Formal and Empirical Issues in Optimality Theoretic Syntax

edited by Peter Sells

ISBN: 1-57586-244-1

Contents

Contributors

Ash Asudeh: Department of Linguistics, Stanford University, Stanford, CA 94305-2150, USA, asudeh@csli.stanford.edu

George Aaron Broadwell: Department of Anthropology, SUNY Albany, Albany, NY 12222, USA, g.broadwell@albany.edu

Hye-Won Choi: Department of Modern Languages and Literatures, SUNY Buffalo, Buffalo, NY 14260-4620, USA, hwchoi@acsu.buffalo.edu

Anette Frank: DFKI, Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany, frank@dfki.de

Tracy Holloway King: NLTT/ISTL, Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA, thking@parc.xerox.com

Jonas Kuhn: Institut für maschinelle Sprachverarbeitung, Universität Stuttgart, Azenbergstraße 12, 70174 Stuttgart, Germany, jonas@ims.uni-stuttgart.de

Hanjung Lee: Department of Linguistics, Stanford University, Stanford, CA 94305-2150, USA, hanjung@csli.stanford.edu

John T. Maxwell III: NLTT/ISTL, Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, USA, maxwell@parc.xerox.com

Yukiko Morimoto: Department of Linguistics, California State University, Fresno, 5245 N. Backer Ave. PB 92, Fresno, CA 93740-8001, ymorimot@csufresno.edu

Peter Sells: Department of Linguistics, Stanford University, Stanford, CA 94305-2150, USA, sells@csli.stanford.edu

Devyani Sharma: Department of Linguistics, Stanford University, Stanford, CA 94305-2150, USA, sharma@csli.stanford.edu

3

Markedness and Word Order Freezing

Hanjung Lee

3.1 Introduction

Studies on word order variation in "free" word order languages fall into two major categories.[1] The dominant approach has been one in which "free" word order or scrambling is driven by grammatical features such as Case and Agr(eement) (e.g., Gurtu 1985; Webelhuth 1989; Mahajan 1990; Speas 1990; Saito 1992; Y. Lee 1993; Miyagawa 1997, among others). In this approach, scrambling is a consequence of Case/Agr-driven movement: a phrase is moved to a certain specifier position or adjoined to a specific functional category (e.g., IP) so that its Case can be licensed or its Agr feature can be checked off. Another approach has been suggested in which scrambling is an instance of semantically-driven movement (e.g., de Hoop 1996; Diesing 1992; Neeleman and Reinhart 1997, among others) or happens as a result of the interaction of syntax and discourse/pragmatics (e.g., King 1995; Costa 1998; Müller 1998; Samek-Lodovici 1998; Choi 1999, 2001,

[1]I am grateful to Joan Bresnan and Peter Sells, who made detailed comments and valuable suggestions on earlier versions of this paper. One of the main ideas developed here—word order freezing involves harmonic alignment of morphosyntactic prominence hierarchies—was first proposed to me by Peter Sells. Special thanks and no blame for my use of their input, comments and suggestions should go to Judith Aissen, Jane Grimshaw, Yukiko Morimoto, Elizabeth Traugott, two anonymous reviewers, the participants in the Stanford OT Syntax Research Group, members of the OT syntax seminar led by Judith Aissen and Joan Bresnan, Winter 2000, and particularly to Ashwini Deo, Devyani Sharma and Shiao Wei Tham. Versions of this paper were presented at the 8th International Symposium on Korean Linguistics (ISOKL 8), Harvard University, July 16–19, 1999, the 30th Conference of the North East Linguistics Society (NELS 30), Rutgers University, October 22–24, 1999, the Joint Stanford/UCSC Workshop on Optimal Typology, UCSC, October 30, 1999, and at the 33rd Annual Linguistics Conference, Language Research Institute, Seoul National University, Korea, December 10, 1999. I'd like to thank the audiences for helpful comments and discussion. This material is based upon work supported by the National Science Foundation under Grant No. BCS-9818077.

Formal and Empirical Issues in Optimality Theoretic Syntax.
Peter Sells (ed.).

among others). In this approach the varied word orders are optional from a purely syntactic point of view: they are motivated by semantic factors, such as specific vs. non-specific interpretations and discourse considerations, such as topic and focus.

In many free word order languages, it is not uncommon to find fixed word order phenomena: a certain canonical word order becomes fixed under special circumstances in which the relative prominence relations of different dimensions of linguistic substance—grammatical functions, semantic roles, case, and positions in phrase structure—do not match, or in which morphology is unable to distinguish the grammatical functions of the arguments. Despite their important implications for syntactic theory relating to the formal mechanisms for capturing word order freedom, fixed word order phenomena, referred to as *word order freezing* (Mohanan 1992; Mohanan and Mohanan 1994), have received relatively little attention in recent literature on word order. While the various approaches to scrambling suggested so far can generally account for the free ordering of constituents, none of the previous approaches explain word order freezing effects found in languages like Hindi and Korean, regardless of whether we assume that scrambling is motivated by Case/Agr or by semantic/pragmatic factors.

This paper presents an Optimality-Theoretic (OT: Prince and Smolensky 1993) account of word order in Hindi and Korean that can account for both the free ordering and fixed ordering of constituents. Specifically, the analysis is developed within the OT-LFG framework which embeds LFG's non-derivational system of correspondence between parallel structures within OT theory of constraint interaction (Bresnan 2000a,b, 2001a; Choi 1999; Kuhn 1999, 2001; Sells 1999, 2000, 2001). A theory like OT-LFG which sees simultaneous competition between parallel, co-present structures as pervasive and constraint ranking as the means to resolve it provides an appropriate framework in which to approach the problem of word order freezing.

This paper is organized as follows: Section 3.2 presents the facts on the two types of word order freezing found in Hindi and Korean, and discusses the problems posed by the data for previous approaches to word order variation. Section 3.3 examines basic clause structures in Hindi and Korean. Section 3.4 and 3.5 present the OT-LFG analysis of word order variation in Hindi and Korean that can account for both the free ordering and fixed ordering of constituents. In section 3.4, I show that the OT approach making crucial use of hierarchy alignment (Prince and Smolensky 1993; Aissen 1999) captures the basic generalization about 'the worst of the worst' type of freezing: highly marked argument types occur only in unmarked position. Then in section 3.5, I show that the word order freezing found in sentences with ambiguous case marking can be explained when the constraint system allows an extension to bidirectional competition (Smolensky 1996b, 1998; Wilson 1996, 1997, 2001). Section 3.6 concludes the paper.

3.2 Word Order Freezing in Hindi and Korean

This section presents data from Hindi and Korean illustrating word order freezing effects, and discusses the problems posed by these data for previous approaches to word order variation.

3.2.1 'The Worst of the Worst' Type of Freezing

In this section I will present basic word order facts about Hindi and Korean, and data demonstrating a loss of word order freedom in clauses which contain highly marked argument types.

Word Order Freezing in Nonvolitional Transitives in Hindi

Hindi is a right-headed language with SOV canonical order. However, unlike Japanese and Korean, the surface order of elements is not strictly head-final. The possible permutations of a simple Hindi sentence are shown in (1).[2] The three elements in a simple sentence can appear in any order.[3]

(1) a. Anuu-ne caand dek^{h}aa.
Anu-ERG moon-NOM see/look at-PERF
'Anu saw the moon.'
b. Caand Anuu-ne dek^{h}aa.
c. Anuu-ne dek^{h}aa caand.
d. Caand dek^{h}aa Anuu-ne.
e. Dek^{h}aa Anuu-ne caand.
f. Dek^{h}aa caand Anuu-ne.

(1a) reflects the 'basic', 'canonical' or 'unmarked' order. The other orders are deviations from this canonical order (Gambhir 1981; Mohanan 1992, 1994a; Mohanan and Mohanan 1994). Such deviations are used to mark a special information structure and are generally associated with shifts in prominence, emphasis and semantic effects (e.g., definiteness effects).

Despite a high level of word order freedom in this language, under certain circumstances, free word order freezes into a fixed, canonical order. For Hindi (Mohanan 1992; Mohanan and Mohanan 1994), and as I will show

[2]The abbreviations used in are as follows: ACC 'accusative', ADJ 'adjunct', ASP 'aspect', AUX 'auxiliary', BCK 'background information', CAUS 'causality', COMP 'complementizer', COMPL 'completive information', COP 'copular', DAT 'dative', DECL 'declarative ending', ERG 'ergative', FOC 'focus', FUT 'future', GEN 'genitive', HON 'honorific affix', INST 'instrumental', IO 'indirect object', LOC 'locative', NF 'nonfinite', NOM 'nominative', OBJ 'object', OBL 'oblique', OM 'object marker', PAST 'past tense', PERF 'perfective', PRES 'present', PRON 'pronoun', SENT 'sentience', SUBJ 'subject', TOP 'topic', VOL 'volitionality'.

