Third, in comprehension directed optimization, the same OT constraints that are used to capture the linking optionality in production can also capture the resulting ambiguity in the Marathi strings that correspond to the

---

*This work was supported in part by SSHRC Doctoral Fellowship 752-98-0424 and was completed as a Research Fellow at UCSC, where I was kindly sponsored by Jim McCloskey. I am also very grateful to the following people for discussion, comments, and suggestions: Farrell Ackerman, Judith Aissen, Arto Anttila, David Beaver, Joan Bresnan, Miriam Butt, Ashwini Deo, Alan Gous, Frank Keller, Paul Kiparsky, Chris Manning, Line Hove Mikkelsen, John Moore, Peter Sells, the participants in the Stanford OT Working Group, and the joint Stanford/UCSC Optimality Theory Syntax class. I would like to particularly thank Ida Toivonen, who has influenced my thoughts on OT and optionality through many long discussions. I accept sole responsibility for any remaining errors.

1It would be more precise to say that a form can be grammatical and still violate constraints.

2OT-LFG is a version of OT syntax which uses Lexical Functional Grammar (Bresnan, 1982; Dalrymple et al., 1995) for its representations and as its theory of \( \text{GEN} \) (see (1) below).
winning candidates in production. Fourth, this OT approach to linking has interesting implications for Dowty’s theory of proto-roles.

In section 2, I briefly present the basic architecture of standard Optimality Theory and how it relates to ineffability, ambiguity, and optionality. Then, in section 3, I discuss four possible ways of capturing optionality in OT. In section 4, I present the facts about optionality in Marathi linking and review Joshi's (1993) Lexical Functional Grammar account of these facts. Then, in section 6, I present an Optimality Theory account that is considerably simpler in some respects, although it also uses LFG and thus benefits from its formal rigour and representational clarity. I go on to illustrate how the OT analysis deals with ambiguity in the comprehension direction. Lastly, I discuss some new insights into proto-role theory that result from this account.

## 2 Ineffability, Ambiguity, and Optionality in OT

The basic architecture of Optimality Theory is as follows (Prince and Smolensky, 1993; Bresnan, 2000):

\[(1) \quad \text{Inputs} \quad \text{The inputs to the grammar are universal. For syntax, the inputs are usually assumed to be semantic forms of some kind (Smolensky, 1998).}
\]

\[\text{GEN} \quad \text{A generator function from an input to a set of candidates which the grammar will evaluate.}\]

\[\text{EVAL} \quad \text{An evaluation function from the set of candidates to a winning output. It consists of:}\]

1. A universal constraint set.
2. A language particular ranking of these constraints.
3. An algorithm for harmonic ordering: the optimal/most harmonic/least marked candidate (= the output for a given input) is one that best satisfies the top ranked constraint on which it differs from its competitors. (Bresnan, 2000; Smolensky, 1996a)

The last point is the most important one: OT grammars will pick the most optimal candidate. As Pesetsky (1997) puts it, an OT grammar only picks winners and it will always pick a winner. However, there are certain cases in which there should be no winners because the input is “ineffable” (i.e. it has no acceptable output). The inability of OT to deal with ineffability led many to initially reject it as a theory of syntax, but steps have been taken to address this problem. Smolensky (1998) presents an OT account of ineffability based on the notion of production/comprehension chains. This notion is based on the idea of dealing with production and comprehension using the same constraint ranking. Production takes a semantic input and gives a syntactic parse as an output, whereas comprehension takes the string corresponding to the syntactic parse (i.e. the “overt” part of the production output) and gives a semantic form as an output. The output of comprehension is the same type of form that is the input for production. We can thus talk about bidirectional optimization using a production/comprehension chain:

\[(2) \quad \begin{align*}
\text{a. } & /I/ \rightarrow \text{production} \rightarrow [S] \rightarrow \text{overt part} \rightarrow "O" \rightarrow \text{comprehension} \rightarrow /I'/ \\
\text{b. } & \text{If } /I'/ = /I/, \text{ then } /I/ \text{ is expressible; If } /I/ \text{ is not expressible it it is ineffable. (Smolensky, 1998)}
\end{align*}\]

---

3A fifth way, which has not been discussed much in the literature (for example, it is not mentioned in Müller’s (1999) rather thorough overview of optionality in OT), is to use underspecification (in the candidates, not the input). I feel that candidate underspecification has severe foundational problems in OT (see Artstein 1998 for a general discussion, as well as Kuhn 2000 for a discussion specific to OT-LFG), but this is not the place to give these issues serious treatment, which they deserve. Therefore, I leave underspecification out of this discussion, as I plan to examine its role in ambiguity and optionality and in OT generally in future work.

4For extensive OT-LFG analyses using bidirectional optimization, see Lee (2001) and Donohue (1999).

5$I$ is the semantic input; $[S]$ is the syntactic parse; “$O$” is the string; $I'$ is the output of comprehension, but is also the same type of object as the input in production.
Once we take this approach into account, although the production grammar still picks a winner, we can formally capture the idea that not all winners are expressible. Thus, bidirectional optimization addresses the charge that OT cannot deal with ineffability and removes a major obstacle in adopting it as a framework for syntax research.

This still does not address the fact that OT picks only one winner. But, there are situations in which more than one candidate should be selected. This occurs when there is either optionality or ambiguity. The latter case is common in semantic comprehension; an example is provided by quantifier scope ambiguity such as in the following sentence, which has the interpretations in (3a) and (3b):

(3) Every student read a book.
   a. For every student x, there is a book y such that x read y. (wide scope universal)
   b. There is a book y such that for every student x, x read y. (narrow scope universal)

An OT grammar that picks only one of (3a) or (3b) has failed to get the facts right.

This problem is not particular to semantic comprehension. It also occurs in phonological comprehension. In some dialects of North American English, the following phonetic form is ambiguous between the two underlying phonological forms in (4a) and (4b).

(4) [tʃɪɾɑːl]
   a. /tʃɪdɑːl/ (≈ “rider”)
   b. /tʃɪtɑːl/ (≈ “writer”)

Again, a properly descriptive grammar should select both the (a) and (b) forms.

In optimality-theoretic syntax, ambiguity can be characterized as a situation in which the constraints should pick more than one semantic candidate as optimal (where the semantic candidates are the inputs to production and the outputs of comprehension; see footnote 5). Since optionality can be characterized as a situation in which an input has more than one optimal output candidate as well, it is just the flip side of ambiguity. That is, ambiguity occurs when there is more than one optimal candidate in the comprehension grammar and optionality occurs when there is more than one optimal candidate in the production grammar. Although the focus of this paper is optionality, the ideas presented here can be carried over to deal with ambiguity. In the next section, I present several ways to handle optionality and ambiguity in OT.

3 Four Approaches to Optionality

Müller (1999) presents a concise but thorough overview of optionality in OT. He classifies proposals to deal with optionality into four types: pseudo-optionality, neutralization, true optionality, and constraint ties. I will opt for a variant of the last approach, but first I will discuss the other three approaches and the reasons for rejecting them.

3.1 Pseudo-optionality

In the pseudo-optionality approach, each option is the winner of a separate optimization competition. Optionality as such does not exist. Müller points out that this predicts that the various options will never be in competition, which is too strong. Optionality often breaks down in certain contexts, but pseudo-optionality cannot get only one alternant in such cases, because the alternants are not in competition by hypothesis. Pseudo-optionality approaches will overgenerate in such cases.

One of Müller’s (1999) examples will make things clearer. Consider English complementizer drop:

---

6There can, in principle, be candidate ties, in which case both candidates will win, but see section 3.3 for criticism of this approach.
Leaving details aside, the pseudo-optionality approach would have each of these candidates be the winner for distinct inputs. But, the complementizer cannot always be dropped (Müller, 1999):

(6) a. It surprised me that the earth is round.
   b. *It surprised me the earth is round.

On this approach, it is impossible to rule out (6b), as the candidate with the complementizer present and the one with the complementizer absent are each the winner of their competition. Therefore, it is not possible for (6a) to block (6b) and the grammar overgenerates.

3.2 Neutralization

The second proposal is neutralization. On this approach, the options also belong to different competitions, but in certain contexts a high-ranking constraint will result in the same surface string being selected in each candidate set. Thus, the difference between the two competitions is neutralized, as the candidate that wins in each has the same string yield. Müller reviews the neutralization approach to English complementizer drop given by Baković and Keer (Baković, 1997; Keer and Baković, 1997; Baković and Keer, to appear). I will consider the version presented in Baković and Keer (to appear), as this is the most recent statement of this approach.

Baković and Keer make three key assumptions in their analysis. The first is that embedded CPs and IPs are distinguished by the feature COMP. A CP is marked [+ COMP] and an IP is marked [− COMP]. Second, they follow Doherty (1993) and Grimshaw (1997) in assuming that a complement clause is a CP if it has an overt complementizer and an IP if it does not. Third, they assume that an embedded clause in the input to GEN can be freely specified as [+ COMP] or as [− COMP].

The OT analysis they propose is quite simple and elegant. It involves only the two basic OT constraint types: faithfulness and markedness constraints. The faithfulness constraint is based on the feature [+ COMP] (Baković and Keer, to appear: (6)).

\[
(7) \text{F\text{AITH}[COMP]: The output value of [COMP] is the same as the input value.}
\]

The markedness constraints fall into two basic categories. There are constraints that target CP (the MARK-CP family) and ones that target IP (the MARK-IP family).

The analysis works as follows. If FAITH[COMP] outranks MARK-XP, where XP refers to either CP or IP, then the winner is the faithful candidate, as illustrated here:

\[
(8) \text{Complementizer optionality: FAITH[COMP] \gg MARK-XP}
\]

<table>
<thead>
<tr>
<th>Input: [+ COMP]</th>
<th>FAITH[COMP]</th>
<th>MARK-CP</th>
<th>MARK-IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [+ COMP] CP</td>
<td>F</td>
<td>*</td>
<td>![ ]</td>
</tr>
<tr>
<td>b. [+ COMP] IP</td>
<td>F</td>
<td>![ ]</td>
<td>![ ]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: [− COMP]</th>
<th>FAITH[COMP]</th>
<th>MARK-CP</th>
<th>MARK-IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [- COMP] CP</td>
<td>![ ]</td>
<td>![ ]</td>
<td>![ ]</td>
</tr>
<tr>
<td>b. [- COMP] IP</td>
<td>![ ]</td>
<td>![ ]</td>
<td>![ ]</td>
</tr>
</tbody>
</table>

\[\text{7In their papers, Baković and Keer spell out the actual constraints that fall into each family of constraints. The Mark-IP and Mark-CP constraints in Baković and Keer (to appear) are essentially those used in the complementizer drop analysis of Grimshaw (1997).}\]

\[\text{8All tableaux in this section are adapted from Baković and Keer (to appear).}\]
However, if either type of markedness constraint is ranked higher than the faithfulness constraint, we get neutralization. That is, no matter whether the input is marked [+ COMP] or [− COMP], the same candidate wins. If MARK-IP outranks FAITH[COMP], then the complementizer is obligatory.

(9) Complementizer obligatoriness: MARK-IP ≫ FAITH[COMP]

<table>
<thead>
<tr>
<th></th>
<th>Input: [+ COMP]</th>
<th>MARK-IP</th>
<th>FAITH[COMP]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>CP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>IP</td>
<td>!</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Input: [− COMP]</td>
<td>MARK-IP</td>
<td>FAITH[COMP]</td>
</tr>
<tr>
<td>a.</td>
<td>CP</td>
<td></td>
<td>!</td>
</tr>
<tr>
<td>b.</td>
<td>IP</td>
<td>!</td>
<td>!</td>
</tr>
</tbody>
</table>

If MARK-CP outranks FAITH[COMP], then the complementizer is absent.

(10) Complementizer absence: MARK-CP ≫ FAITH[COMP]

<table>
<thead>
<tr>
<th></th>
<th>Input: [+ COMP]</th>
<th>MARK-CP</th>
<th>FAITH[COMP]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>CP</td>
<td>!</td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>IP</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td></td>
<td>Input: [− COMP]</td>
<td>MARK-CP</td>
<td>FAITH[COMP]</td>
</tr>
<tr>
<td>a.</td>
<td>CP</td>
<td>!</td>
<td>!</td>
</tr>
<tr>
<td>b.</td>
<td>IP</td>
<td>!</td>
<td>!</td>
</tr>
</tbody>
</table>

In this fashion, Baković and Keer get the right results for the cases they examine, lending some empirical bite to the neutralization approach.

But, there are various arguments against both the particular analysis that Baković and Keer present and the neutralization approach in general. I will start with considering the problems with this particular analysis. First, there is no independent motivation for the feature [+ COMP] beyond its role in the analysis of complementizer drop. Second, this feature is purely a diacritic to indicate whether something is a CP or an IP. That is, CPs are [+ COMP] and IPs are [− COMP]. Therefore, the feature is completely unnecessary, since the category information encodes the required distinction. However, category information is not usually considered to be in the input for a syntactic OT competition. This brings us to the third objection: the input is standardly assumed to be a semantic form of some kind, whether it is a predicate logic argument structure (Grimshaw, 1997), or a partially specified LFG functional structure (Bresnan, to appear). If this is the case, then the purely formal feature [+ COMP] should not and could not occur in the input.

There are also problems with the neutralization approach in general (Müller, 1999). Fundamentally, this approach increases the complexity of inputs, candidate sets, and candidate competition. First, because inputs can differ only on some formal feature, there will be many minimally different inputs. As a direct consequence, there will be more candidate sets to consider. Second, the candidates are allowed to vary from their inputs indiscriminately (at least with respect to these formal features). So, the constraints on GEN must be much weaker than standardly assumed in OT syntax, allowing it to generate all kinds of candidates and “everything competes with everything else in each competition” (Müller, 1999: 7). Third, this approach fosters spurious ambiguities when inputs are neutralized. For one competition the faithful candidate is optimal, while in the other competition (which differs only for the value of the input feature in question), the unfaithful candidate wins. That is, when faced with a neutralized sentence, a language learner cannot determine what the specification of the formal feature is in the input. As neutralization is a common issue in phonology, Prince and Smolensky (1993: 192) propose the mechanism of ‘Lexicon Optimization’ as a solution for the learnability problem. In OT syntax this term is not appropriate, so Müller (1999: 8) coins the more general term ‘Input Optimization’. Informally, the way this works is that the language learner selects the input for the most
harmonic output as the correct input. But, as Müller points out, this introduces a notion of second-order optimization into the theory: Input Optimization is selecting the best competition in a set of competitions, where each competition is the result of an optimization over a candidate set. This is quite complex from both a computational and learnability perspective.

3.3 True Optionality

A third way to select multiple winners in OT is by postulating true optionality, where this means multiple candidates that have exactly the same constraint profile. This is the approach that Grimshaw (1997) uses to account for English complementizer drop. Basically, in the cases where the complementizer is optional, neither optimal candidate violates any of the constraints, or else both candidates violate exactly the same constraints the same number of times. Schematically, the competition looks like this:⁹

(11)  IP and CP propositional complements

<table>
<thead>
<tr>
<th></th>
<th>C₁</th>
<th>C₂</th>
<th>C₃</th>
<th>C₄</th>
<th>C₅</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>☜</td>
<td>☜</td>
<td>☜</td>
<td>☞</td>
<td>☞</td>
</tr>
<tr>
<td>b.</td>
<td>☜</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
<td>☞</td>
</tr>
</tbody>
</table>

Both candidate (a) and candidate (b) are equally optimal by definition, as they have identical constraint profiles. Thus, there is no way for an OT grammar to pick between them.

This approach has the advantage that it introduces no extraneous mechanisms or features into the theory. But, there are two strong arguments against its adoption. First, it is very unlikely⁰ that the two candidates will remain tied in a large OT grammar with many constraints. If any constraint, no matter how lowly ranked, is violated by one candidate but not the other, then the optionality will disappear. It would seem to be an insurmountably difficult task to fine-tune a grammar to the point that no constraint differentiates just the candidates required for optionality. Similarly, a language learner would have to receive no information at all throughout the entire course of acquisition that would differentiate these two candidates according to even just one constraint.

The second argument against this approach, presented in Baković and Keer (to appear), is that true optionality makes strange typological predictions. If we assume that the inputs and candidates are universal¹¹ (Prince and Smolensky, 1993; Smolensky, 1996b) and that the only locus of typological variation is constraint ranking, then the

⁹I gloss over Grimshaw’s actual constraints, as the details of her analysis are not necessary for the arguments I give here. The tableau is based on Grimshaw (1997: 411) and Müller (1999: 4).

¹⁰In a previous version of this paper I presented an argument against true optionality based on simple probability theory. There are various problems with the argument (as pointed out to me by Chris Manning and Alan Gous), but it is still somewhat instructive. Given a constraint, Constraint₁, and two candidates, Candidate₁ and Candidate₂, there are two possible outcomes: either Candidate₁ and Candidate₂ are tied for Constraint₁ (both candidates have a no marks for this constraint or an identical number of marks), or there is no tie. In general, the probability of two candidates tying on n constraints is \( p_{tie}(n) = \frac{1}{2^n} \). It is easy to tell that, even with a small number of constraints, the probability of a candidate tie (i.e. true optionality) is very low. For example, with just twenty constraints, the probability of a tie is \( \frac{1}{1048576} \).

¹¹I am not necessarily advocating such an assumption here. Ultimately, I expect both assumptions to be untenable in a more mature version of the theory.
true optionality approach predicts that whatever optionality there is in one language should exist in all languages. The reason is simple: if two candidates are not distinguished by their constraint profiles, then no matter how the constraints are ordered they will still be indistinguishable; since the only difference between languages is constraint ranking, by hypothesis, the two candidates will be equally optimal in all languages. Of course, in any given language other candidates could be more optimal than either of the two considered here. However, these two arguments make it clear that the true optionality approach is not a good solution, because it raises more problems than it solves.

3.4 Constraint Ties

The approaches considered so far assume a standard OT grammar whose constraint set forms a total order. The final approach to optionality I consider is that of allowing constraint ties,\textsuperscript{12} which effectively relaxes the requirement of a total order by permitting equally ordered (or unordered) constraints (i.e. constraints that are “tied” for a position in the ordering).

3.4.1 Partial Orderings

One of the properties of a total order is that of connectedness: every constraint must be ranked relative to every other constraint (Anttila, 1997a). If we do not require this property, then we can characterize an OT grammar as a partial order. In this case some of the constraints will not be ranked relative to some others. A schematic view of a total ordering and a partial ordering of three constraints is given in (12):

\begin{center}
\begin{tabular}{c|c}
(12) & Total Ordering & Partial Ordering \\
\hline
A & B & C \\
B & C & \\
C & & \\
\end{tabular}
\end{center}

The total ordering states that A \(\gg\) B \(\gg\) C, whereas the partial ordering states that A \(\gg\) B and that A \(\gg\) C. However, in the partial ordering there is no ordering between B and C. It is neither the case that B \(\gg\) C nor that C \(\gg\) B.

Assuming \(n\) constraints that are partially ordered, \(n!\) possible total orderings of just these constraints can be derived. In this case, we get the two total orderings A \(\gg\) B \(\gg\) C and A \(\gg\) C \(\gg\) B. Insofar as the total orderings derived will constitute different constraint rankings, different candidates can be selected as optimal on different resulting total orders. Thus, partial orders offer the most general method for relaxing the total ordering on constraints and deriving some optionality in output.

There are two major problems with general partial orders. To understand the first problem, consider a partial order like the following:\textsuperscript{13}

\begin{center}
\begin{tabular}{c|c}
(13) & \\
\hline
A & B C \\
B & C \\
C & \\
\end{tabular}
\end{center}

This partial ordering corresponds to twenty-five total orderings, which can be represented with these three schemas:

\textsuperscript{12}This is also the most widely utilized approach to optionality. Relevant references include Anttila (1997a,b); Boersma (1997, 1998, to appear a); Boersma and Hayes (1999); Pesetsky (1997, 1998); Prince and Smolensky (1993). See Müller (1999) for further references and discussion.

\textsuperscript{13}The elide between C and Z indicates a total ordering of twenty-four constraints, with C as the highest-ranking constraint.
The problem is evident: since the total ordering from C to Z must be maintained, the various resolutions give radically differing importance to the constraint B. The two extreme cases are provided by (14a) and (14c). In (14a), B is very highly ranked and can only be violated in order to satisfy A. However, in (14c), the same constraint B is ranked very low and will not play a significant role in selecting a candidate. This is a very serious conceptual problem with partial orders. Furthermore, it will make modelling optionality very difficult: if the partial ordering is between highly-ranked constraints in a large constraint set, the two total orders will be radically different, most likely yielding no optionality after all. Thus, there is not only a conceptual problem, but we also encounter again the problem of constructing a grammar where other constraints do not interfere in the optionality, as is the case with the true optionality account.

The second problem with this approach is that there is no non-exponential learning algorithm for partially ordered OT grammars (Boersma and Hayes, 1999; Boersma, to appear b; Kager, 1999). Neither Tesar and Smolensky’s (1998) nor Boersma and Hayes’s (1999) learning algorithm works for such grammars. In fact, the former learning algorithm only works for totally ordered OT grammars without ties (I return to this point in the following sections). As there are learning algorithms for the other approaches I review below, they are to be preferred to partially ordered grammars.

3.4.2 Stratifiable Partial Orderings

Anttila (1997a; 1997b) proposes a slightly modified version of partial ordering to handle optionality in Finnish genitive plurals.14 His approach is essentially a hybrid of total and partial orderings: constraints fall into sets which are totally ordered with respect to each other; however, within the sets there can be partial ordering. I will not go into the details of Anttila’s analysis, as it is quite rich and complex, but the three highest ranking sets he proposes are given here.15

<table>
<thead>
<tr>
<th>Set 1</th>
<th>Set 2</th>
<th>Set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>*X.X</td>
<td>*L</td>
<td>*H/I</td>
</tr>
<tr>
<td></td>
<td>*H</td>
<td>*Í</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*L.L</td>
</tr>
</tbody>
</table>

This stratified partial ordering corresponds to the following twelve total orderings:

1. *X.X >> *L >> *H >> *H/I >> *Í >> *L.L
2. *X.X >> *I >> *H >> *H/I >> *L.L >> *Í
3. *X.X >> *L >> *H >> *Í >> *H/I >> *L.L
4. *X.X >> *I >> *H >> *Í >> *L.L >> *H/I
5. *X.X >> *L >> *H >> *Í >> *H/I >> *L.L
7. *X.X >> *H >> *L >> *H/I >> *Í >> *L.L
8. *X.X >> *L >> *H >> *H/I >> *Í >> *L.L
9. *X.X >> *H >> *L >> *Í >> *H/I >> *L.L
10. *X.X >> *H >> *Í >> *L.L >> *Í >> *H/I
11. *X.X >> *H >> *L >> *H/I >> *Í
12. *X.X >> *H >> *L >> *Í >> *L.L >> *Í >> *H/I

---

14 The term ‘stratified partial ordering’ comes from Boersma (to appear b).
15 What the constraints mean is not relevant here. The curious reader is invited to consult Anttila’s work.
As the partial orderings are only within each set/stratum, this approach does not suffer from the first problem of (general) partial orderings discussed above. In other words, the various total orderings are minimally different from each other and no one constraint varies tremendously in its ranking position.

Using this approach, Anttila (1997a,b) gets very interesting results in both the synchronic and diachronic dimensions of grammar. Furthermore, he makes accurate predictions about the actual distribution of alternants of the genitive plural in Finnish corpora. I contend that it is problematic to expect grammars to predict actual percentages of forms in corpora, as performance factors surely mediate between the forms produced by the grammar and forms listed in a corpus. However, supposing that we do wish to use the grammar to predict actual percentages, we quickly run into a strange consequence of stratifiable partial orderings. Suppose that a given candidate is optimal under three of the twelve orderings in (16). The prediction is that 25% of the forms attested should be this candidate. Boersma and Hayes (1999) point out that there are cases of free variation in which one variant is considerably more frequent than the other. An example is English words that phonemically end in /t, d/ and /n/. These are normally realized with no schwa and a syllabic /n/, as in Swedish [swi:dn], but are occasionally realized with the schwa, as in [swi:dan]. Let us assume purely for the sake of argument that the rarely attested form occurs once for every thirty-two times the commonly attested form does. For a grammar with three partially ordered constraint sets, the first and second of which contain two constraints and the third of which contains four constraints, there are 96 (2! 2! 4!) total orderings. Thus, the rarer form would have to be optimal on three of the total orderings. Or, for a grammar with thirty-two constraint sets, each of which contains only one constraint, the rarer form would have to be optimal on one total ordering. In any case, we have a situation in which the frequency of occurrence of the forms produced by the grammar partially determines the number of constraints in the grammar and their organization into strata. This is a very peculiar state of affairs, as a grammar with fewer constraints should in principle be preferred to a grammar with more constraints, providing they make identical predictions about grammaticality.

Even if we put this conceptual problem aside, there is a second problem which has to do with learnability (Boersma, to appear b; Boersma and Hayes, 1999; Kager, 1999). Tesar and Smolensky’s (1998) learning algorithm cannot handle stratifiable partial orderings: it either fails to acquire them or it never terminates (Boersma and Hayes, 1999). However, as we will see in the next section, stratifiable partial orderings are just a special case in stochastic Optimality Theory, which uses a modified notion of constraint ranking, which associates constraints with a ranking value on a continuous scale. Stochastic OT has an associated learning algorithm: the Gradual Learning Algorithm (Boersma, to appear a; Boersma and Hayes, 1999). Since the continuous ranking approach avoids the first problem noted above (see below), is learnable, and can handle stratified partial orderings as a special case, it will be the approach I adopt.

3.4.3 Stochastic Optimality Theory

Boersma (1997, 1998, to appear a) and Boersma and Hayes (1999) present a methodology for representing optionality in a stochastic version of Optimality Theory. The method involves assigning a ranking value to each constraint. That is, each constraint has a ranking value along a continuous scale and at constraint evaluation time the ranking of a constraint is randomly distributed around this ranking value (see figure 1 below). The upshot of this is that if two constraints are closely ranked, there will be optionality in which constraint is ranked higher than the other. The frequency of a given ranking (e.g. whether C₁ outranks C₂), depends on the ranking difference. If the ranking difference is high, the constraint with the higher ranking value has a very high probability of outranking the constraint with the lower ranking value. If the ranking difference is 0, each constraint will outrank the other fifty percent of the time.