[3]The canonical case for animate objects in Hindi is ACC, and the canonical case for inanimate objects is NOM. Verbs that are neutral to the animacy of their objects like *dek^{h}* 'see' can take either ACC or NOM depending on the animacy of their object (nominative case in Hindi has no phonological realization). This phenomenon of selective case marking on objects is a highly principled and widespread one, found in a variety of languages typologically remote from Hindi. See Aissen (2000) for a formal articulation of the idea that the higher in prominence an object is, the more likely it is to be overtly case marked.

in the next section, Korean, one environment for restricted word order variation occurs when a sentence contains highly marked types of subject and object (i.e. transitive verbs without prototypical agent-patient argument relations). The other environment will be described shortly in section 3.2.2. In Hindi, an example of a verb class that takes marked argument types is 'unaccusative transitives' or 'nonvolitional transitives' (Mohanan 1994a, section 7.1; Mohanan and Mohanan 1994, section 4). Of the two arguments of the nonvolitional transitives one is sentient and the other may be sentient or nonsentient. The obligatory sentient argument of these verbs always has dative case, as in (2) and (3).[4] While the ergative subject in (1) carries the meaning of volitional action, the dative-marked arguments in (2) and (3) are nonagentive and nonvolitional. Unlike the objects of volitional transitives, the theme arguments in (2) and (3) must be nominative even if animate: they cannot be accusative because they do not have the semantic property of being an entity toward which an action or event is directed by a volitional inceptor of the action or event (Mohanan (1994a, section 4.4) contains a more detailed discussion of the nominative arguments of 'unaccusative transitives' or 'nonvolitional transitives').

(2) Anuu-ko caand dik^{h}ii.
Anu-DAT moon-NOM appear-PERF
'Anu saw the moon.'
(Lit. 'To Anu the moon appeared/became visible.')

(3) Vijay-ko Ravii milaa.
Vijay-DAT Ravi-NOM find/encounter-PERF
'Vijay met Ravi unexpectedly.'

Unlike volitional transitives, nonvolitional transitives in Hindi are subject/object alternating verbs, i.e. either of the two arguments of the verb may be construed to be subject, the other the object.[5] Two syntactic tests for grammatical subjecthood in Hindi, namely the binding of the reflexive and subject obviation of the pronoun to take subject antecedents, can be used to demonstrate that either argument of nonvolitional transitives can be the grammatical subject. As argued in Mohanan (1994a:122–127), the Hindi reflexive takes either a grammatical subject or a logical subject (i.e. the thematically highest argument role) as its antecedent, and the Hindi pronoun cannot take the grammatical subject of its clause as its antecedent.

[4]The dative subject construction in Hindi has been studied in detail in Bahl (1967), Davidson (1969), Mohanan (1994a) and Verma and Mohanan (1990), among others. DAT case on the subject may be induced by any of three types of predicates. The first is a small set of 'nonvolitional transitives' as in (2) and (3). Belonging to the second type are noun+verb complex predicates. A third source of DAT case on the subject comes from modality meanings such as 'urge' and 'oblige', derived from complex verbals involving auxiliaries (Mohanan 1994a:142). Only the first type allows its arguments to alternate between subject and object.

[5]The Marathi counterparts of Hindi nonvolitional transitives also show this property; see Joshi (1993) and Asudeh (2001).

The sentences in (4) show that the nonvolitional experiencer argument *Anuu* is the grammatical subject: the reflexive *apnii* takes it as its antecedent (4a); the pronoun *uskii* cannot be coreferent with it (4b).[6]

(4) a. Anuu-ko Niinaa apnii bastii-mẽ dikʰii.
Anu-DAT Nina-NOM self-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in self's$_{i/*j}$ neighborhood.'

b. Anuu-ko Niinaa uskii bastii-mẽ dikʰii.
Anu-DAT Nina-NOM PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

In (5a), either the theme or the nonvolitional experiencer can be the eligible antecedent of the reflexive. Given the principle governing the interpretation of the reflexive in Hindi, it follows that the theme *Niinaa* is the grammatical subject in (5a) and that the experiencer *Anuu*, the logical subject in both (4) and (5a), is the grammatical subject in (4). The subject obviation test supports this conclusion; in (5b) the pronoun cannot be coreferent with the theme argument *Niinaa*. Therefore, it is the grammatical subject.

(5) a. Niinaa Anuu-ko apnii bastii-mẽ dikʰii.
Nina-NOM Anu-DAT self-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in self's$_{i/j}$ neighborhood.'

b. Niinaa Anuu-ko uskii bastii-mẽ dikʰii.
Nina-NOM Anu-DAT PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{i/*j}$ neighborhood.'

An interesting fact is that the two grammatical function analyses, shown in (4) and (5) above, are consistent only with certain restricted word orders. According to Mohanan and Mohanan (1994:175), the analysis of the nonvolitional experiencer as the grammatical subject in Hindi is required for all word orders except that in which the theme precedes the experiencer preverbally; this word order, shown in (5), can only receive the analysis of the theme as the grammatical subject, and in fact it is the only possible word order when the subject is a theme and the object is a nonvolitional experiencer.[7] This point is illustrated by the sentences in (6), which are acceptable only under the analysis of the experiencer as the subject. In the word orders shown in (6), the pronoun can be coreferent only with the theme because it is not the subject. This evidence suggests that the analysis of the theme as the subject is incompatible with the five orders

[6]Evidence for the objecthood of the nominative argument in (4) comes from the facts of gapping in Hindi. In order to be gapped, an element must be identical to the gapper in both grammatical function and case. The nominative argument in (4) can both gap as well as be gapped by uncontroversial grammatical objects. See Mohanan (1994a:142) for further details and examples.

[7]The passive of triadic predicates shows the same word order pattern: the order of the subject and object becomes fixed as SOV order when the subject is a theme and the object is a goal (Mohanan 1992; Mohanan and Mohanan 1994).

in (6) and that the order of the subject and object becomes fixed as SOV order in the marked linking pattern when the subject is a theme and the nonvolitional experiencer is the object.[8]

(6) a. Niinaa uskii bastii-mẽ dik^{h}ii Anuu-ko.
Nina-NOM PRON-GEN neighborhood-LOC appear-PERF Anu-DAT
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

b. Anuu-ko Niinaa uskii bastii-mẽ dik^{h}ii.
Anu-DAT Nina-NOM PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

c. Anuu-ko uskii bastii-mẽ dik^{h}ii Niinaa.
Anu-DAT PRON-GEN neighborhood-LOC appear-PERF Nina-NOM
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

d. Dik^{h}ii Niinaa Anuu-ko uskii bastii-mẽ.
appear-PERF Nina-NOM Anu-DAT PRON-GEN neighborhood-LOC
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

e. Dik^{h}ii Anuu-ko Niinaa uskii bastii-mẽ.
appear-PERF Anu-DAT Nina-NOM PRON-GEN neighborhood-LOC
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

The word order pattern in Hindi nonvolitional transitives discussed so far is summarized in (7). We see that all orders except the OSV order in (a2) are possible with the unmarked linking pattern shown in (7a) where the experiencer is the subject and the theme is the object. In contrast, the order of the subject and the object is frozen in unmarked order (i.e. SOV) in the marked linking pattern when the theme is the subject and the object is a nonvolitional experiencer.

(7) Table 1. Word order pattern in Hindi nonvolitional transitives

a.	SUBJ OBJ | | *exp* *th*	a1. $S_{exp}O_{th}V$ a2. *$O_{th}S_{exp}V$ a3. $S_{exp}VO_{th}$ a4. $O_{th}VS_{exp}$ a5. $VS_{exp}O_{th}S$ a6. $VO_{th}S_{exp}$
b.	SUBJ OBJ \ / *exp* *th*	b1. $S_{th}O_{exp}V$ b2. *$O_{exp}S_{th}V$ b3. *$S_{th}VO_{exp}$ b4. *$O_{exp}VS_{th}$ b5. *$VS_{th}O_{exp}$ b6. *$VO_{exp}S_{th}$

[8]Examples with the reflexive are not presented here because reflexive binding in Hindi does not refer to the grammatical subject exclusively, and both the theme and the experiencer can bind the reflexive in the word orders shown in (6).

As mentioned above, the arguments of Hindi nonvolitional transitives may be associated with two different grammatical function realizations. Interestingly, a close examination of the word order pattern summarized in (7) reveals that these two grammatical function realizations do not share the same surface string (strictly speaking, the precedence relation of the subject and the object). In other words, the string expressions of Hindi nonvolitional transitives is not ambiguous; they are associated with only one grammatical function structure. For example, the string *Niinaa Anuu-ko dik^{h}ii* receives an SOV analysis in (b1) but not the OSV one in (a2); similarly, the string *Anuu-ko Niinaa dik^{h}ii* is only grammatical for the SOV analysis in (a1) but not for the OSV one in (b2).