Thus, constraint evaluation is stochastic in this system. Constraints do not deterministically outrank other constraints, but rather do so probabilistically. This probability is calculated using a Gaussian (i.e. normal) distribution, and the formula is not trivial. For the mathematical details, see Boersma (1997, 1998, to appear a). However, the effects for two constraints — C₁ and C₂ — are illustrated in the following table (Boersma, 1997):

\[
\begin{array}{cccccccccccc}
| r_1 - r_2 | & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\
| P \quad | & 0.5 & 0.46 & 0.24 & 0.14 & 0.079 & 0.039 & 0.017 & 0.007 & 0.002 & 7 \cdot 10^{-4} & 2 \cdot 10^{-4} \\
\end{array}
\]
The top line indicates the ranking difference between the ranking values of the two constraints, $r_1$ and $r_2$. The bottom line indicates the probability, $P$, of the second constraint outranking the first one.\footnote{The ranking values assigned to constraints do not have to be integers; finer grained probabilities can be assigned using positive real numbers.}

If two constraints are ranked closely enough, on different evaluations one will be ranked higher than the other. This yields optionality. However, if two constraints are ranked fairly far apart, the probability of different rankings becomes small. That is, the ranking becomes close to obligatory (Boersma, 1997). It is important to realize that at evaluation time there will always be a total ordering. Since the probability of the first constraint outranking the second is relative to each evaluation, in any given evaluation the ranking is determinate.

It should be evident that stratifiable partial orderings (Anttila, 1997a,b) are simply a special case in this approach. A stratified partial ordering is derived in a constraint ranking where several constraints have exactly the same ranking value. For example, consider the following schematic tableau, which has the ranking value for each constraint indicated below the constraint’s name.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
 & $C_1$ & $C_2$ & $C_3$ & $C_4$ & $C_5$ \\
\hline
150 & 100 & 100 & 100 & 50 \\
\hline
\end{tabular}
\end{table}

Bearing in mind that each constraint is associated with the same normal distribution, constraints $C_2$, $C_3$, and $C_4$, are equally likely to occur in any of the six (3!) possible orders. The other two constraints will not enter into the optionality, as their ranking values are too far from the constraints ranked at 100. This is equivalent to a stratified partial ordering with three strata: $C_1 \gg \{C_2, C_3, C_4\} \gg C_5$.

In my analysis of optionality in Marathi linking (section 6), I assign the optionally ranked constraints the same ranking value. So I effectively use a stratified partial ordering. The reasons for using the stochastic OT model instead of a simple stratified partial ordering were given at the end of the last section. First, the stochastic model does not depend on a direct relationship between the number of constraints in the grammar (and how they are divided into strata) and the frequency of the optional forms. In the stochastic model, the frequency of forms depends only on the ranking value of the constraints in question.\footnote{Another relevant factor is the standard deviation of the normal distribution associated with the constraints, but as this is the same for every constraint, it is a constant.} Therefore, on this approach it is possible to have a simpler grammar, with fewer constraints, and still predict the same frequency data.\footnote{I take this to be a conceptual point about the undesirability of frequency data dictating the number of constraints in the grammar. However, as I stated in section 3.4.2, I do not believe that the grammar should ever predict exact frequencies, as there will be many intervening factors which will distort the frequency of forms attested.}

The stochastic model also has a learning algorithm — the Gradual Learning Algorithm (GLA; Boersma to appear a; Boersma and Hayes 1999) — and grammars stated in this model are therefore learnable. Stratified partial orderings have only been demonstrated to be learnable by the GLA (Boersma, to appear b). Until a separate learning algorithm that does not depend on a stochastic model of OT is shown to learn stratifiable partial orderings, there is no choice but to implement them using continuous ranking and stochastic Optimality Theory.

Thus, continuously-ranked constraints offer a way to represent optionality which avoids the problems associated with previous approaches. The pseudo-optionality approach’s problem with overgeneration does not arise, because...
here the various options belong to the same competition. The problem of maintaining identical constraint profiles in the true optionality approach does not arise because the candidates win on different rankings and do not need to have identical constraint profiles. The various problems with the neutralization approach are dispensed with, because there is only one input and it does not require formal features, as will be demonstrated. Lastly, in this section I have demonstrated that this approach is the soundest and most general way of implementing the notion of tied constraints, as it is equivalent to stratifiable partial orderings as a special case, but is also learnable. In the following sections I review the facts about Marathi that I will be accounting for, as well as Joshi’s (1993) insightful LFG account of these facts. In section 6, I present my own analysis, which uses continuously-ranked constraints in a stochastic Optimality Theory.

4 Subject/Object Alternation in Marathi

Marathi\(^{19}\) presents an interesting case for studying optionality in Optimality Theory. Certain verbs in this language show an alternation in which argument can be realized as the subject or object. That is, given two arguments, each can be mapped to the subject while the other is mapped to the object. The verb classes that exhibit this property are non-volitional transitives, such as ‘saapâdhṛ’ (to find) and ‘nawâdhṛ’ (to like), passives of underived ditransitives, such as the passive of ‘depe’ (to give), and passives of derived ditransitive causatives, such as the passive of causativized ‘paâdzṛ’ (to feed).\(^{20}\) In the next section I will present examples of all of these verbs and the evidence that Joshi uses to argue for the alternation of their core grammatical functions.

4.1 The Data

Here are examples\(^{21}\) of the verb classes mentioned in the previous section.

(19) Non-volitional transitives

a. sumaa-laa ek pustak milaale.
   Suma-D/A one book got
   Suma got a book.
   (‘milne’ (to receive) ∈ ‘find’ class)

b. sumaa-laa aâi̯saa upadeś patlaa.
   Suma-D/A mother’s advice agreed
   Mother’s advice became acceptable to Suma.
   (‘paṭne’ (to accept mentally) ∈ ‘like’ class)

(20) Passive ditransitive

baâḷa-laa kâau dilaa gelaa.
baby-D/A food given was
Baby was given food.

(21) Passive ditransitive causative

ramaâ-laa duď paâdzle gele.
Rama-D/A milk drink.CAUS was
Rama was fed milk.

\(^{19}\)Marathi is an Indo-Aryan language spoken in Maharashtra and adjacent states in India. It has roughly 64,000,000 speakers. Ethnographic data from Ethnologue (http://www.sil.org/ethnologue/countries/Inda.html#MRT; checked 17.03.99).

\(^{20}\)See appendix A for a list of verbs in these classes.

\(^{21}\)All Marathi examples are taken from Joshi (1993) and I have used her transcriptions, with one exception. She gives various case feature transcriptions (see footnote 22), for the suffix -laa, whereas I normally gloss it as dative/accusative, since the distinctions are never relevant in this paper. See appendix D for example citation information and appendix B for an index of relevant gloss abbreviations.
Joshi (1993) separates the non-volitional transitives into the ‘find’ class (19a) and the ‘like’ class (19b).

4.2 Evidence for the Subject/Object Alternation

Joshi (1993: 37–44) lists three syntactic tests for grammatical subjecthood in Marathi: the conjunction reduction test, the prenominal present participial relative test, and the un clause controller test. She uses all three tests to demonstrate the subject/object alternation in the verb classes under consideration, but in this section I will only review the conjunction reduction test.

The conjunction reduction test is not solely a test for subjects, but rather distinguishes between grammatical functions in general. The phenomenon can be characterized as follows: if two clauses with coreferent arguments are conjoined, the coreferent argument in the second clause must be null, but only if the two arguments have the same grammatical function and have the same case marking (Joshi, 1993: 37–38). 22 For illustration of these requirements on conjunction reduction the reader is invited to consult Joshi (1993). I will show that only subjects can gap subjects, as this will be important in demonstrating the subject/object alternation momentarily.

The following examples have the properties required. 23 Example (22) shows that subjects can gap subjects and must do so if the cases match, example (23) shows that subjects cannot gap objects, and example (24) shows that objects cannot gap subjects.

(22) a. [alka-ni mini-łaa maarle] pan [niilimaa-łaa wiihirit d\textsuperscript{b}akalle].
   Alka-ERG mini-D/A hit but Nilima-D/A well-in pushed
   Alka\textsubscript{i} hit Mini, but \textit{Ø} pushed Nilima in the well.

b. *[alka-ni mini-łaa maarle] pan [tini niilimaa-łaa wiihirit d\textsuperscript{b}akalle].
   Alka-ERG mini-D/A hit but she-ERG Nilima-D/A well-in pushed
   Alka\textsubscript{i} hit Mini, but she\textsubscript{i} pushed Nilima in the well.

(23) a. *[d\textsuperscript{a}zaamb\textsuperscript{b}}-ji-ci pikleli p\textsuperscript{b}ale k\textsuperscript{b}aali paḍ\textsuperscript{di}i] aapi [raam-ni go\textsuperscript{a}a keli].
   Jamun-GEN ripe fruits down fell and Ram-ERG collect did
   [Ripe fruits of the Jamun tree\textsubscript{i} fell on the ground, and Ram collected \textit{Ø}.

b. [d\textsuperscript{a}zaamb\textsuperscript{b}}-ji-ci pikleli p\textsuperscript{b}ale k\textsuperscript{b}aali paḍ\textsuperscript{di}i] aapi [ti raam-ni go\textsuperscript{a}a keli].
   Jamun-GEN ripe fruits down fell and those Ram-ERG collect did
   [Ripe fruits of the Jamun tree\textsubscript{i} fell on the ground, and Ram collected them\textsubscript{i}.

(24) a. *[pustake raam-ni ṭeblaa-war ṭ\textsuperscript{e}wli] paṃ [k\textsuperscript{b}aali paḍ\textsuperscript{di}i].
   books ram-ERG table-on kept but down fell
   Ram put books\textsubscript{i} on the table but \textit{Ø} fell down.

b. [pustake raam-ni ṭeblaa-war ṭ\textsuperscript{e}wli] paṃ [ti k\textsuperscript{b}aali paḍ\textsuperscript{di}i].
   books ram-ERG table-on kept but they down fell
   Ram put books\textsubscript{i} on the table but they\textsubscript{i} fell down.

These examples have demonstrated how conjunction reduction functions in Marathi. Next I will show how this test provides evidence for the subject alternation.

4.3 Subject Alternation

First, the conjunction reduction test will be used to demonstrate that either argument in non-volitional transitives (‘like’ and ‘find’ verbs) can be the grammatical subject. This follows from conjunction reduction because, as demonstrated

---

22 It is important to realize that the two arguments have to have the same case marking, but can have differing case features, or abstract case. For example, the dative/accusative case and the ability case in Marathi are both marked with the suffix -laa, but Joshi (1993: 18–30) argues that they must be distinguished as separate cases.

23 In all these examples, the case-matching requirement has been controlled for.
in (22), “when the subjects of conjoined clauses are coreferent and have matching cases, the second subject is null” (Joshi, 1993: 53). The following examples illustrate that either argument of a non-volitional transitive can be gapped in the second conjunct.

(25)  
\textbf{‘like’ class}

\textit{a. Dative/accusative argument gapped}

\[
\text{rajat-laa aai-laa exći {taakun yewawlə naahi]} \text{ pan [hyaa citrakaarəa-imeline maagʃə pradarʃə kʰup Rajat-Abi. mother-D/A alone leaving come able.not but this artist-GEN last exhibition much aawadəla hota]. liked’ had}
\]

\textit{Rajat did not feel like leaving mother alone and coming, but }Ø\textit{ had much liked the artist’s last exhibition.}

\textit{b. Nominative argument gapped}

\[
\text{sumaa goḍ hasli] aani [rajat-laa aawadə]. Suma sweet smiled and } \text{Rajat-D/A liked}
\]

\textit{Suma smiled sweetly and Rajat liked }Ø\textit{.}

These sentences illustrate that either argument of the non-volitional transitive ‘aawadə’ (to like) can be gapped by a coreferent subject in a preceding conjoined clause. This indicates that both arguments of a ‘like’ class verb can serve as the grammatical subject.

The following sentences show that the arguments of ‘find’ class verbs have the same properties.

(26)  
\textbf{‘find’ class}

\textit{a. Dative/accusative argument gapped}

\[
\text{raajkumaarəa-laa waaḷwaŋṭəat sodəla gela] \text{ pan [parat gaay-imeline rastaa saapaḍlaa]. prince-D/A desert.in’ left was but back go-GEN way found}
\]

\textit{The prince was left in the desert, but }Ø\textit{ found his way back.}

\textit{b. Nominative argument gapped}

\[
\text{raajkumaarəa-ci angtəi talaayat paḍli] aani [ekaa kolyaa-laa saapaḍlii]. prince-GEN ring lake.in fell and one fisherman-D/A found}
\]

\textit{The prince’s ring fell in the lake and one fisherman found }Ø\textit{.}

Again, since the subject in the first clause can gap either of the ‘find’ verb’s arguments, this shows that either argument can be the grammatical subject.

The conjunction reduction test can also be used to demonstrate that either argument in a passive of a ditransitive or ditransitive causative can be the grammatical subject.