The word order pattern in Hindi nonvolitional transitives exhibits intriguing interactions of various types of prominences as well—prominences on the dimensions of grammatical function, thematic role and position in constituent structure. Crosslinguistically, the subject is canonically associated with the highest thematic role (e.g., agent or experiencer), and the object is canonically associated with a non-highest thematic role (e.g., patient or theme). Therefore, it is more marked for a subject to be a theme rather than an experiencer, and for an object to be an experiencer rather than a theme. The most marked situation obtains when the subject and the object are mapped onto marked thematic roles, *and* one (or both) of the subject and the object is in a marked position in the phrase structure. This situation is manifested in the five ungrammatical structures in (7b) in Table 1, where either a theme subject is not occurring in its unmarked clause-initial position as in (b2), (b4), (b5) and (b6), or an experiencer object is not in its unmarked preverbal position as in (b3), (b5) and (b6). In connection with the lack of ambiguity of the Hindi clauses with the nonvolitional experiencer and the theme, what is revealing about this case is the way the universal typological preference for the unmarked structure reveals itself in the Hindi data. As we will see in the analysis part, of the two grammatical function realizations shown in (7) above, the allowed structure is the less marked one in which the linking of arguments to grammatical functions and the relative order of the two arguments is also not marked.

The descriptive generalizations that emerge can be stated as follows:

(8) a. Generalization 1: **Avoidance of the Worst of the Worst**
Canonical SOV order becomes fixed in a Hindi nonvolitional transitive with a theme subject and an experiencer object. That is, such arguments should not be associated both with a marked semantic role and with a marked position.

b. Generalization 2: **Avoidance of Ambiguous Strings**
There is only one grammatical surface realization (i.e. linear order) for the alternative grammatical function realizations of the arguments of a nonvolitional transitive.

'The worst of the worst' (Smolensky 1995) generalization in (8a) is not easily captured by derivational approaches to word order variation. In the GB theory (Chomsky 1981; Lasnik and Saito 1992), free word order results from optional free adjunction and head movement (e.g., Gurtu 1985, Mahajan 1990, Speas 1990; Srivastav 1991, among others). The basic motivation for syntactic movement is based on the idea that word order at a more abstract underlying level of representation is the unmarked one, and that more marked surface word orders are derived from this by transformational derivations, which are encoded by means of indexed traces. In the Minimalist Program (Chomsky 1993, 1995), all syntactic operations are obligatory and take place only if driven by some independent requirement in the grammar (e.g., formal features of Case or Agr). The freezing effects in Hindi pose a problem for these derivational approaches. As pointed out in Mohanan and Mohanan (1994), in order to account for the freezing effects in (6) in an analysis involving movement, it will be necessary to prohibit both NP movement and head movement if the theme is the Spec of AgrSP and the nonvolitional experiencer is the Spec of AgrOP. Explaining word order freezing in a way that can relate derivational markedness to the relative markedness of arguments along various dimensions may not be an impossible task, but it is worth exploring an alternative theory in which markedness is built into grammars in the form of violable universal constraints, not simply a criterion external to the grammar, evaluating its complexity.

In this paper, I put forward an OT analysis of restricted word order variation in Hindi which captures both generalizations stated in (8). It will be shown that generalization (8a) follows from the general model of hierarchy alignment in OT (Prince and Smolemsky 1993; Aissen 1999) (section 3.4.1). The ungrammaticality of (a2) in (7) is part of the generalization about ambiguity avoidance and the preference for the unmarked structure, and the analysis to be developed here has it as a direct consequence without having to stipulate it using further constraints (sections 3.4.1 and 3.5.2).

Honorification and Word Order Freezing in the Dative-Subject Construction in Korean

'The worst of the worst' type of word order freezing is observed in Korean as well, a canonical SOV language where the surface order of constituents is strictly head-final. The freedom of word order in sentences with a ditransitive verb *cwu-* 'give' is illustrated in (9).

(9) a. Mary-ka ai-eykey senmwul-ul cwu-ess-ta.
Mary-NOM child-DAT present-ACC give-PAST-DECL
'Mary gave a present to the child.'
b. Mary-ka senmwul-ul ai-eykey cwu-ess-ta.
c. Ai-eykey Mary-ka senmwul-ul cwu-ess-ta.
d. Ai-eykey senmwul-ul Mary-ka cwu-ess-ta.

e. Senmwul-ul Mary-ka ai-eykey cwu-ess-ta.
f. Senmwul-ul ai-eykey Mary-ka cwu-ess-ta.

With existential-possessive predicates and psych-predicates, the subject is marked with the dative case marker. In the dative subject construction too, either order of the two arguments is allowed as shown in the examples in (10) and (11).

(10) a. Mary-eykey kay-ka philyoha-ta.
Mary-DAT dog-NOM need-DECL
'Mary needs a dog.'
b. Kay-ka Mary-eykey philyoha-ta.
'Mary needs a dog.'

(11) a. John-eykey kohyang-i kulip-ta.
John-DAT hometown-NOM be missable-DECL.
'John misses his hometown.'
b. Kohyang-i John-eykey kulip-ta.
'John misses his hometown.'

However, freedom of word order disappears in the dative subject constructions when the honorified subject cooccurs with the honorific verbal agreement marker *-si*. This loss of word order freedom is illustrated in the examples in (12) and (13).

(12) a. Lee kyoswunim-**kkey** kay-ka philyoha-**si**-ta.
Prof. Lee-DAT.HON dog-NOM need-HON-DECL
'Prof. Lee (honorified) needs a dog.'
b. *Kay-ka Lee kyoswunim-**kkey** philyoha-**si**-ta.
'Prof. Lee (honorified) needs a dog.'

(13) a. Halmeni-**kkey** kohyang-i kuliwu-**si**-ta.
grandmother-DAT.HON hometown-NOM be missable-HON-DECL
'Grandmother (honorified) misses her hometown.'
b. *Kohyang-i halmeni-**kkey** kuliwu-**si**-ta.
'Grandmother (honorified) misses her hometown.'

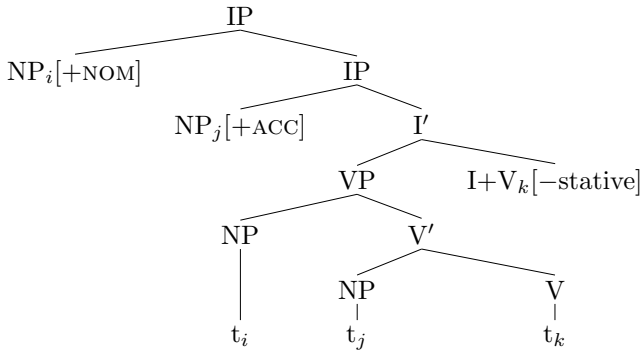
Examples (12b) and (13b) are another instances of 'the worst of the worst': the subject referring to an honorified being is neither in an unmarked nominative case nor in an unmarked clause-initial position. In sharp contrast, the usual (di-)transitive clauses with a nominative subject do not show any word order freezing when the subject is honorified. The theoretical point of this discussion that needs to be captured is summarized in (14).

(14) Generalization 3: Canonical word order becomes fixed when the subject triggering honorific agreement on the verb is not in nominative case. That is, the subject triggering honorific agreement on the verb should not be in both a marked case and in a marked position.

As far as I know, no account has been proposed to explain this generalization, and in fact various approaches to scrambling mentioned so far have little to say about this. Here I will briefly discuss the problems posed by the Korean data for derivational approaches to word order variation.

In her dissertation on scrambling in Korean, Y. Lee (1993) developed a very thorough analysis of word order variation with a number of important consequences. Assuming the VP internal subject hypothesis, she proposes that all arguments have to move out of VP and are adjoined to IP to be assigned Case, resulting in an S-structure representation like (15).

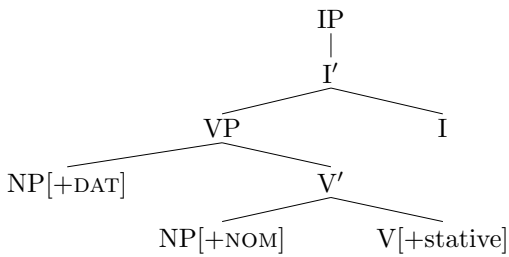
(15)



After head movement of the verb to I, both the nominative Case licenser, INFL, and the accusative Case licenser, INFL[−stative], are in the same position (the feature [−stative] is due to the verb) (Y. Lee 1993:68). This leads us to expect that the subject and the object may be arranged in any order, giving rise to scrambling effects.