(27)  
\textbf{Passive ditransitive}

\textit{a. Dative/accusative argument gapped}

\[
\text{mini-laa kʰurcit basawle gele] aani [ek pustak dile gele]. Mini-D/A chair.in sit.made was and one book given was}
\]

\textit{Mini was seated in a chair and }Ø\textit{ was given a book.}

\textit{b. Nominative argument gapped}

\[
\text{ek pustak kʰaali paḍle] aani [mini-laa dile gele]. one book down fell and Mini-D/A given was}
\]

\textit{One book fell down and }Ø\textit{ was given to Mini.}
Passive ditransitive causative

a. **Dative/accusative argument gapped**
   
   [mini-laₐ gₐri nele gele] aanī [te ḍud₈ paₐd³le gele].
   Mini-D/A home taken was and that milk drink.cause was
   Miniᵢ was taken home and Ø₁ fed milk.

b. **Nominative argument gapped**
   
   [te ḍud₈ taaple] aanī [mini-laₐ paₐd³le gele].
   that milk warm.became and Mini-D/A drink.cause was
   The milkᵢ became warm and Ø₁ was fed to Mini.

Once again a subject can gap either of the two arguments of these verbs. Therefore, either argument can be the subject.

In this section, I presented Joshi’s evidence that either argument in the verb classes with subject/object alternation can be the grammatical subject. In the next section, I will show that either argument can also be the grammatical object.

### 4.4 Object Alternation

Testing for the object alternation in Marathi is slightly more complex than testing for the alternating subject. It involves using a controlled adjunct, the *taanaa* participle adverbial clause, in tandem with conjunction reduction. The basic characterization of the *taanaa* adverbial is that it has an obligatorily null subject which can only be controlled by an argument which has a non-oblique grammatical function (i.e. subjects and objects). Joshi’s strategy is to use conjunction reduction to establish that one argument of the relevant alternating verb is a subject and then to show that the other argument can control a *taanaa* adverbial clause. Under the assumption that there can only be one subject per clause in Marathi, and given that only non-obliques can control the gap in these adverbials, it follows that the controller must therefore be a grammatical object.

In the following examples, conjunction reduction is used to establish that one of the alternating verb’s arguments is the grammatical subject. At the same time the *taanaa* control test shows that the other argument must be a core grammatical function. Since there is an uncontroversial subject, this other argument must be the object.

(29) **Non-volitional transitive (‘like’ class)**

a. **Dative/accusative object as controller**
   
   Princess Rajat-D/A bath do.PART saw and disappear became
   While Øᵢ/j taking a bath Rajatᵢ saw the princessᵢ and Øⱼ disappeared.

b. **Nominative object as controller**
   
   Rajat-D/A princess bath do.PART saw and imprisoned done was
   Rajatᵢ saw the princessᵢ Øᵢ/j bathing and Øᵢ was imprisoned.

In sentence (29a), *raajakanyaa* is gapped under conjunction reduction with the subject of *naahi₇i* (disappeared). Therefore, *raajakanyaa* is the subject of *disli* (saw). Since *rajat* can still control the adverbial *kartaanaa*, it must be the grammatical object of *disli*. Likewise, in (29b), *rajat-laₐ* is coreferent with the gapped subject of *girafdaar* and must therefore be the subject of *disli*. The adverbial *kartaanaa* can be controlled by *raajakanyaa* and *raajakanyaa* is therefore the grammatical object of *disli*. Thus, either argument of the non-volitional transitive ‘like’ verb *disпе* can be the grammatical object.

The test yields the same results for ‘find’ verbs, passive ditransitives, and passive ditransitive causatives. In each case, conjunction reduction controls for the subject, and either argument can be the grammatical object. The relevant examples follow, but I will not explain each one in detail.
Non-volitional transitives (‘find’ class)

a. Dative/accusative object as controller

> kaahi mor rajat-laaj [janglaat pʰir-taanaa] aadʰalle aani [ekaa kʰāṇbʰaraa-t-āʃi naahise dʰaale]. some peacocks Rajat-D/A jungle.in roam-PART came.across and one moment-in-only disappear happened Rajat came across some peacocks while Ø roaming in the jungle, and Ø in a moment disappeared.

b. Nominitative object as controller

> rajat-laaj kaahi mor [janglaat naaf-s-taanaa] aadʰalle pan [tcyaancyaa dźawaľ dźawaawle naahii]. Rajat-D/A some peacocks jungle.in dance-PART came.across but they.gen near go.able not Rajat came across some peacocks when Ø dancing in the jungle, but Ø could not go near them.

Passive ditransitives

a. Dative/accusative object as controller

> aai-ni magawleli pustaka sakaal-cyaa [papaalaa-ni aali] aani [[saalet dźaa-taanaa] sumaa-laaj daakʰawli mother-ERG ordered books morning post-with came and school.to go-PART Suma-D/A shown geli]. were [The books that mother had ordered]i came with the morning post and Øi were shown to Suma, while Øj going to school.

b. Nominative object as controller

> sumaa-laaj madʰya-raatri utʰawla gela aani [hiwaalya-tsa pahila barfa [paṭtaanaa] daakʰawla gela]. Suma-D/A middle.of-night woken was and winter.of first snow falling shown was Suma was woken up at midnight and Øi was shown the first snow of winter when Øj falling.

Passive ditransitive causatives

a. Dative/accusative object as controller

> dudʰ garam dźʰaala aani [raam-laaj [dzʰoptaanaa] paadźla gela]. milk warm became and Ram-D/A sleep.PART drink.CAUS was The milk got warm and Ram was fed Øi while Øj going to sleep.

b. Nominative object as controller

> raam-laaj bʰuk laagli aani [dudʰ [garam astaanaa] paadźla gela]. Ram-D/A hunger came and milk hot be.PART drink.CAUS was Ram became hungry and Øi was fed some milkj while Øj hot.

In the last two sections, I have presented Joshi’s evidence that in Marathi non-volitional transitives, passive ditransitives, and passive ditransitive causatives either argument can be the grammatical subject while the other is the grammatical object. Therefore, these verb classes demonstrate true optionality in linking their arguments to grammatical functions. In the next section, I outline the relevant aspects of Joshi’s LFG analysis, before turning to my own OT analysis in section 6.

5 Joshi’s Analysis

In her thesis, Joshi uses a combination of Lexical Mapping Theory (LMT; Bresnan and Kanerva 1989; Bresnan and Zaenen 1990), which is a component of LFG, Dowty’s proto-role theory (Dowty, 1991), and an unordered argument structure representation to give rules that provide predicates with their required linking properties.
Lexical Mapping Theory provides two syntactic features \([\pm r (estricted)]\) and \([\pm o (bjective)]\) which cross-classify grammatical functions. Joshi (1993: 113) assumes the following classification:\(^{24}\)

\[
\text{SUBJ: } [-r, -o], \text{ OBJ: } [-r, +o], \text{ OBL: } [+r, -o]
\]

(Joshi, 1993: 113)

In Joshi’s analysis, the LMT features of an argument are determined by the proto-roles associated with the argument. The proto-properties she assumes are:

(34) **Proto-Agent Properties**

a. volitional involvement in the event or state
b. sentience (and/or perception)
c. causing an event or change of state in another participant
d. movement (relative to the position of another participant)
e. exists independently of the event named by the verb

(35) **Proto-Patient Properties**

a. undergoes change of state/location
b. incremental theme
c. causally affected by another participant
d. stationary relative to movement of another participant
e. does not exist independent of the event, or not at all
f. not entailed to possess sentience/perception

(Dowty 1991: 572; Joshi 1993: 80)

I follow Joshi in referring to these properties as PAa, PAb, PPa, PPb, and so forth.

Proto-property PPf is not one of the original Dowty (1991) properties, but is proposed by Joshi (1993). But, this property is quite problematic and I do not adopt it. First, it is not clear what its motivation is. Although, as Joshi (1993: 90) notes, lack of sentience may be a reasonable proto-patient property, the entailment should be non-sentience, as opposed to PAb. The reason Joshi opts for the present formulation instead, is to get certain facts about Marathi linking right (see (39) below). Thus, the problem is a lack of independent motivation.

Second, this proto-property works completely differently from the other ones.\(^{25}\) In Joshi’s formulation, this proto-property will necessarily be present whenever PAb is absent. In other words, any argument of a verb that is not entailed to have the proto-agent property of sentience (PAb) will *automatically* receive property PPf. It is clear, as Dowty (1991) notes, that some proto-properties are related, in particular PAc-d and PPC-d, but they are related positively, not negatively. For example, if one argument is a causer (PAc), then there must also be an argument that is the causee (PPc). Thus, the presence of the first property entails the presence of the second, and vice versa.\(^ {26}\) However, PPf is different because its presence is entailed by the *absence* of another property. Joshi has therefore introduced a completely new dynamic into Dowty’s system, changing it considerably.

She then gives two syntactic feature rules for Marathi:

---

\(^{24}\)The restricted object, OBJb, which has the features \([+r, +o]\), is omitted. Joshi argues that Marathi is a symmetric language (i.e. it does not distinguish between a direct and indirect object).

\(^{25}\)Farrell Ackerman and his students (p.c.) have made this point to me, in a slightly different manner.

\(^{26}\)Of course, the argument that receives a certain proto-entailment can be suppressed. For example, a passive of a causative will have the causer suppressed, but the entailments still hold of the base predicate to which the passive is related.
She also proposes the following function mapping principle:

(37)  
\[
\begin{align*}
\text{(i) } & [+\text{volitional}] \\
& \quad \Rightarrow \text{subject; otherwise:} \\
\text{(ii) } & [-r] \\
& \quad \Rightarrow \text{subject}
\end{align*}
\]

b. Other arguments are mapped onto the lowest compatible function on the markedness hierarchy \text{SUBJ} > \text{OBJ, OBL}.\textsuperscript{27}

Finally, she assumes the following well-formedness condition:

(38) The Subject Condition: every predicate must have a single token of the grammatical function \text{SUBJ}.

The basic idea of Joshi’s analysis is that proto-role information determines syntactic features and these in turn determine grammatical function. I will now give an example of how this works in Joshi’s system. The example is for the non-volitional verb \textit{aawadhe} (Joshi, 1993: 117).

(39) Mapping of grammatical functions in the verb \textit{aawadhe} (‘like’)

<table>
<thead>
<tr>
<th>Undergoes change of state (PPa)</th>
<th>Not E.T.B. Sentient (PPf)</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{aawadhe} ‘to like’</td>
<td>{ARG [-r], ARG [-r]}</td>
<td>ARG STR</td>
</tr>
</tbody>
</table>

Due to the syntactic feature rules in (36a), the two arguments get the feature assignment \([-r]\), since they have proto-patient properties. Given this feature assignment, either argument can be a subject, according to (37aii). Finally, each argument can equally be an object, according to (37b), since the lowest role in the grammatical function hierarchy that they are compatible with is \text{OBJ} (obliques are \([+r]\)). Thus, she successfully accounts for the subject/object alternation using these mechanisms. The other alternating predicates are handled in a similar manner.

This example also illustrates how the problematic proto-patient property, PPF, is necessary for Joshi’s analysis. It is clear that the second argument of \textit{aawadhe} (‘like’) has no proto-patient properties other than PPF. It does not undergo a change of state or location (PPa). It is not an incremental theme (PPb). It is not causally affected by the other argument (PPc). It is not necessarily stationary relative to the other argument (PPd). Finally, it exists independently of

\textsuperscript{27}This hierarchy arises from the assumption that, in LMT, negative features are less marked than positive ones (Joshi, 1993: 105).
the liking event (PPe). However, if the second argument gets no proto-patient property, it does not get the feature \([-r]\) by (36a), but rather gets assigned \([-o]\) by (36b). If it gets assigned \([-o]\) then it must be the subject, since objects are \([+o]\). But then the optionality is unaccounted for and the analysis would wrongly predict that the verb aawadjne has no subject/object alternation and that the stimulus ‘likee’ must be the subject (perhaps analogously to English please, as in The food pleased me). This is clearly not the case, according to the evidence that Joshi presents. This verb exhibits the subject/object alternation and the likee can be the object as well as the subject.

Looking at it another way, Joshi’s (1993) theory yields the following two generalizations:

(40) a. If a predicate has a volitional argument, that argument is realized as the subject; otherwise any argument of the verb that possesses a proto-patient property can be the grammatical subject. (Joshi, 1993: 92)

b. If there is no volitional argument and there is more than one argument that has a proto-patient property there will be subject/object alternation.

Both of these principles are somewhat strange. Although Joshi clearly demonstrates the importance of the proto-agent property of volitionality, why should having a proto-patient property be the second-most important criterion for subjecthood? The whole point of Dowty’s (1991) theory is that proto-patient entailments are characteristic of objects, not subjects. Instead, (40) proposes that there is one proto-agent property that really matters for subjecthood, and that otherwise subjects are principally characterized by proto-patient properties. This is not at all in accordance with (Dowty, 1991) and it is extremely hard to see how both versions of the theory (Dowty’s (1991) and Joshi’s (1993)) could possibly be true.

Joshi’s analysis is insightful in its use of protorole information and its avoidance of an ordered argument structure. However, it is a parametric account and the particular instantiation of the parameters she gives are largely parochial to the analysis of Marathi. Thus, her account as such does not make strong crosslinguistic predictions. Furthermore, she stipulates the volitionality of the subject, rather than deriving it from general principles. She also proposes that proto-patient properties characterize subjects. Lastly, her analysis requires the use of the problematic proto-patient property PPF to get the descriptive facts right.

In the following section, I provide an OT account of grammatical function selection in Marathi. This account will use a small set of general, universally viable constraints to derive the correct linking of arguments in Marathi. In addition, it derives the volitionality of the subject from the harmonic alignment of two prominence scales and subsequent constraint interaction, and it does not use proto-patient property PPF.

### 6 An Optimality-theoretic Analysis

In this section I will provide an OT analysis for the following features of grammatical function selection in Marathi:

(41) a. If a predicate has a volitional argument, that argument is realized as the subject.

b. If there is no volitional argument there will be subject/object alternation.

The principal difference between this formulation and the one in (40) is that proto-patient properties are not assumed to be characteristic of subjects. Rather, the alternation facts will be shown to arise from a finer-grained analysis that uses a small set of crosslinguistically viable constraints.

Aissen (1999) has motivated the use of prominence scales and harmonic alignment to capture markedness universals. I will adopt her scale Subj(ec) > Nonsubj(ec), where “>” means “more prominent than”. I also propose the new scale Proto-Agent > Proto-Patient (P-A > P-P). “Proto-Agent” and “Proto-Patient” should be understood as cover terms for the set of properties listed in section 5. Thus, this scale is really a partial ordering, with the P-A properties unordered relative to each other, and likewise for the P-P properties. The idea is that investigating various languages will provide evidence for further ordering these properties. In fact, the importance of volitionality in Marathi indicates
that this property is possibly the highest ranked of the Proto-Agent properties. The fully articulated scale I will adopt is P-A_{vol} > P-A_{vol} > P-P.  

Now, I need to express the notion that subjects tend to bear proto-agent properties, and nonsubjects tend to bear proto-patient properties. This will be captured using the prominence scales for proto-role properties and grammatical functions. There is a formal method in OT for aligning two scales (Prince and Smolensky, 1993; Aissen, 1999), called harmonic alignment.

(42) Alignment: Given a binary dimension $D_1$ with a scale $X > Y$ on its elements $X, Y$, and another dimension $D_2$ with a scale $a > b \ldots > z$ on its elements, the harmonic alignment of $D_1$ and $D_2$ is the pair of Harmony scales:
   a. $H_a: X/a \succ X/b \succ \ldots \succ X/z$
   b. $H_b: Y/z \succ \ldots \succ Y/b \succ Y/a$

The constraint alignment is the pair of constraint hierarchies:
   a. $C_a: *X/a \succ \ldots \succ *X/b \succ *X/z$
   b. $C_b: *Y/a \succ \ldots \succ *Y/b \succ *Y/z$

(Prince and Smolensky, 1993: 136)

The result of the harmonic alignment is that it is more harmonic for $X$’s to be $a$’s and for $Y$’s to be $z$’s. This is captured in the constraint alignment by having $*X/a$ and $*Y/z$ be the lowest ranked constraints in their hierarchies. It is important to realize that these hierarchies are fixed universally. For example, no language may parochially have the ranking $*X/b \succ *X/z$. On the other hand, the resulting hierarchies may be interspersed, so long as the ordering for particular hierarchies is respected.

The alignment of the two scales that I am using will serve as a more concrete example:

(43) Subj $>$ Nonsubj, P-A_{vol} $>$ P-A_{vol} $>$ P-P

a. Harmonic alignment:
   Subj/P-A_{vol} $\succ$ Subj/P-A_{vol} $\succ$ Subj/P-P
   Nonsubj/P-P $\succ$ Nonsubj/P-A_{vol} $\succ$ Nonsubj/P-A_{vol}

b. Constraint hierarchies:
   *Subj/P-P $\succ$ *Subj/P-A_{vol} $\succ$ *Subj/P-A_{vol}
   *Nonsubj/P-A_{vol} $\succ$ *Nonsubj/P-A_{vol} $\succ$ *Nonsubj/P-P

The resulting constraint hierarchies are such that it is worst for a subject to have proto-patient properties and for a nonsubject to have the proto-agent property of volitionality. This does not mean that objects cannot be volitional; it just means that the least preferred situation is one in which a predicate entails the property of volitionality for an argument that gets mapped to a nonsubject grammatical function.

With these hierarchies in place, I can now present the constraint set that I will be using in this analysis:

(44) Constraints

28 Davis (1996) postulates that the proto-agent property of causation (PPc) is also a more important factor than the other proto-properties. I leave this aside here.

29 I have noted the specific volitionality property by subscripting P-A with vol. Similarly, I use the label P-A_{vol} as a shorthand for all the P-A properties except volitionality.

30 In fact, this is a very important distinction that Dowty (1991) discusses as well. The entire theory of proto-roles is about what the meaning of a predicate entails for that predicate’s arguments. Incidental properties of the actual fillers of these argument slots are not under discussion here.
a. MAX(proto)
Proto-role information in the input is realized in the output.

b. DEP(proto)
Proto-role information in the output is present in the input.\(^{31}\)

c. THEMATIC\(_{subj}\)
Every clause has a thematic subject.\(^{32}\)

d. *Nonsubj/P-A\(_{vol}\)
Avoid nonsubjects with the proto-agent property of volitionality.\(^{33}\)

e. *Subj/P-P
Avoid subjects with proto-patient properties.

f. *Subj/P-A\(_{vol}\)
Avoid subjects with proto-agent properties other than volitionality.

g. *Nonsubj/P-A\(_{vol}\)
Avoid nonsubjects with proto-agent properties other than volitionality.

h. *Nonsubj/P-P
Avoid nonsubjects with proto-patient properties.

The non-stochastic ranking of these constraints is: \(^{34}\)

\[
(45) \quad \text{Marathi ranking} \\
\text{THEMATIC}_{subj} \gg \text{MAX(proto)} \gg \text{*Nonsubj/P-A\(_{vol}\)} \gg \text{DEP(proto)} \gg \{ \text{*Subj/P-P, *Subj/P-A\(_{vol}\), *Nonsubj/P-A\(_{vol}\), *Nonsubj/P-P} \}
\]

In order to prove that a ranking is not just sufficient, but also necessary, a set of ranking arguments must be presented. As noted in Prince and Smolensky (1993: 106–107), this constitutes “a potential empirical argument that [Constraint] dominates [Constraint].”

The following tableau (Prince and Smolensky, 1993: 106, (160)) will be useful in explaining the form of a ranking argument.

(46) Ranking Argument

<table>
<thead>
<tr>
<th></th>
<th>C(_1)</th>
<th>C(_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(input)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. (\not\in) opt</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b. subopt</td>
<td>*(_1)</td>
<td></td>
</tr>
</tbody>
</table>

\(^{31}\)I also assume that there is a markedness constraint, *STRUC(proto), since faithfulness constraints usually have related markedness constraints. However, this constraint plays no role in the analysis presented here.

\(^{32}\)This constraint is related to the constraint SUBJ, which requires that every clause has a subject, thematic or otherwise [SUBJ is therefore analogous to the Extended Projection Principle of Government and Binding Theory or the Subject Condition of LFG’s Lexical Mapping Theory, a version of which is given in (38) above]. If a language has SUBJ and either DEP(proto) or *STRUC(proto) — the markedness constraint appropriate for proto-role properties — outranks THEMATIC\(_{subj}\), it will realize clausal subjects linked to arguments without proto-role properties as expletives. However, if THEMATIC\(_{subj}\) outranks SUBJ, there will be no expletive subjects. Either subjects with no proto-role properties will lead to ungrammaticality, or another argument will be co-opted as subject. The latter case occurs in Marathi verbs that exhibit the subject/object alternation. Marathi does not have expletive subjects (Pandharipande, 1997: 133), so I will assume that THEMATIC\(_{subj}\) outranks SUBJ and I do not include the latter constraint in my analysis.

\(^{33}\)Lødrup (1999) uses similar constraints to THEMATIC\(_{subj}\) and *Nonsubj/P-A\(_{vol}\) in his OT analysis of linking in Norwegian presentational focus constructions. However, he uses thematic roles, such as Theme and Agent, rather than proto-roles. His paper also contains more extensive discussion of the interaction between SUBJ and THEMATIC\(_{subj}\) (see footnote 32).

\(^{34}\)I present the actual stochastic version in (48) below.
In this tableau, \textit{opt} is a candidate that is attested in the data set, and is therefore known to be a winning candidate. The higher ranking constraint must not have any violations for \textit{opt}, whereas the lower ranking constraint must have one or more violations. The other candidate, \textit{subopt}, competes with \textit{opt}. The higher ranking constraint must be violated by this candidate, whereas the lower ranking constraint should not be violated. Since it is known that \textit{opt} must be an output, this configuration shows that \(C_1\) must outrank \(C_2\). If the reverse ranking, \(C_2 \gg C_1\), were picked, \textit{opt} would be erroneously rejected in favour of \textit{subopt}. Therefore, ranking arguments are very important in establishing the empirical adequacy of a given ranking and to show that a given ranking is both necessary and sufficient for the data set under consideration. Appendix C contains a full set of arguments for the ranking in (45).

I am using Boersma’s probabilistically ranked constraints (see section 3.4.3) to capture optionality, so the constraints in ranking (45) must be augmented with ranking values. There are four constraints that will capture the output, this configuration shows that \(C\) subject/object optionality in Marathi: *Subj/P-P, *Subj/P-A, *Nonsubj/P-A, and *Nonsubj/P-P. Since I do not have information about which argument maps to which grammatical function more frequently, I will make the baseline assumption that there is true optionality. In other words, I will assign these constraints the same ranking value, so that the ranking difference is 0. This yields 4! (= 24) different optional rankings; however, due to the universal hierarchies in (43b), only six of these outcomes are possible without violating either the hierarchy *Subj/P-P \(\gg\) *Subj/P-A or *Nonsubj/P-A \(\gg\) *Nonsubj/P-P. The six constraint orderings are:

\[
\begin{align*}
(47) & \quad a. & *\text{Subj/P-P} \gg *\text{Subj/P-A} \gg *\text{Nonsubj/P-A} \gg *\text{Nonsubj/P-P} \\
& \quad b. & *\text{Subj/P-P} \gg *\text{Nonsubj/P-A} \gg *\text{Subj/P-A} \gg *\text{Nonsubj/P-P} \\
& \quad c. & *\text{Subj/P-P} \gg *\text{Nonsubj/P-A} \gg *\text{Nonsubj/P-P} \gg *\text{Subj/P-A} \\
& \quad d. & *\text{Nonsubj/P-A} \gg *\text{Subj/P-P} \gg *\text{Subj/P-A} \gg *\text{Nonsubj/P-P} \\
& \quad e. & *\text{Nonsubj/P-A} \gg *\text{Subj/P-P} \gg *\text{Nonsubj/P-P} \gg *\text{Subj/P-A} \\
& \quad f. & *\text{Nonsubj/P-A} \gg *\text{Nonsubj/P-P} \gg *\text{Subj/P-P} \gg *\text{Subj/P-A}
\end{align*}
\]

Given the ranking value difference of 0, each of these rankings has a \(\frac{1}{6}\) probability of occurring.

A note of explanation is in order about the nature of hierarchies and their relationship to the stochastic model of OT that I am assuming here. By definition, hierarchies reduce the number of factorial rerankings to just those that respect the hierarchies, as demonstrated in (47). There are two ways to think about this, both of which are equivalent for my purposes. We could think of the hierarchies as a constraint on the generation of constraint rankings. On this view, the hierarchy-violating rankings would never be generated. Or, we could allow the all factorial rerankings to be generated and then filter out any that violate the hierarchies. With respect to acquisition, this means that when the language learner is reranking her constraints in response to language data, in order to respect the hierarchies the illicit rankings must either not be generated or there must be a filtering of rankings that violate the hierarchies before each evaluation step in learning.\(^{35}\)

This property of hierarchies, that they are effectively meta-constraints on constraints, is a fundamental one and must be acknowledged by any OT analysis that uses them, no matter whether the analysis uses regular OT or stochastic OT. Thus, hierarchies add a mechanism to OT, but this is not particular to stochastic OT. Although we must be cautious in expanding our theories, the descriptive, explanatory, and typologically predictive value of hierarchies are clear, as most notably demonstrated by Aissen (1999). However, the use of hierarchies and the filtering out of hierarchy-violating constraint rankings — in the manner discussed here — has not been countenanced in the OT learnability literature. Their effect on learnability proofs has yet to be tested thoroughly. In the stochastic OT model I am assuming, there is at present no way to implement the filtering effect I have just outlined. The assumption I am making is that although all

\(^{35}\)Ida Toivonen (p.c.) has pointed out that the hierarchies are not necessarily part of the grammar. They may be functionally/cognitively motivated constraints, then they do not need to be expressed grammatically at all, but rather constrain the space of possible languages. In that case, no language that violated the constraints would be instantiated, simply because such languages are cognitively impossible. This is an interesting alternative and an important architectural issue, but in the present analysis the hierarchies are to be thought of as purely formal grammatical constraints, with no functional motivation. Thus, I am making the architectural assumption that hierarchies are simply another formal aspect of optimality-theoretic grammars.
of the hierarchically ranked constraints have been assigned the same ranking value, on any actual evaluation only one of the hierarchy-respecting rankings in (47) will be instantiated. This is precisely equivalent to Boersma’s stochastic OT model augmented with filtering for hierarchies, and gives a similar effect to Anttila’s (1997b) stratified partial orderings, but without the problems discussed in section 3.4.2.