Y. Lee (1993) assumes that dative case is an inherent case, assigned at D-structure, which looks like (16).

(16)

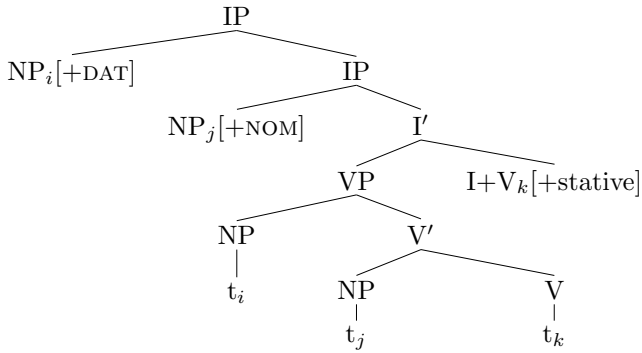


In (16) the subject of the dyadic stative predicates is assigned dative by the verb, and the object is assigned nominative (which is assumed to be the default case in Korean) under government by INFL. Since there is no Case-driven movement here, Y. Lee's (1993) analysis predicts that there will be no scrambling, and therefore wrongly predicts fixed word order for all dative subject constructions in Korean.

The problem is not solved even if we assume that dative case is not an inherent case but a structural Case, as suggested by Yoon (1996). Under

this assumption, then, the experiencer or possessor subject NP may be assigned dative Case either in its base-generated position, SpecVP, or in IP-adjoined position just like the accusative objects of non-stative transitive verbs (see (15)). Both the subject and the object move out of VP to be assigned Case, resulting in an S-structure representation like (17).

(17)



Thus, in this system scrambling is due to the fact that a dative subject can be assigned its Case either in its D-structure position or in IP-adjoined position. Since both the nominative Case licenser, INFL, and the dative Case licenser, INFL[+stative], are in the same position, the subject and the object may now be arranged in any order. However, this analysis predicts free word order for both the standard case frame (nominative-accusative) and the dative-nominative frame, missing the generalization that word order freezing happens in the dative subject construction only in a specific context—when the subject triggers honorific agreement on the verb. The only difference between the frozen and the non-frozen sentences is that only the former (e.g., (12a) and (13a)) contain the honorific marker *-si* on the verb, but not the latter (e.g., (10) and (11)). Given the lack of any theory neutral evidence for the existence of AgrP and various Spec-head relations in Korean, it is difficult to ascribe the contrast between the frozen and non-frozen sentences to their difference in syntactic structure and derivational complexity.[9]

In sum, the word order freezing phenomena in Hindi and Korean examined in this section strongly suggest that there is another important dimension that constrains word order independently of discourse prominence or formal features such as Case and Agr, namely the relative *markedness* of argument types. A central problem then is how to formally relate the rela-

[9]The phrasal head analysis of inflectional morphemes in Korean has been proposed by Whitman (1989), Ahn (1991), and many others. Ahn (1991) also suggests that the honorific marker *-si* and the negative marker *an* are syntactic heads of AgrP and NegP respectively. However, consistent with lexicalist theories of grammar, I will assume that Korean lacks functional projections and that verbal inflectional morphemes and case markers are affixes (showing lexical properties), following Han (1991), Sells (1995), Cho and Sells (1995), Kim (1999), and Choi (1999).

tive markedness of arguments along multiple dimensions to word order. In section 3.4, I will show that a nonderivational analysis of syntactic markedness like the one offered by OT captures the markedness generalizations about the freezing effects in Hindi and Korean in a way that acknowledges the universal basis of these effects and at the same time accounts for the language-particular ways in which these effects are realized. As we will see, 'the worst of the worst' type of word order freezing discussed in this section follows naturally from the interaction of markedness constraints derived through hierarchy alignment, constraints on the realization of information structure and constraints on canonical ordering.

3.2.2 Morphological Ambiguity and Word Order Freezing

The data presented in the preceding section have illustrated word order freezing in clauses which contain highly marked argument types. A different type of word order freezing is observed in Hindi and Korean in sentences where case markings on nominal arguments of a single predicate are identical. An example of the double nominative construction in Hindi is given in (18). The subject *patt^{h}ar* 'stone' is nominative because the transitive verb *tod* 'break' is not in perfective aspect, and the inanimate object *botal* 'bottle' is also nominative.[10] The examples in (18) and (19), taken from Mohanan (1992), show that the order of the two nominative constituents is "frozen" in SOV. This happens in a null context and in certain special discourse contexts (e.g., in an all focus context). Reversing the order of the two arguments in (18) yields a new sentence in (19) in canonical SOV order rather than maintaining meaning in an OSV order.

(18) Patt^{h}ar botal todegaa.
stone-NOM bottle-NOM break-FUT
(i) 'The stone will break the bottle.'
(ii) *'The bottle will break the stone.'

(19) Botal patt^{h}ar todegaa.
bottle-NOM stone-NOM break-FUT
(i) 'The bottle will break the stone.'
(ii) *'The stone will break the bottle.'

The multiple functions of some of the case markers in Hindi provide a further source of word order freezing in sentences with multiple nominals bearing the same case markers. For instance, the dative case marker and the accusative case marker in Hindi are identical: they are both *-ko*. The verb

[10]The choice between zero and overt case marking in Hindi is determined not by lexical subclasses of nouns but by their syntactic and semantic properties. Therefore the situation in Hindi double nominative construction is not an instance of case syncretism found in languages such as German and Russian, where certain lexical subclasses of nouns have identical case forms systematically associated with two or more distinct case features (Arnold Zwicky, p.c., April 1999). Word order freezing in Russian sentences with two syncretized arguments will be briefly discussed in section 3.5.2.

sik^{h} 'teach' in (20) can take either an ergative or a nominative subject. However, when a modal of obligation is superimposed on it, the modal induces dative case on that subject, as in (20).

(20) Raam-ko ilaa-ke bacce-ko gaanaa sik^{h}aanaa hai.
Ram-DAT Ila's child-ACC music-NOM teach-NF be-PRES
(i) 'Ram has to teach music to Ila's child.'
(ii) *'Ila's child has to teach music to Ram.'

In (20) the dative subject *Raam* canonically precedes the two nonsubject arguments, the accusative goal object and the nominative theme object. That is, the initial *-ko* marked NP is interpreted as the agent but not as the goal fronted to the clause-initial position.[11]

Similar facts are found in multiple nominals with the case marker *-se*, which indicates instrument, source, path, the demoted subjects of passive, and so on. In (22), the passive of (21), both the demoted agent and the source bear the case marker *-se*. Grammatical function and thematic role are often closely aligned in Hindi and therefore it is difficult to distinguish which of the two influences ordering. However, the examples in (22) provide justification for the proposal made by Mohanan (1992, 1994a) and others that it is in fact the thematic hierarchy that determines canonical order. In (22a) and (22b), the thematic role of the arguments does not match their grammatical functions in terms of hierarchy: in spite of the fact that the initial *-se* marked NP is an ADJUNCT function, which is lower on the grammatical function hierarchy (i.e. SUBJ > OBJ > OBL > ADJ (Bresnan 1994; Mohanan and Mohanan 1994; Bresnan 2001b)) than the nominative subject, it canonically precedes the subject. Here again, the initial *-se* marked NP is not interpreted as fronted OBLIQUE.

(21) Coor-ne kal Ravii-se paise curaae.
thief-ERG yesterday Ravi-from money-NOM steal-PERF
'The/a thief stole money from Ravi yesterday.'

(22) a. Coor-se kal Ravii-se paise curaae gae.
thief-INST yesterday Ravi-from money-NOM steal-PERF go-PERF
(i) 'Money was stolen from Ravi yesterday by the/a thief.'
(ii) *'Money was stolen from the thief yesterday by Ravi.'
b. Ravii-se kal coor-se paise curaae gae.
Ravi-INST yesterday thief-from money-NOM steal-PERF go-PERF
(i) 'Money was stolen from the/a thief yesterday by Ravi.'
(ii) *'Money was stolen from Ravi yesterday by the thief.'

[11] Mohanan (1994b) has proposed a case OCP principle, a constraint that disfavors identical case marking on more than one nominal in Hindi. Multiple occurrences of nominals marked with *-ko* and *-se* allowed in these examples are not subject to the case OCP principle, which is a restriction on case markings only, not on case features. Repetition of a case marking is allowed as long as the markings refer to distinct case features and the nominals marked with the same case markings are not adjacent.

Therefore following Mohanan (1992, 1994a) and also Sharma (1999), I assume that the canonical or unmarked word order in Hindi conforms to the thematic role hierarchy:

(23) Thematic Role Hierarchy
(Bresnan and Kanerva 1989; Bresnan and Zaenen 1990)
agent > beneficiary > experiencer/goal > instrument > patient/theme > locative

A similar phenomenon is found in the Korean double nominative construction. The Korean dyadic stative predicates exemplified in (10)–(13) above allow a different case realization on the experiencer (Gerdts and Youn 1988; Kim 1990; Hong 1991; Yoon 1996). It may be associated with nominative case as in (24)–(25) as well as dative case as in (10)–(13).

(24) Mary-ka kay-ka philyoha-ta.
Mary-NOM dog-NOM need-DECL
'Mary needs a dog.'

(25) John-i kohyang-i kulip-ta.
John-NOM hometown-NOM be missable-DECL
'John misses his hometown.'

In sentences where the case markings on both the subject and the object NP constituents are identical (i.e. nominative), their order is fixed as SOV. For instance, the two examples in (26) are only grammatical as interpreted in an SO order, and reversing the order of the nominal constituents of (26a) yields a new sentence (26b) in SOV order.

(26) a. Mary-ka ku kay-ka coh-ta.
Mary-NOM that dog-NOM like-DECL
(i) 'Mary likes that dog.'
(ii) *'That dog likes Mary.'

b. Ku kay-ka Mary-ka coh-ta.
that dog-NOM Mary-NOM like-DECL
(i) 'That dog likes Mary.'
(ii) *'Mary likes that dog.'

The Hindi and Korean examples above reveal the following generalizations, due to Mohanan (1992):

(27) Generalization 4: Canonical word order determined by the grammatical function hierarchy or the thematic role hierarchy becomes fixed if the case markings on two nominal arguments of a single predicate are identical under two alternative thematic role interpretations of the nominals.

Similar freezing effects have been observed in previous studies in various languages with fairly free word order, mostly without an explanatory anal-

ysis[12] (e.g., Rudin 1985 and Siewierska and Uhlirova 1998 for Bulgarian; Lenerz 1977 for German; Kuno 1980 for Japanese; England 1983 for Mayan languages; Foley 1986 for Papuan languages; Siewierska and Uhlirova 1998 for Polish; Jakobson 1963, Bloom 1999 for Russian), and occasionally cited as a problem for previous generative approaches to word order and case. One of the problematic aspects of the word order freezing is that it is not absolute but it is always possible to supply a context that brings out the interpretation "disallowed" in a null context (e.g., the OSV reading). In fact, it is usually fairly easy to find real-life examples where preferences for the unmarked word order and interpretation in sentences involving morphological ambiguity are overridden by constraints on the realization of information structure or selectional restrictions on verb arguments, and as a result the "disallowed" interpretation is the intended one. This raises the more general question of what role is played by syntactic constraints and by other sources of "soft" information in the order and interpretation of arguments. Section 3.5 shows that the emergence of the unmarked word order and interpretation in sentences with ambiguous case marking finds a natural analysis in OT, a general theory of constraint interaction which is designed from the beginning to incorporate soft universal constraints.

3.3 Deriving Free Word Order in OT-LFG

So far I have presented data from Hindi and Korean suggesting that theories of word order need to incorporate markedness and express the relation between alternative means of realizing grammatical roles, e.g., word order and case marking. I have also suggested certain generalizations that need to be captured by any general theory of word order. OT is a framework well-suited to the task of formally modeling the markedness generalizations because in this theory markedness is built into grammars in the form of violable universal constraints. In this section, I first briefly discuss the core ideas in OT and present major constraints on word order and basic clause structure, proposed in recent works on constituent ordering.

OT as a general theory of constraint interaction has been applied to a number of areas of linguistic research since its extraordinary success in the domain of phonology. For the domain of syntax, a growing body of work shows that many of the motivations for the OT approach to phonology are paralleled in syntax. Throughout this paper, I assume the formal framework of Lexical-Functional Grammar (LFG) recast within the OT framework (OT-LFG) (Bresnan 2000a,b, 2001a; Choi 1996, 1999; Kuhn 1999, 2001; Sells 1999, 2000, 2001). In this section, for illustrative purposes, we will consider

[12] Notable exceptions include Bloom (1999), who proposes an explicit, formal LFG account of word order freezing in Russian, and Kiparsky (1998), who has developed an account of restricted word order variation found in dative subject constructions and double nominative constructions in Hindi, Korean and Icelandic, based on case licensing theory.

how the basic clause structure and free word order in Hindi and Korean are derived in the framework of OT-LFG.

Input

In OT a grammar is a function mapping each linguistic input to its correct structural description or output. Within the OT-LFG framework inputs are taken to be a (possibly underspecified) feature structure representing (i) a predicator, (ii) the proto-role properties (P-ROLE PROP) of its argument(s), Proto-Agent (P-A) and Proto-Patient (P-P) (Dowty 1991), represented with features [VOL(ITIONALITY)], [CAUS(ALITY)], [SENT(IENCE)], etc. (Asudeh, 2001), and (iii) other morphosyntactic and semantic information (e.g., TENSE and ASPECT) in a language independent form (Bresnan, 2001a). The universal input is modeled by sets of f(unctional)-structures. Following Choi (1996, 1999), I further assume that the input also contains a description of the information status of the verb arguments represented with two features [±PROM(INENT)] and [±NEW]. The input that has *Anuu-ne caand dek^{h}aa* 'Anu saw the moon' (=(1a)) in Hindi as its optimal realization would be the feature structure in (28).

(28)

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Anu'} \\ \text{PROM} & + \\ \text{NEW} & - \\ \text{P-ROLE PROP} & \begin{bmatrix} \text{VOL} & + \\ \text{CAUS} & - \\ \text{SENT} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'moon'} \\ \text{P-ROLE PROP} & \begin{bmatrix} \text{VOL} & - \\ \text{CAUS} & - \end{bmatrix} \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{'see<P-A}_x\text{, P-P}_y\text{>'} \end{bmatrix}$$

(P-A: topic, P-P: discourse neutral)

I assume a four-way distinction of discourse functions based on these two features in (29), following the distinction posited by Butt and King (1996).[13]

[13]Choi's (1996) original use of the two discourse features is different from Butt and King (1996) in that she differentiates between contrastive focus and completive focus, as in (i). I do not discuss contrastive focus in this paper.

(i) Sentence = focus, ground

	+Prom	−Prom
−New	Topic	Background
+New	Contrastive Focus	Completive Focus

(29)

	+Prom	−Prom
−New	Topic	Background
+New	Focus	Completive Information

Topic and background share the feature [−new] since they both express the ground material (i.e. the material that a sentence shares with the preceding sentences), while focus and completive information are classified as [+new] information since they both introduce new information into the discourse. The more prominent information units of a sentence, i.e. topic and focus, are differentiated in terms of the feature [prom] from their less prominent pairs, background and completive information respectively.

In Hindi and Korean topics occur in clause-initial position, which I assume to be a position adjoined to S, following the analysis of scrambling in free word order languages (e.g., Y. Lee 1993; Choi 1999; cf. King 1995 for Russian; Dviwedi 1994, Sharma 1999 for Hindi; Butt and King 1996 for Urdu and Turkish).[14]

Apart from topic, the two most commonly employed discourse functions in Hindi are focus and postposed background (to be described shortly). The major function of focus is to provide new information relevant for the discourse structure. If there is only one focused constituent in the sentence, then it must appear immediately before the verb (Butt and King 1996; Sharma 1999). This is illustrated in (30).[15]

(30) a. Niinaa-ne Mohan-ko [tofii]$_{FOC}$ d-ii.
Nina-erg Mohan-dat toffee-nom give-perf
'Nina gave TOFFEE to MOHAN.'

b. #Niinaa-ne [Mohan-ko]$_{FOC}$ tofii d-ii.
Nina-erg Mohan-dat toffee-nom give-perf
'Nina gave toffee to MOHAN.'

Certain syntactic and information structural differences distinguish two types of nonprominent information: preverbal (completive information) and postverbal nonprominent information (background information). Consider the Urdu/Hindi sentences in (31)–(32), taken from Butt and King (1996).

[14]The term 'scrambling' here is used to refer to the phenomenon of multiple possibilities of the order of phrasal constituents which carry argument functions rather than to the formal operation.

[15]In Hindi, in addition to the preverbal focus, in-situ focusing of a phrase is possible in a sentence with multiple foci. This in-situ focus is always interpreted as contrastive focus, as illustrated in the example in (i) (taken from Butt and King 1996, footnote 13). I will not discuss multiple foci and in situ contrastive focus in this paper.

(i) (Aadnaan-keliye nahii) Naadyaa-ne [Hassan-keliye]$_{CF}$ [tofii]$_{F}$ xarid-ii.
Adnan-for Nadya-erg Hassan-for toffee-nom buy-perf
'Nadya bought TOFFEE for HASSAN (not for Adnan).'

(31) Naadyaa kahãã-se aa rahii hai?
Nadya-NOM where-from come STATE be.PRES
'Where is Nadya coming from?'

(32) a. Naadyaa to abhii [tofii]$_{COMPL}$ [bazaar-mẽ]$_{FOC}$
Nadya-NOM indeed just toffee-NOM market-LOC
xarid rahii t^{h}ii.
buy STATE be.PRES
'Nadya was just buying toffee at the market.'

b. #Naadyaa to abhii [bazaar-mẽ]$_{FOC}$ xarid rahii t^{h}ii [tofii]$_{BCK}$

Since *bazaar-mẽ* 'at the market' provides the information which answers the question in (31), it is focused, while the object *tofii* 'toffee' represents completive information, which expresses purely new information that is neither topicalized nor focused. In this context, the utterance in (32b), where *tofii* 'toffee' appears postverbally, is infelicitous because it presumes the toffee to be a familiar entity which represents known background information.