With this in mind, (48) presents the constraint ordering again, but marked (with subscripts) for ranking values. There is no evidence for any optionality in the other constraints and they have therefore been assigned values that are far apart. This will have the result that any ordering other than the one stated is extremely unlikely to occur.

(48) **Stochastic Marathi ranking**

\[
0.10 \text{THEMATIC}_{subj} \gg 0.90 \text{MAX}(\text{proto}) \gg 0.80 \ast \text{Nonsubj/P-A}_{vol} \gg 0.70 \text{DEP}(\text{proto}) \gg 0.60 \ast \text{Subj/P-P} \gg 0.60 \ast \text{Nonsubj/P-A}_{-vol} \gg 0.60 \ast \text{Nonsubj/P-P} \gg 0.60 \ast \text{Subj/P-A}_{-vol}
\]

The probability of any of the constraints separated by a ranking difference of 10 switching at evaluation time is 0.0002, (assuming a noise factor of 2). However, the probability of any two of the final four constraints switching is 0.5; so, any of the six rankings in (47) is equally likely to occur.

Now that the constraints are in place, the next step is to characterize the form of the inputs and the candidates. I follow Bresnan (to appear) and Kuhn (2001) in assuming the input to be a basic semantic form, including argument structure information. However, I also assume that argument structure is a separate level of representation within LFG grammars, called a-structure, and that it is related to c-structure by the \( \alpha \) projection function and to f-structure by the \( \lambda \) projection function (Butt et al., 1997). Since proto-properties pertain to argument structure, they are represented at the a-structure level and are also part of the input. Thus, the inputs I’m assuming are pairs of f-structures and a-structures. As an illustration, the input for the sentence in (49) would be the pair in (50), where the first member is an f-structure and the second is an a-structure.

(49) **tyaa mulaa-laate aaiskrim aawadte.**

That boy likes that ice cream.

(50) \[
\begin{aligned}
&\text{PRED} & \text{‘aawadte’} \\
&\text{GF}_1 & \begin{cases}
\text{PRED} & \text{‘mulaa’} \\
\text{PERS} & 3 \\
\text{NUM} & \text{SG} \\
\text{GENDEIXIS} & \text{MASC} \\
\end{cases} \\
&\text{GF}_2 & \begin{cases}
\text{PRED} & \text{‘aaiskrim’} \\
\text{PERS} & 3 \\
\text{NUM} & \text{SG} \\
\text{DEIXIS} & + \\
\end{cases} \\
\text{TNS} & \text{PRESENT} \\
\end{aligned}
\begin{aligned}
\text{REL} & \text{aawadte} \\
&\text{ARG}_1 & \begin{cases}
\text{PROTO-AGENT} & \{\text{PAb, PAe}\} \\
\text{PROTO-PATIENT} & \{\text{PPa}\} \\
\end{cases} \\
&\text{ARG}_2 & \begin{cases}
\text{PROTO-AGENT} & \{\text{PAe}\} \\
\text{PROTO-PATIENT} & \{\} \\
\end{cases}
\end{aligned}
\]

I have abbreviated the proto-role properties as PAa–e (i.e. proto-agent properties a to e) and PPa–e, where the letters correspond to the properties listed in section 5, (34) and (35), save for the proto-patient property that Joshi added (PPf), which I rejected.

This kind of input is cumbersome in tableaux. So I will instead use inputs that encode the same proto-role information in a more concise manner:

(51) \[
\langle \text{arg1[PAb,PAe,PPa]} \text{ arg2[PAe]} \rangle
\]

\(^{36}\)Recall from section 3.4.3 that there is a random variable with a Gaussian distribution of values associated with each constraint ranking and that this yields a slightly different ranking at each evaluation.
This abbreviated format represents only the part of the input that is relevant for my analysis. The correspondence should be clear on inspection. Although I’m assuming an unordered argument structure, I will number the arguments so that which grammatical functions they have mapped to will be more obvious.

The inputs I use generally contain the proto-role information attributed to them in Joshi (1993). However, she does not attribute the full set of proto-properties to predicates in all cases, whereas I have endeavoured to include all of a predicate’s arguments’ proto-properties. Also, I reject the proto-patient property PPF, ‘not entailed to be sentient’. Thus, my inputs are in certain respects richer than Joshi’s, but only include the necessary proto-role entailments of the predicates in question, as motivated in Dowty (1991). Appendix A gives the proto-properties for each of the alternating verbs listed by Joshi (1993).

The candidates and outputs will be triples of c-structures, f-structures and a-structures. Evidently, this will consist of fixing the underspecified GF feature to a fully specified grammatical function, SUBJ and OBJ in particular. Since I am using the architecture proposed by Butt et al. (1997), this essentially amounts to specifying the λ projection function, which is the linking function.

Once again, though, these large f-structure and a-structure pairs are too cumbersome for tableaux, so I will abbreviate the relevant parts of the candidates like so:

(52) \[
\begin{array}{c}
\text{SUBJ} \quad \text{arg1[PAb, PAe, PPa]} \\
\text{OBJ} \quad \text{arg2[PAe]}
\end{array}
\]

In this example, one argument has been linked to the subject function, while the other has been linked to the object function. The necessary constraints and ranking are in place now, and I have characterized the form of the inputs and candidates.

### 6.1 Optionality in Marathi Linking

First, I will show how these constraints capture the subject/object alternation in non-volitional transitives. In the tableaux, I have suppressed the constraint ranking values. Instead, I have represented the fact that *Subj/P-P, *Subj/P-A<vol>, *Nonsubj/P-A<vol>, and *Nonsubj/P-P can alternate with numerals above these constraints. These numerals are then used to indicate which constraint ranking results in which optimal candidate. For example, the notation (1 \(\gg\) 2 \(\gg\) 3 \(\gg\) 4) next to the pointing finger is shorthand for (*Subj/P-P \(\gg\) *Subj/P-A<vol> \(\gg\) *Nonsubj/P-A<vol> \(\gg\) *Nonsubj/P-P).

The first case to consider is that of verbs in the ‘like’ class and the first two verbs in the ‘find’ class (see appendix A). These verbs all have one argument that is entailed to be sentient (PAb), exists independently of the event (PAe), and undergoes a change of state (PPa). The second argument is entailed to exist independently of the event (PAe) and has no other proto-properties.

---

37 Bresnan (to appear) and Kuhn (2001) propose that the candidate set in both production and comprehension directed optimization is actually a set of c-structure, f-structure pairs (i.e. \(\{x, y\} \mid x\) is a c-structure and \(y\) is an f-structure). Thus the candidates and outputs assumed here are triples, since they also contain a-structures. However, we can ignore the c-structures in the production grammar, since the linking optionality only has to do with grammatical function information in the f-structure. In section 6.3, which deals with ambiguity in the comprehension grammar, the c-structures will be discussed further.
Non-volitional transitives 1

‘like’ class
‘find’ class: saapadpe (to find)
‘find’ class: aadpo (to come across)

Due to the high-ranking THEMATIC\textsubscript{subj} constraint, candidates (a) and (b) are knocked out immediately, since they have subjects with no associated proto-properties. This also results in several MAX\textsubscript{proto} violations, as these candidates do not realize proto-role information from the input. Candidate (c) does have a subject with a proto-property, thus not violating THEMATIC\textsubscript{subj}, but it is eliminated by MAX\textsubscript{proto}. In fact, any candidate that does not realize all of the proto-role information in the input will be eliminated in this manner. Therefore, in subsequent tableaux I do not show such candidates. Similarly, candidate (d) has a fatal violation of DEP\textsubscript{proto}, since the proto-property PPb is present on its subject, but is not present in the input. I will also suppress further DEP\textsubscript{proto} violators.

This leaves candidates (e) and (f). In this tableau and any other tableaux with 1 \gg 3, candidate (f) is the winner. However, on any of the three other legitimate rankings of the optional constraints (where 3 \gg 1), candidate (e) will be the winner. Thus, these constraints yield subject/object alternation for non-volitional transitives, since the only difference between the candidates is in which argument gets mapped to which grammatical function.

It should be noted that nothing in this analysis distinguishes between objects and obliques, which is another important facet of linking in Marathi, since it is a symmetric language. I assume that separate case distribution principles will make the necessary distinctions, as obliques are always marked with oblique case or with a postposition (Pandharipande 1997; Joshi 1993; see section C, example (2c)).

In fact, as suggested by Joan Bresnan (p.c.), in order to capture various generalizations about subjects and objects ranking higher than obliques according to certain criteria, it may be necessary to assume a prominence scale such as Core > Oblique, in addition to the Subject > Nonsubject scale assumed here. We could align the proposed new scale with the proto-role scale as well, thus enabling us to make statements about the realizations of thematic information in obliques. The resulting constraints, along with the constraints for stating the case facts mentioned above, could then be used in accounting for the object symmetry property, as well as possibly related facts, such as the causee in an active causative being realized as an oblique.

A third alternative is to incorporate the OT constraints proposed by Alsina (2000) for distinguishing first and second objects in asymmetric
The next case I consider is that of the final two ‘find’ class verbs, *milìpe* (to get) and *laab³pe* (to come to possess). These verbs share the same proto-properties as the other ‘find’ verbs, but the second argument is in addition entailed to undergo a change of location (PPa).

(54) **Non-volitional transitives 2**

<table>
<thead>
<tr>
<th>'find' class: milìpe (to get)</th>
<th>'find' class: laab³pe (to come to possess)</th>
</tr>
</thead>
</table>

The addition of a single proto-property has resulted in a drastic change in distribution. Now the first candidate is selected on five out of six rankings, while the second candidate is selected only on the ranking 1 \(\gg\) 2 \(\gg\) 3 \(\gg\) 4. Foreshadowing the discussion in section 7 somewhat, there is a difference between candidates (a) and (b) in terms of proto-properties that explains why candidate (a) is favoured. Note that arg2 has one proto-agent property and one proto-patient property; so, it is just as good a subject as it is an object. However, arg1 has two proto-agent properties and one proto-patient property. This means arg1 is a better subject than object. Candidate (a) has arg1 mapped to **SUBJ** and arg2 mapped to **OBJ**, which is what we would expect according to proto-role theory, as arg1 is a better subject, while arg2 is as good an object as it is a subject.

Next I turn to the first set of passive ditransitives. These verbs all have one argument (arg2) that is entailed to be sentient (PAb), exists independently of the event (PAe), undergoes a change of state (PPa), is causally affected by the demoted first argument (PPc), and is stationary relative to the movement of the third argument (PPd). The other argument (arg3) is entailed to exist independently of the event (PAe), moves relative to the second argument (PAd), undergoes a change of location (PPa), and is causally affected by the demoted first argument (PPc).

languages and for distinguishing objects from obliques in symmetric and asymmetric languages.
Passives of ditransitives 1

depe (to give)
paał^{	ext{a}}\text{awpe} (to send)
b^{	ext{a}}\text{arawpe} (to feed)
wíkpe (to sell)

Candidate (a) is selected on only one out of six constraint rankings, and candidate (b) is selected on the other five rankings. Again, the candidate that is favoured is the one whose mappings respect the predictions of proto-role theory. Arg3 has two proto-agent properties and two proto-patient properties, while arg2 has three proto-patient properties and only two proto-agent properties. This means that arg2 makes a better object than subject, while arg3 is as good a subject as it is an object. Candidate (b) maps arg2 to OBJ and arg3 to SUBJ.

The second set of passive ditransitives, presented in tableau (56), consists of the semantically similar verbs linhipe (to write (a letter)) and saangpe (to tell). These verbs have one argument (arg3) that is causally affected by the demoted first argument, which is the writer or teller (PPc), and that does not exist independently of the event (PPe), as the letter or statement comes into existence through being written or told. The other argument (arg2) is entailed to be sentient (PAb), exists independently of the event (PAe), undergoes a change of state (PPa), and is causally affected by the demoted first argument (PPc).
Unlike the other passive ditransitives, arg2 is not stationary relative to the movement of arg3. Therefore, arg2 lacks the proto-patient property PPd. This means that arg2 for these passive ditransitives has two proto-agent properties and two proto-patient properties, meaning there is nothing to distinguish it as an object versus a subject. On the other hand, arg3 only has proto-patient properties, making it a good object but a bad subject. We see again that candidate (a) which is preferred on five out of six rankings, is the one where proto-role theory is respected and arg2 is SUBJ while arg3 is OBJ.