Candidates

Given an underspecified input f-structure (like the f-structure in (28)), a set of output candidates are generated by the generator GEN. Here I assume that candidate sets consist of pairs of a c(onstituent)-structure and its corresponding f-structure, which is subsumed by the input f-structure (Bresnan 2000a; Kuhn 1999, 2001). Candidates are evaluated against the input with respect to a set of ranked constraints, and all constraints are universal and violable.

Constraints

Clause structure and word order are constrained by potentially conflicting constraints in several parallel structures of grammar. To derive the canonical word order and deviations from this order in Hindi, I adopt the constraints proposed in previous works based on OT-LFG, in particular by Choi (1999) and Sells (1999, 2000, 2001). Those that are particularly relevant for present purposes are given below. The interaction of two alignment constraints in (33b) and (33c) gives basic subtypes of clause structure without problematic recourse to complementary Left and Right alignment constraints (see Sells (1999, 2000, 2001) for details). For example, ranking Spine-R over VHD-L will give right-branching languages. Head-final languages like Hindi, Japanese and Korean, which lack the structural functional head of I, instantiate fully right-branching, with a single co-head at the bottom (Sells 1999).

(33) Constraints on Clausal Skeleton (Sells 1999, 2000, 2001):
  a. The co-head of the clause is any node which is part of the Extended Projection (Grimshaw 1991, 1997), including V, V′, VP, I, I′, C and C′.
  b. Spine-R: co-head aligns right in its local subtree.
  c. Vhd-L: X^0 verbal head aligns left in its local subtree.

The ordering of a verb's arguments in Hindi results from the interacting competing sets of constraints on word order: constraints on canonical ordering based on the hierarchies of grammatical functions and thematic roles (34); and information structuring constraints (35) distinguishing the contextual dimensions of discourse prominence and novelty, each marked by a binary feature. Here information structuring constraints proposed by Choi (1999) (i.e. Prom and New) are reinterpreted as f-structure alignment constraints à la Sells (1999, 2000, 2001b).

(34) Canonical Phrase Structure Constraints Canon (Choi 1999):
  a. Canon_{GF}(f-s/c-s correspondence): Grammatical functions align with their canonical argument positions in c-structure according to the function hierarchy.[16]
  (subj > d.obj > i.obj > obl > adjunct (Bresnan 1994; Mohanan and Mohanan 1994; Bresnan 2001b))
  b. Canon_{θ}(a-s/c-s correspondence): Non-verbal arguments at c-s align according to the thematic hierarchy.
  (Agent > Beneficiary > Experiencer/Goal > Instrument > Patient/Theme > Locative (Bresnan and Kanerva 1989; Bresnan and Zaenen 1990; Bresnan 2001))

(35) Information Structuring Constraints:
  a. Top-L: Topic aligns left in the clause.
  b. Foc-L: Focus aligns left in the clause.
  c. Bck-R: Background information aligns right in the clause.
  d. Compl-L: Completive information aligns left in the clause.

The faithfulness constraints in (36) ensure that the optimal candidate faithfully represents the proto-role information and the discourse information in the input:[17]

(36) Faithfulness Constraints:
  a. Ident-IO(p-role): The value of the proto-role features in the input (e.g., [vol], [caus], [sent], etc.) is preserved in the output.

[16] Strictly speaking, in Choi (1996, 1999), Canon_{GF} is split into cn1: 'subj should be structurally more prominent (e.g., c-command) than non-subj functions' and cn2: 'Non-subj functions align reversely with the c-structure according to the function hierarchy'.

[17] The Max-io class of faithfulness constraints, which penalize the deletion of features in the input, are omitted because they do not crucially distinguish the candidates under discussion here.

b. Dep-IO(prom): The feature [prom] in the output is present in the input.
c. Dep-IO(new): The feature [new] in the output is present in the input.

The discourse motivation for locating background information at one end of the clause and other discourse information at the other seems transparent. For Hindi, the dominance ranking is as (37). With this ranking, a topic always more to the left than a focus in the same clause; a background is always more to the right than spine elements in the same clause.

(37) Ranking for Hindi: Faithfulness constraints in (36) ≫ Bck-R, Top-L ≫ Compl-L≫ Foc-L≫ Spine-R ≫ Canon$_\theta$≫ Canon$_{GF}$≫ Vhd-L

Crucially, the ranking for Hindi in (37) can predict that when the arguments do not differ in informational status, the canonical constraints will take effect, leading to sov order; when there are differences, the canonical sov order will however violate information structuring constraints, so that competitors with a noncanonical ordering may win out.

Korean differs from Hindi in one respect:[18] Canon$_{GF}$ has a stronger effect than Canon$_\theta$. This difference between Hindi and Korean can be handled by modifying the ranking of two types of canonical phrase structure constraints as follows:

(38) Ranking for Korean: Faithfulness constraints in (36) ≫ Top-L≫ Compl-L≫ Foc-L≫ Spine-R ≫ Canon$_{GF}$≫ Canon$_\theta$≫ Vhd-L

Now, suppose that the experiencer argument *Anuu* of the Hindi volitional transitive verb *dek^{h}aa* 'saw' is topic (i.e. prominent given information) and the theme *caand* 'moon' is focus (i.e. prominent new information). In this context the input is as in (28) above. This results in the optimal output (39a) (=(1a)), going through the constraint competition in (39) (violations of ordering constraints are computed by counting constituents from the left.[19]). The ranking relation of the constraints separated by the comma is not specified here. Canon$_\theta$ is omitted here, since it has the same effect as Canon$_{GF}$ in this case. The constraints on faithfulness to proto-role information and discourse information in (36) will also be omitted in subsequent tableaux until they play a crucial role in section 3.5, and

[18] Another difference is that clause-final backgrounding in Korean is more restricted than in Hindi. I do not include Bck-R in (38) as its effects are not relevant here.

[19] Violations of Vhd-L are counted within a local subtree (here VP) (Grimshaw 1997). As argued convincingly in Sharma (1999), no arguments appear within VP in Hindi, whether in specifier or complement position, and all arguments are generated directly under S. In other words, a VP does not contain the verb and its complements. Instead, the only VP-internal elements are those which are preverbally focused.

for our purposes here, we consider only candidates faithfully representing proto-role information and discourse information in the input in (28).

(39) Tableau 1. Volitional Transitives in Hindi

		BCK-R,	TOP-L	COMPL-L	FOC-L	Spine-R	CANON$_{GF}$	VHD-L
	CANDIDATES:							
☞	a. [$_S$ S$_{TOP}$ [$_S$ [$_{VP}$O$_{FOC}$V]]]		0		1	0	0	1
	b. [$_S$ [S$_{TOP}$ [$_{VP}$V]]$_S$ O$_{FOC}$]		0		2	1	0	0
	c. [$_S$ O$_{FOC}$ [$_S$ S$_{TOP}$ [$_{VP}$ V]]]		1		0	0	1	1
	d. [$_S$ [O$_{FOC}$ [$_{VP}$V]]$_S$ S$_{TOP}$]		2		0	1	1	0
	e. [$_S$ [$_{VP}$ V] S$_{TOP}$O$_{FOC}$]		1		2	2	0	0
	f. [$_S$ [$_{VP}$ V] O$_{FOC}$S$_{TOP}$]		2		1	2	1	0

As noted previously, the varied word orders in Hindi are optional from a purely syntactic point of view: although all orders shown in (39) above are in principle available, each is preferentially brought out by a particular context to mark a particular information structure. In the present framework this can be captured by considering the role of the input (Choi 1996, 1999). For example, the candidate (39c) with OSV order corresponds more faithfully to (40); the candidate (39d) with OVS order, to (41) (in the f-structures below the proto-role properties associated with each argument role are omitted for simplicity). In other words, according to this analysis OSV becomes optimal for expressing the topical status of the object and the newness of the subject under the same ranking that yields (39a) as the optimal output; OVS is optimal for expressing the nonsalient status of the subject as background information.

(40)
$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{`Anu'} \\ \text{PROM} & - \\ \text{NEW} & + \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{`moon'} \\ \text{PROM} & + \\ \text{NEW} & - \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{`see<P-A}_x\text{, P-P}_y\text{>'} \end{bmatrix}$$

(P-A: focus, P-P: topic)

(41)
$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Anu'} \\ \text{PROM} & - \\ \text{NEW} & - \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'moon'} \\ \text{PROM} & + \\ \text{NEW} & - \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{'see<P-A}_x\text{, P-P}_y\text{ >'} \end{bmatrix}$$
(P-A: background, P-P: topic)

To summarize, the varied word orders in Hindi and Korean and their related interpretations find a natural analysis in the OT-LFG framework combining the ideas of imperfect correspondence and violable constraints. The next step is to explain why orderings in sentences with more marked types of subjects (e.g., nonvolitional subject and dative subject) are more restricted, and word order is even frozen in SOV order in the most marked situation (see the discussion in section 3.2.1).

3.4 Markedness and Word Order Freezing

This section presents an OT-LFG account of the 'worst of the worst' type of freezing effects in Hindi and Korean. Central to my account is the use of harmonic alignment (Prince and Smolensky 1993; Aissen 1999) and local conjunction in OT (Smolensky 1995). I show that under the analysis I develop here, word order freezing in "free" word order languages is not the exception, but just one of the options allowed by the universal scales of grammatical function, semantic role, case, etc.

3.4.1 Hindi

In this section I demonstrate that the 'worst of the worst' type of the freezing effects in Hindi outlined in section 3.2.1 follows naturally from the general model of harmonic alignment developed in Prince and Smolensky (1993) and adopted in Aissen (1997, 1999). The formal definition of harmonic alignment is given in (42) (Prince and Smolensky 1993:136).

(42) Suppose a binary dimension D_1 with a scale X > Y on its element {X, Y}, and another dimension D_2 with a scale a > b > ... > z on its elements. The harmonic alignment of D_1 and D_2 is the pair of Harmonic scales:
H_X: X/a ≻ X/b ≻ ... ≻ X/z
H_Y: Y/z ≻ ... ≻ Y/b ≻ Y/a

The *constraint alignment* is the pair of constraint hierarchies:
C_X: *X/z ≫ ... ≫ *X/b ≫ *X/a
C_Y: *Y/a ≫ *Y/b ≫ ... ≫ *Y/z

Harmonic alignment is an important source of constraints in OT, and was introduced originally to express the relation between syllable position and sonority: the more prominent position (the nucleus) attracts segments which are more sonorous, while less prominent positions (the margins) attract less sonorous segments. This paper extends an idea first proposed in Aissen (1997) and developed further in Artstein (1998) and Aissen (1999, 2000), that harmonic alignment plays a central role in the domain of morphosyntax to express the relative markedness of different associations of morphosyntactic prominence hierarchies. The basic idea is that subject function plays a role in the clause analogous to that played by the peak in syllable structure: it is the most prominent grammatical function and thereby attracts elements which are relatively prominent on other dimensions such as semantic role (Aissen 1999:9).

Now, we can apply the definition above to the three prominence scales in (43)—the grammatical function (GF) scale, the semantic role scale and the scale on c-structure position. I will adopt the scale Subject > Nonsubject (43a) proposed by Aissen (1999) and the scale Proto-Agent (P-A) > Proto-Patient (P-P) (43b), where ">" means "more prominent than". For present purposes, I use the last two elements of the decomposed prominence scale P-A_{vol} > P-A_{-vol} > P-P proposed by Asudeh (2001). In the predicates I discuss here, the experiencer argument corresponds to a nonvolitional Proto-Agent (P-A_{-vol}), and the theme argument to Proto-Patient (P-P). Furthermore, I use the additional scale of structural position (Lee 1999).

(43) Universal Scales
   a. GF: SUBJ > Non-SUBJ
   b. Semantic Role: P-A_{-vol} > P-P
   c. Position: Initial > Non-initial

If the scales in (43) are harmonically aligned, we obtain the pairs of combined harmony scales in (44):

(44) Harmony Scales derived through Harmonic Alignment
   a. H_1: SUBJ/P-A_{-vol} $\succ$ SUBJ/P-P
   b. H_2: NSUBJ/P-P $\succ$ NSUBJ/P-A_{-vol}
   c. H_3: SUBJ/Initial $\succ$ SUBJ/NInitial
   d. H_4: NSUBJ/NInitial $\succ$ NSUBJ/Initial

The first two harmony scales concern the association between grammatical function and semantic role, and assert that the unmarked situation is for the subject to be Proto-Agent, and for the object to be Proto-Patient (The connective "$\succ$" is read as "more harmonic than"). The last two harmony scales involve the alignment of the grammatical function hierarchy and the structural position hierarchy. The basic insight is that the unmarked situation is for the subject to be in initial position, and for the nonsubject

to be in noninitial position. The corresponding constraint alignments are the pairs of structural markedness constraint hierarchies in (45):

(45) Constraint Subhierarchies
  a. C_1: *SUBJ/P-P $\gg$ *SUBJ/P-A$_{-vol}$
  b. C_2: *NSUBJ/P-A$_{-vol}$ $\gg$ *NSUBJ/P-P
  c. C_3: *SUBJ/NInitial $\gg$ *SUBJ/Initial
  d. C_4: *NSUBJ/Initial $\gg$ *NSUBJ/NInitial

Each subhierarchy in (45) expresses the universal markedness relation (e.g., a clause with a P-P subject will lose out to a clause with a P-A subject, other things being equal). The important property of the constraint hierarchies in (45) is that while the ranking of constraints within a subhierarchy is fixed (e.g., *SUBJ/P-P always outranks *SUBJ/P-A$_{-vol}$), individually they may be variously ranked with respect to other constraints.

However, showing that a Proto-Patient subject and a subject in noninitial position are more marked than a Proto-Agent subject and a subject in initial position is not enough, because both cases are still allowed in Hindi. In order to capture the idea that if the subject is both a Proto-Patient and in a noninitial position at constituent structure, it is the worst of the worst, we can use the mechanism of *local conjunction* (Smolensky 1995:4).[20]

(46) The Local Conjunction of C_1 and C_2 in domain D, C_1 & C_2 is violated when there is some domain of type D in which both C_1 and C_2 are violated. Universally, C_1 & $C_2 \gg C_1$, C_2.

To derive the pattern of universal markedness reflected in freezing effects in Hindi, let us consider the conjunction of the subhierarchy C_3 in (45c) with the high-ranked constraint *SUBJ/P-P in (45a). This results in the new constraint subhierarchy in (47). The high-ranked constraint in (47) expresses the basic idea that if the subject is a highly marked Proto-Patient argument, it should not be in marked noninitial position. This most marked configuration excluded by this constraint obtains in sentences like (6b,c,d,e), repeated below as (48a,b,c,d). Recall from section 3.2.1 that these examples are acceptable only under the analysis of the nonvolitional experiencer as the subject, which cannot be the antecedent of the pronoun; the analysis of the theme as the subject in Hindi is compatible only with the theme-experiencer-verb order. This evidence suggests that when the Proto-Patient (e.g., theme) of a nonvolitional transitive verb is the subject and the Proto-Agent (e.g., experiencer) is the object, their order is fixed as SOV. This is a clear case of avoidance of the worst of the worst.

(47) Conjoining *SUBJ/P-P with C_3:
  C_5: *SUBJ/P-P & *SUBJ/NInitial $\gg$ *SUBJ/P-P & *SUBJ/Initial

[20] Local conjunction was first used in syntax in Legendre, Wilson, Smolensky, Homer and Raymond (1995), and has been extended to various domains of morphosyntax.

(48) a. Anuu-ko Niinaa uskii bastii-mẽ dik^{h}ii.
Anu-DAT Nina-NOM PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

b. Anuu-ko uskii bastii-mẽ dik^{h}ii Niinaa.
Anu-DAT PRON-GEN neighborhood-LOC appear-PERF Nina-NOM
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

c. Dik^{h}ii Niinaa Anuu-ko uskii bastii-mẽ.
appear-PERF Nina-NOM Anu-DAT PRON-GEN neighborhood-LOC
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

d. Dik^{h}ii Anuu-ko Niinaa uskii bastii-mẽ.
appear-PERF Anu-DAT Nina-NOM PRON-GEN neighborhood-LOC
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

Another type of structure excluded in Hindi is one like (6a), repeated here as (49), where the subject is in initial position but the object is postposed. Such a structure is marked with regard to the positioning of both the argument and the head, i.e. in that the verbal head is not the final element of the clause.

(49) Niinaa uskii bastii-mẽ dik^{h}ii Anuu-ko.
Nina-NOM PRON-GEN neighborhood-LOC appear-PERF Anu-DAT
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

To derive this word order pattern in Hindi, first we need a simple restriction on the positioning of the co-head in (33b) (Sells 1999, 2000, 2001), repeated below in (50):

(50) Spine-R: co-head aligns right in its local subtree.

Now applying the same scheme to the conjunction of the subhierarchy C_2 from (43) with the constraint Spine-R in (50), we get the following new subhierarchy in (51), which concerns the markedness of the nonsubject:

(51) *Nsubj/P-A$_{-vol}$ & Spine-R $\gg$ *Nsubj/P-P & Spine-R

The higher-ranked constraint in (51) expresses the basic idea that the most marked situation obtains when the object gets linked to a marked semantic role (i.e. nonvolitional Proto-Agent) *and* the co-head of the clause is not in clause-final position. If this conjoined constraint dominates Bck-R, then marked types of objects would be restricted to the unmarked position, preceding the verb.