This leaves us only with the passives of ditransitive causatives to consider. These are again broken into two sets, according to their proto-properties. The first set has one argument (arg2) that is entailed to be sentient (PAb), exists independently of the event (PAe), undergoes a change of state (PPa), and is causally affected by the demoted causer (PPc). The other argument (arg3) is entailed to exist independently of the event (PAe), but bears no other proto-properties.
Notice that the input in this tableau is similar to the one for non-volitional transitives 1 in (53), except that the first non-suppressed argument (arg1 in (53), arg2 in (57)) has one more proto-patient property in (57). In both cases the final argument makes a better subject than object, as its only proto-property is PAe. And the first (non-suppressed) argument is the best object, as it is the only argument with any proto-patient properties. But, the first argument also has more proto-agent properties than the final argument does. So we have a situation in which the same argument is both the best subject and the best object. Since the grammar attempt to assign the same argument to both subject and object, each optimal candidate wins in half the cases, with one winning when 3 ∋ 1 and the other winning when 1 ∋ 3.

Lastly, let us consider the final passive ditransitive causative, $paa\ddot{z}pe$ (to feed), which has slightly different proto-properties from the causatives in (57). Arg2 has the same entailments, but arg3 has additional entailments beyond that of existing independently (PAe). The other entailments are movement relative to the position of arg2 (PAd), undergoing change of location (PPa), being the incremental theme (PPb), and being causally affected by arg2.

### Passives of ditransitive causatives 2

<table>
<thead>
<tr>
<th>$\text{paa}\ddot{z}pe$ (to feed)</th>
</tr>
</thead>
</table>
| \[
\{Ø arg2[PAb, PAe, PPa, PPc] \arg3[PAd, PAe, PPa, PPb, PPc]\}
| **Tableaux (53) to (58)** show that the OT grammar outlined here does indeed get the subject/object alternation in all three relevant cases. Thus, fact (41b) has been accounted for: when there is no volitional argument there is subject/object alternation. The analysis also accounts for fact (41a), which states that if there is a volitional argument, it must be the subject. This is where *Nonsubj/P-A vol — the top ranked constraint on the nonsubject grammatical function/proto-role alignment — comes in.\footnote{In this tableau, I have glossed over all of the proto-roles except that of volitionality, representing them with ellipses (...). Similarly, I have suppressed DEP(proto) and the variable ranking constraints, as these play no role in selecting the winner.}
Candidate (a) shows that the property of volitionality, if present in the input, must be realized in the output, or else offending candidates are eliminated by MAX(proto). Candidates (b) and (c) show that once the property of volitionality is realized, it must be realized on the subject, as realizing it on the object results in a violation of *Nonsubj/P-A₁. In cases where the input has two arguments marked as volitional, such as in the verb 'kheşpe' (to play with), the object will be marked for volitionality, even though this violates *Nonsubj/P-A₁, because not realizing the property would result in a violation of MAX(proto), which is worse (see tableau (2a) in appendix C).

Thus, this analysis correctly accounts for the fact that subjects in Marathi are volitional, and that in cases where there is no volitional argument, there will be a choice in linking arguments to grammatical functions. In section 7, I compare this approach to Joshi’s. But first, let us consider how this grammar deals with ambiguity in the comprehension direction.

### 6.3 Ambiguity

In section 2, I argued that optionality and ambiguity in OT can be viewed as opposite sides of the same coin. If we view the difference as one between production and comprehension directed optimization, then the inputs of each kind of optimization correspond to the outputs of the other (Smolensky, 1996c, 1998). Thus, both optionality and ambiguity could be viewed as one input to the optimization having multiple outputs, with the nature of inputs and outputs varying in the appropriate way.

Recall from section 2 that one way of handling ineffability in OT is through the use of bidirectional optimization. If we assume Smolensky’s (1998) formulation of bidirectional optimization, then the inputs to production are semantic forms (like the f-structure/a-structure pair in (50)) and the inputs to comprehension are strings. The constraints and their ranking are identical in both production and comprehension, though.

From the string inputs to the comprehension directed optimization, the function Gen will yield the candidate set of (semantic) f-structures. However, we want to make sure that only those f-structures that match the input string are generated. As mentioned briefly in section 6 (see footnote 37), a standard assumption in OT-LFG is that the candidate set in both production and comprehension directed optimization is a set of c-structure and f-structure pairs (Bresnan, to appear; Kuhn, 2001) and I have extended this to include a-structures. In comprehension, we need to consider the string yield of the c-structure members of these triples. Only those triples whose c-structure yields match the input string are generated (Kuhn, 2001).

As motivated in Kuhn (2001) and Johnson (1998), consideration of the string’s yield results in an asymmetry between generation and parsing which means further steps have to be taken for parsing. In generation, an input to Gen gives candidates with all kinds of (ungrammatical) strings as the yields of their c-structure members. Kuhn (2001)

---

<table>
<thead>
<tr>
<th>(arg₁[PAa, . . .] arg₂[. . .])</th>
<th>THEMATIC (arg₂)</th>
<th>MAX(proto)</th>
<th>*Nonsubj/P-A₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. SUBJ [arg₂[. . .]] OBJ [arg₁[. . .]]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. SUBJ [arg₂[. . .]] OBJ [arg₁[PAA, . . .]]</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. SUBJ [arg₁[PAA, . . .]] OBJ [arg₂[. . .]]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
gives the example of the input for a wh-question having as one of the generated strings *She do read what*. Now, if we were to try and parse this string, every candidate in the candidate set will have this string as its yield. So no matter which candidate is optimal, the candidate should actually be ungrammatical and should not be generated.

Kuhn (2001) therefore proposes a filtering function for the resulting optimal candidates of comprehension to get back the kind of input f-structures that are used in generation (i.e. semantic forms). Then a backwards generation step is executed and only the optimal candidate(s) with a string yield that matches the yield of the string being parsed are selected. If no such candidate exists (as in the example), the input string to comprehension is rejected as ungrammatical. This is essentially bidirectional optimization.

6.3.1 Ambiguity in Marathi Linking

The Marathi data that I have been examining does present a case of ambiguity: the strings that display optionality in linking (i.e. non-volitional transitives and passive ditransitives) are ambiguous regarding which noun phrase in the string is assigned to the subject role and which is assigned to the object role. This can be viewed as a kind of structural or parsing ambiguity.\(^{40}\)

Let us now observe how the constraint system developed for non-volitional transitives in production works in comprehension.

\(^{40}\)At first this may not seem like a true semantic ambiguity, since the meaning is constant. But, once we embed these ambiguous sentences in broader contexts which show discourse effects, we should be able to observe the semantic effects of this ambiguity. For example, *SUBJ* is a discourse function, whereas *OBJ* is not. Therefore, any context which is sensitive to discourse functions will be ambiguous if it involves one of the strings under consideration here, as either NP could be the subject discourse function.
Non-volitional transitives

(60)

```
Non-volitional transitives
```
The grammatical functions SUBJ and OBJ have been underspecified to GF₁ and GF₂. By the notational convention discussed on page 22, this is equivalent to:

```
(62) \langle \text{arg1}[\text{PAa,P Ae,PPa}] \text{ arg2}[\text{PAe}] \rangle
```

The reader can see by referring to tableau (53) that this is indeed the input for non-volitional transitives in the production grammar. Similar results can be shown for passive ditransitives and passive ditransitive causatives. I will not go through these other results here.

Thus, the OT grammar fragment presented here can account for both the syntactic production optionality in Marathi linking and the comprehension ambiguity for the same constructions. This shows that production optionality and comprehension ambiguity are formally the same in OT grammars. In other words, the system of continuously-ranked constraints that was used to capture optionality can also be used to capture ambiguity. In this paper, I’ve illustrated this by looking at the same domain. However, I see no principled reason why other types of ambiguity that do not have optionality in the production direction could not be handled in the same manner.

### 7 Comparing the OT Approach with Joshi’s Approach

Joshi’s analysis not only accounts for the subject/object alternation, it also accounts for the fact that Marathi is a symmetric language, and for several intriguing facts about logical subjects and logical objects, the latter of which is Joshi’s own innovation. The analysis developed here only accounts for the alternation facts. With respect to just these facts, there are five major differences between my analysis and Joshi’s and I will argue that these provide reasons for preferring this analysis. I discuss the first four points in this section and the fifth in section 7.1.

One major difference is that the OT analysis uses a small number of crosslinguistically viable constraints with an ordering specific to Marathi, whereas Joshi’s syntactic feature rules (in (36)) and function mapping principles (in (37)) are part of a parametric approach to language variation, with the relevant parameters set for Marathi. It may seem that the two approaches are more or less equivalent, but while it is not obvious from Joshi’s exposition what the relevant dimensions of variation between languages might be, the OT approach makes specific claims (through constraint reordering) about language variation. Of course, this is not a specific problem with Joshi’s analysis, but rather an instance of the problem of relating parameters across languages and deciding what the limits of the space of parametric variation are.

The second difference is that in the OT analysis the alternation between subject and object linking and the fact that volitional arguments are realized as subjects are both derived from the alignment of the grammatical function and proto-role prominence scales. This alignment yielded the automatic consequence that volitional arguments are subjects, without ever stipulating it directly. However, Joshi’s LFG account does stipulate this property directly (in (37a)).
Third, Joshi’s analysis uses proto-patient properties as criteria of subjecthood. Recall the generalizations in (40) from section 5:

(40)  
   a. If a predicate has a volitional argument, that argument is realized as the subject; otherwise any argument of the verb that possesses a proto-patient property can be the grammatical subject. (Joshi, 1993: 92)
   b. If there is no volitional argument and there is more than one argument that has a proto-patient property there will be subject/object alternation.

Once the proto-agent property of volitionality does not apply, subject choice is entirely dependent on proto-patient properties. But proto-patient properties were proposed for object selection, not subject selection (Dowty, 1991). Joshi’s proposal therefore runs contrary to the theory that she presupposes. The OT analysis rejects the proposal that proto-patient properties determine subject choice and thus maintains proto-role theory’s original insights.

The fourth difference is that Joshi proposes a new proto-patient property, “not entailed to be sentient” (PPf), which I do not adopt. As discussed in section 5, there are two problems with this proto-property. First, it would fit better with proto-role theory if this property was “entailed to be non-sentient”, which is a stronger entailment and in direct opposition to the proto-agent property of sentience (PAb). Many of the proto-agent and proto-patient properties are in direct opposition to each other (Dowty, 1991), in particular PAc-d and PPe-d, and this would be another such case. This brings us to the second problem: any argument of a verb that lacks the proto-agent property of sentience will automatically have the property of not being entailed to be sentient. The proto-patient property that Joshi proposes is directly entailed by this absence and is superfluous, because its presence is completely predictable. Note that this contrasts with PAc-d and PPe-d, because these are such that if, for example, one argument has the entailment of the relevant PA property, then a different argument gets the opposing entailment. Thus, the new property does not fit into proto-role theory and its motivation is suspect.

7.1 Implications for Proto-Role Theory

The fifth difference between my analysis and Joshi’s also has to do with Dowty’s proto-role theory and has interesting implications for that theory. Originally, the idea behind Dowty’s analysis was that the argument with more proto-agent properties gets linked to the subject function, whereas the argument with more proto-patient properties gets linked to the object function. Any kind of alternation in linking would be predicted to only occur if there is a tie on the number of relevant properties. But, in both Joshi’s analysis and this one, this is no longer strictly true. If it were true, only the dative-marked argument in non-volitional transitives (e.g. the ‘finder’ in the ‘find’ class), could be realized as the subject, since this argument has two proto-agent properties, whereas the other argument only has one (see tableau (54)).

In Joshi’s analysis, this feature of Dowty’s theory is simply given up. However, in the analysis presented here it is still preserved, but only as a tendency, rather than as an absolute property. Looking back at tableau (54) for the ‘find’ class of non-volitional transitives, one alternate is picked five out of six times, whereas the other one is picked only one in six times. The striking fact is that the alternate that is picked most often is the one where the subject has more proto-agent properties than the object. Similarly, the opposing tendency of objects to have more proto-patient properties is illustrated in tableau (55). In that tableau, one candidate is again realized five times out of six, whereas the other is realized only once out of every six times. The candidate that is picked most often is the one in which the object, rather than the subject, has the most proto-patient properties. Thus, the argument with the most proto-agent

---

43How we interpret these predicted frequencies is an open question. It seems unrealistic to interpret them as raw frequencies. I believe they should instead be interpreted as encoding the strength of the tendency in question.

44The tendency is purely due to the difference in the number of proto-agent properties, since each argument in the winning candidates in (54) has exactly one proto-patient property.

45In the same way that proto-patient properties are kept constant across the two arguments in (54), each argument of the winning candidates in (55) has two proto-agent property. The tendency in argument realization noted in this case is therefore entirely due to the differing number of proto-patient properties.
properties is preferentially but not obligatorily linked to the subject, and the argument with the most proto-patient properties is again only preferentially linked to the object.

Another consequence of this analysis is that the optionality in linking is predicted to disappear if there is an asymmetry between arguments on both the number of proto-agent properties and the number of proto-patient properties. That is, if the first argument has more proto-agent properties than the second and the second argument has more proto-patient properties, then the first argument must be linked to the subject and the second argument must be linked to the object, as would be the case in Dowty’s original theory. This can be demonstrated easily by using similar inputs to the ones that yielded optionality. I will show that adding a proto-patient property to the second argument in the ‘find’ class of non-volitional transitive would remove the optionality.

According to the hierarchies in (43b), *Subj/P-P must outrank *Subj/P-A_{vol} and *Nonsubj/P-A_{vol} must outrank *Nonsubj/P-P. Therefore, either *Subj/P-P or *Nonsubj/P-A_{vol} must be the highest ranking constraint at evaluation time. Whenever *Subj/P-P is ranked highest, candidate (b) wins. This is the candidate which has the argument with most proto-patient properties linked to the object and the one with most proto-agent properties linked to the subject. And this candidate also wins whenever *Nonsubj/P-A_{vol} is ranked highest, because it is also the candidate in which the object has the least proto-agent properties. Therefore, there is no optionality: the argument with the most proto-agent properties is linked to the subject and the argument with the most proto-patient properties is linked to the object.

**8 Conclusion**

In this paper, I have shown how continuously-ranked constraints in a stochastic model of OT can be used to describe optionality and ambiguity in syntax. I argued that this method was preferable to other approaches to optionality proposed in the OT literature. Finally, I applied the method to describing a specific case of syntactic optionality and ambiguity: argument linking in certain Marathi predicates.

The analysis developed here builds on that of Joshi (1993) in using proto-role information. However, while Joshi has to give up certain facts about proto-roles and grammatical function selection originally proposed by Dowty (1991), this analysis preserves those insights as statistical tendencies. Furthermore, this analysis uses crosslinguistically viable constraints and only a simple mechanism of constraint alignment. Thus, the OT analysis of subject/object alternation in Marathi has three attractive features. First, it uses a small number of crosslinguistically viable constraints. Second,
one mechanism — harmonic alignment — is used to account for both the subject/object alternation and the fact that volitional arguments must be subjects. Third, and most interestingly, this analysis preserves as a tendency Dowty’s observation that the argument with most proto-agent properties is realized as the subject and that the argument with the most proto-patient properties is realized as the object.

I have shown that it is possible to represent linking using a small set of general, violable constraints. In addition, I have illustrated how optionality can be handled in OT syntax. I have also shown how ambiguity in comprehension grammars can be modelled using The same methods. This gives some indication that Optimality Theory can in fact deal with ineffability, optionality and ambiguity and is a viable syntactic theory.
Appendices

A Verb Classes

<table>
<thead>
<tr>
<th>Verbs</th>
<th>ARG1</th>
<th>Proto-Properties</th>
<th>ARG2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-volitional transitives (‘like’ class)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aawadn</td>
<td>to like</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>samadzne</td>
<td>to realize</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>umagnne</td>
<td>to understand</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>disene</td>
<td>to notice</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>purge</td>
<td>to accept (mentally)</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>ruste</td>
<td>to agree (aesthetically)</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>maanawne</td>
<td>to suit (mentally, medically)</td>
<td>[PAb, PAe, (PPa)]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>kalne</td>
<td>to hear/learn</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>Non-volitional transitives (‘find’ class)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>saapadn</td>
<td>to find</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>aaq8alne</td>
<td>to come across</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>milne</td>
<td>to get</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe, PPa]</td>
</tr>
<tr>
<td>laabhnge</td>
<td>to come to possess</td>
<td>[PAb, PAe, PPa]</td>
<td>[PAe, PPa]</td>
</tr>
<tr>
<td>Passive ditransitives</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dene</td>
<td>to give</td>
<td>[PAb, PAe, PPa, PPc, PPd]</td>
<td>[PAd, PAe, PPa, PPc]</td>
</tr>
<tr>
<td>paa8awne</td>
<td>to send</td>
<td>[PAb, PAe, PPa, PPc, PPd]</td>
<td>[PAd, PAe, PPa, PPc]</td>
</tr>
<tr>
<td>b8arawne</td>
<td>to feed</td>
<td>[PAb, PAe, PPa, PPc, PPd]</td>
<td>[PAd, PAe, PPa, PPc]</td>
</tr>
<tr>
<td>wiknne</td>
<td>to sell</td>
<td>[PAb, PAe, PPa, PPc, PPd]</td>
<td>[PAd, PAe, PPa, PPc]</td>
</tr>
<tr>
<td>linhine</td>
<td>to write (letter)</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PPa, PPc]</td>
</tr>
<tr>
<td>saangne</td>
<td>to tell</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PPa, PPc]</td>
</tr>
<tr>
<td>Passive ditransitive causatives (ingestives)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aikawne (aike)</td>
<td>to hear, CAUS (to hear)</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>hungawne (hungne)</td>
<td>to smell, CAUS (to smell)</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>ˇsikawne (ˇskne)</td>
<td>to teach (to learn)</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PAe]</td>
</tr>
<tr>
<td>paadzne (pine)</td>
<td>to feed (to drink)</td>
<td>[PAb, PAe, PPa, PPc]</td>
<td>[PAd, PAe, PPa, PPb, PPc]</td>
</tr>
</tbody>
</table>

B Marathi Glosses

D/A = dative/accusative case  
ERG = ergative case  
GEN = genitive case  
ABIL = ability case  
PART = participle  
CAUS = causative  
ABLE = ability verbal morpheme  

(Joshi, 1993: viii)

48 The first two columns of this table are taken from Joshi (1993: 52–60) and are not meant to be exhaustive. The last two columns are the proto-properties I have assumed.

49 The basic verb from which the causative is formed is given in parentheses.

50 See Joshi (1993: 26–30) for arguments for collapsing these two cases into one.
### C Ranking Arguments

(1) a. \( \text{THEMATIC}_{\text{subj}} \gg \text{MAX}(\text{proto}) \)

<table>
<thead>
<tr>
<th>( \langle \text{arg}1[\text{PAa,PAe,PPa,PPf}] \rangle )</th>
<th>( \text{MAX}(\text{proto}) )</th>
<th>( \text{THEMATIC}_{\text{subj}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. OBJ [( \text{arg}1[\text{PAa,PAe,PPa,PPf}] )]</td>
<td>( *! )</td>
<td></td>
</tr>
<tr>
<td>b. ( * ) SUBJ [( \text{arg}1[\text{PAa,PAe,PPa}] )]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. output = ‘gelaa’ (went)

c. raam pail-tiraa-laa gelaa.
Ram other-bank-LOC went
*Ram went to the other bank (of river).*
(Joshi, 1993: 24, (30a))

(2) a. \( \text{MAX}(\text{proto}) \gg *\text{Nonsubj/P-Avol} \)

<table>
<thead>
<tr>
<th>( \langle \text{arg}1[\text{PAa,PAb,PAe}] \text{ arg2}[\text{PAa,PAb,PAe}] \rangle )</th>
<th>( \text{MAX}(\text{proto}) )</th>
<th>( *\text{Nonsubj/P-Avol} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. SUBJ [( \text{arg}1[\text{PAa,PAb,PAe}] )] OBL [( \text{arg2}[\text{PAa,PAe}] )]</td>
<td>( *! )</td>
<td></td>
</tr>
<tr>
<td>b. ( * ) SUBJ [( \text{arg}1[\text{PAa,PAb,PAe}] )] OBL [( \text{arg2}[\text{PAa,PAb,PAe}] )]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. output = ‘khe\(\text{\'l}\)ne’\(^{51}\) (to play with)

c. sudh\(\text{\'a}\) an\(\text{\'u}\)\(\text{\'s}\)/barobar khe\(\text{\'l}\)ne
Sudha.inst.3sf/with play.pres.3sf
*Sudha plays with Anu.*\(^{52}\)
(Pandharipande, 1997: 295, (847))

\(^{51}\)Also bh\(\text{\'a}\)\(\text{\'n}\)ne (to fight with), and bol\(\text{\'n}\)e (to talk with) (Pandharipande, 1997: 295).
\(^{52}\)This example is from Pandharipande (1997) and uses her transcription and gloss styles.
(3)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a.</strong></td>
<td>*Nonsubj/P-A&lt;sub&gt;vol&lt;/sub&gt; ⇒ DEP(proto)</td>
<td>MAX(proto)</td>
</tr>
<tr>
<td></td>
<td>arg1[PAb,PAe,PPa], arg2[PAa,PAe]</td>
<td></td>
</tr>
<tr>
<td>a.</td>
<td>SUBJ [arg1[PAb,PAe,PPa]]</td>
<td>OBJ [arg2[PAa]]</td>
</tr>
<tr>
<td>b.</td>
<td><strong>SUBJ</strong> [arg1[PAb,PAe,PPa]]</td>
<td>OBJ [arg2[PAe,PPf]]</td>
</tr>
<tr>
<td>b. output = 'aawadñe' (to like)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. tyaa mulaa-la te aiskrim aawadte.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*That boy likes that ice cream.*
(Joshi, 1993: 55, (42a))

(4)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a.</strong></td>
<td>DEP(proto) ⇒ *Subj/P-P</td>
<td>MAX(proto)</td>
</tr>
<tr>
<td></td>
<td>arg1[PAb,PAe,PPa], arg2[PAe,PPa,PPb,PPf]</td>
<td></td>
</tr>
<tr>
<td>a.</td>
<td>SUBJ [arg1[PAb,PAe,PPa]]</td>
<td>OBJ [arg2[PAe,PPa,PPb,PPf]]</td>
</tr>
<tr>
<td>b.</td>
<td><strong>SUBJ</strong> [arg1[PAb,PAe,PPa]]</td>
<td>OBJ [arg2[PAe,PPa,PPb,PPf]]</td>
</tr>
<tr>
<td>b. output = 'aawadñe' (to like)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(5)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a.</strong></td>
<td>DEP(proto) ⇒ *Subj/P-A&lt;sub&gt;vol&lt;/sub&gt;</td>
<td>MAX(proto)</td>
</tr>
<tr>
<td></td>
<td>arg1[PAb,PAe,PPa], arg2[PAe,PPa,PPb,PPf]</td>
<td></td>
</tr>
<tr>
<td>a.</td>
<td>SUBJ [arg1[PPa]]</td>
<td>OBJ [arg2[PAe,PPa,PPb,PPc,PPf]]</td>
</tr>
<tr>
<td>b.</td>
<td><strong>SUBJ</strong> [arg1[PAb,PAe,PPa]]</td>
<td>OBJ [arg2[PAe,PPf]]</td>
</tr>
<tr>
<td>b. output = 'aawadñe' (to like)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\(6\)

a. \(\text{DEP(proto)} \gg *\text{Nonsubj/P-A}_{\text{not}}\)

\[
\langle \text{arg1}[\text{PAb},\text{PAe},\text{PPa},\text{PPb}] \; \text{arg2}[\text{PAe},\text{PPf}]\rangle
\]

<table>
<thead>
<tr>
<th></th>
<th>MAX(\text{proto})</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-A</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>SUBJ [arg1[\text{PAb},\text{PAe},\text{PPa},\text{PPb},\text{PPc}]]</td>
<td>OBJ [arg2[\text{PPf}]]</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. output = ‘aawad\text{\textprime}’ (to like)

\(7\)

a. \(\text{DEP(proto)} \gg *\text{Nonsubj/P-P}\)

\[
\langle \text{arg1}[\text{PAa},\text{PAb},\text{PAd},\text{PAe}] \; \text{arg2}[\text{PPf}]\rangle
\]

<table>
<thead>
<tr>
<th></th>
<th>MAX(\text{proto})</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
<th>DEP(\text{proto})</th>
<th>*Nonsubj/P-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>SUBJ [arg1[\text{PAa},\text{PAb},\text{PAd},\text{PAe},\text{PPa}]]</td>
<td>OBJ [arg2[[]]]</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b. output = ‘saat\text{\textprime}awge’ (to store/collect)

c. tyaa mulaa-ni kaahi naa\text{n\textprime}i saat\text{\textprime}awli.

That boy-\text{ERG} some coins stored

\text{That boy stored (collected) some coins.}

(Joshi, 1993: 195, (51a))

### D Citation Information for Marathi Examples

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(19a)</td>
<td>53 (36a)</td>
<td>(19b)</td>
<td>53 (37a)</td>
<td>(20)</td>
<td>60 (51b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(21)</td>
<td>60 (53b)</td>
<td>(22a)</td>
<td>38 (2a)</td>
<td>(22b)</td>
<td>38 (2b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(23a)</td>
<td>39 (4a)</td>
<td>(23b)</td>
<td>39 (4b)</td>
<td>(24a)</td>
<td>39 (5a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(24b)</td>
<td>39 (5b)</td>
<td>(25a)</td>
<td>54 (38b)</td>
<td>(25b)</td>
<td>54 (39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(26a)</td>
<td>54 (40b)</td>
<td>(26b)</td>
<td>55 (41b)</td>
<td>(27a)</td>
<td>61 (54b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(27b)</td>
<td>61 (55)</td>
<td>(28a)</td>
<td>62 (56b)</td>
<td>(28b)</td>
<td>62 (56c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(29a)</td>
<td>59 (48b)</td>
<td>(29b)</td>
<td>58 (48a)</td>
<td>(30a)</td>
<td>59 (49b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(30b)</td>
<td>59 (49a)</td>
<td>(31a)</td>
<td>64 (61a)</td>
<td>(31b)</td>
<td>65 (61b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(32a)</td>
<td>65 (62a)</td>
<td>(32b)</td>
<td>65 (62b)</td>
<td>(49)</td>
<td>55 (42a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
References


41