The ranking that emerges for Hindi is given in (52).

(52) Ranking for Hindi:
Conjoined markedness constraints:
*subj/P-P & *subj/NInitial, *Nsubj/P-A$_{-vol}$ & Spine-R $\gg$
Linking constraints: *subj/P-P, *Nsubj/P-A$_{-vol}$ $\gg$
Information structuring constraints $\gg$ Canonical phrase structure constraints

Note here that the locally-conjoined constraints are ranked higher than their component constraints such as *Nsubj/P-A$_{-vol}$, which penalize the marked linking pattern. Crucially, the higher ranking of the three conjoined constraints on markedness of argument types over the information structuring constraints (e.g., Top-L and Foc-L) has the effect of restricting the word order freedom motivated by discourse prominence and the newness of arguments (see the discussion in section 3.3): marked argument types (e.g., Proto-Patient subject and nonvolitional Proto-Agent object) must occur in unmarked position in the clause.

Now, let us assume a discourse context in which the Proto-Agent is topic and the Proto-Patient is focus. This particular context renders an input like (53). This input then results in the optimal output (a1), going through the constraint competition in (54). The comma in the tableau indicates that there is no crucial ranking between the constraints separated by it. Also, candidates are again schematically represented, and faithfulness constraints, constraints on clausal skeleton and component constraints of the high-ranked conjoined constraints are omitted, as their effects are not relevant here. Candidates (a1) to (a6) are associated with the same f-structure, where the Proto-Agent argument 'Anu' is canonically mapped to the subject, and the Proto-Patient 'Nina' to the object. Similarly, candidates (b1) to (b6) are paired with the same f-structure with the opposite linking. Also, candidates labeled the same number share the same c-structure string. For example, both candidate (a1) and (b1) share the same string *Anuu-ko Ni-inaa dik^{h}ii.*

(53) Input:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Anu'} \\ \text{PROM} & + \\ \text{NEW} & - \\ \text{P-ROLE} & \text{PROP} \begin{bmatrix} \text{VOL} & - \\ \text{CAUS} & - \\ \text{SENT} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'Nina'} \\ \text{PROM} & + \\ \text{NEW} & + \\ \text{P-ROLE} & \text{PROP} \begin{bmatrix} \text{VOL} & - \\ \text{CAUS} & - \end{bmatrix} \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{'appear<P-A}_x\text{, P-P}_y\text{>'} \end{bmatrix}$$

(P-A: topic, P-P: focus)

(54) Tableau 2. Linking and word order in Hindi nonvolitional transitives

CANDIDATES:	*SUBJ/P-P & *SUBJ/NInitial,	*NSUBJ/P-A_{-vol} & Spine-R	TOP-L	FOC-L	$CANON_{GF}$
☞ a1. $S_{P-A/TOP}O_{P-P/FOC}V$			0	1	
a2. $O_{P-P/FOC}S_{P-A/TOP}V$			1	0	*
a3. $S_{P-A/TOP}VO_{P-P/FOC}$			0	2	
a4. $O_{P-P/FOC}VS_{P-A/TOP}$			2	0	*
a5. $VS_{P-A/TOP}O_{P-P/FOC}$			1	2	
a6. $VO_{P-P/FOC}S_{P-A/TOP}$			2	1	*
b1. $O_{P-A/TOP}S_{P-P/FOC}V$	*!		0	1	*
b2. $S_{P-P/FOC}O_{P-A/TOP}V$			1	0	
b3. $O_{P-A/TOP}VS_{P-P/FOC}$	*!	*!	0	2	*
b4. $S_{P-P/FOC}VO_{P-A/TOP}$		*!	2	0	
b5. $VO_{P-A/TOP}S_{P-P/FOC}$	*!	*!	1	2	*
b6. $VS_{P-P/FOC}O_{P-A/TOP}$	*!	*!	2	1	

Due to the two high-ranking constraints that penalize highly marked types of arguments occurring in noncanonical position, candidates (b1), (b3), (b4), (b5) and (b6) are ruled out immediately: nonsubject initial candidates (b1), (b3), (b5) and (b6) with Proto-Patient subjects are eliminated by the constraint *SUBJ/P-P & *SUBJ/NInitial. Candidates (b3), (b4), (b5) and (b6) have postverbal arguments, violating the constraint *NSUBJ/P-A_{-vol} & Spine-R.[21] Among (a1), (a3), (a4), (a5), (a6) and (b2), (a1) is the best; it satisfies more higher-ranking constraints than any other candidates.[22]

Now, let's consider a context in which the nonvolitional Proto-Agent is a focus and Proto-Patient is a topic. This is illustrated in the following question-answer examples in (55). Suppose that a speaker A asked another speaker B the question in (55) and that the examples in (56) are possible answers to it in that they provide the hearer with information as to who saw Nina, namely Anu. The *what about Niinaa*? phrase, following Vallduví (1992), is used to identify the topic, namely the prominent old information which is the center of interest in the current discourse. Since *Anuu-ko* 'to Anu' provides the information which answers the question, it is focused.

[21] I assume that when gradient alignment constraints like Spine-R are locally conjoined with other markedness constraints, the complex constraint is violated iff both its component constraints are violated at least once. That is, I interpret the conjoined constraint *NSUBJ/P-A_{-vol} & Spine-R as violated by the four candidates (b3), (b4), (b5) and (b6), although (b5) and (b6) are worse than (b3) and (b4) with respect to Spine-R.

[22] The ungrammaticality of (a2) will be explained in section 3.5.2

(55) Aur Niinaa? Niinaa kisko dik^{h}ii?
and Nina? Nina-NOM who-DAT appear-PERF
'What about Nina? Who saw Nina?'
(Lit. 'To whom did Nina appear?')

(56) a. [Niinaa]$_{TOP}$ [Anuu-ko]$_{FOC}$ dik^{h}ii.
Nina-NOM Anu-DAT appear-PERF
'ANU saw Nina.' (Lit. 'Nina appeared to ANU.')
b. *Anuu-ko Niinaa dik^{h}ii.
c. *Niinaa dik^{h}ii Anuu-ko.
d. *Anuu-ko dik^{h}ii Niinaa.
e. *Dik^{h}ii Niinaa Anuu-ko.
f. *Dik^{h}ii Anuu-ko Niinaa.

Among the six examples in (56), only (56a) is an appropriate answer where the topic appears canonically sentence initially, and the focus immediately before the verb. In contrast, in the sentences marked as ungrammatical the topic and focus are not in their canonical position, and as a result they are ungrammatical in the context of (55) as expected given the requirement that topic is clause initial and focus is immediately before the verb.[23]

Suppose that the sentences in (57b,c) are uttered by speaker B as a response to speaker A's question in (57a), followed by (56).

(57) a. Kahan?
'Where?'
b. Apnii bastii-mẽ
self-GEN neighborhood-LOC
'in self's neighborhood'
(Nina (theme) = apnii, Anu (experiencer) ≠ apnii)
c. Uskii bastii-mẽ
PRON-GEN neighborhood-LOC
'in her neighborhood'
(Nina (theme) ≠ uskii, Anu (experiencer) = uskii)

As (57b,c) show, for the Hindi speakers I consulted, Nina is the only eligible antecedent of the reflexive *apnii* within the context of (55) and (57b); it is also the only element in the clause that cannot be coreferent with the pronoun *uskii*. Therefore the facts on word order and coreference in (55)–(57) suggest that the theme argument *Niinaa*, not the experiencer *Anuu*, is the grammatical subject when the former is topic and the latter is focus.

Let us now see how the OT account explains word order freezing in the discourse context in which the nonvolitional Proto-Agent is a focus and Proto-Patient is a topic. In this context the input is as (58). In this

[23]The examples in (56c) and (56d) are felicitous only in the context in which the preverbal NP is topicalized with the verb as an informational unit (Devyani Sharma, p.c., September 1999).

context, however, noncanonical linking becomes optimal under the same ranking, as illustrated in (59).

(58) Input:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Anu'} \\ \text{PROM} & + \\ \text{NEW} & + \\ \text{P-ROLE} & \text{PROP} \begin{bmatrix} \text{VOL} & - \\ \text{CAUS} & - \\ \text{SENT} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'Nina'} \\ \text{PROM} & + \\ \text{NEW} & - \\ \text{P-ROLE} & \text{PROP} \begin{bmatrix} \text{VOL} & - \\ \text{CAUS} & - \end{bmatrix} \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{'appear<P-A}_x\text{, P-P}_y\text{>'} \end{bmatrix}$$

(P-A: focus, P-P: topic)

(59) Tableau 3. Noncanonical linking becomes optimal under the same ranking

CANDIDATES:	*SUBJ/P-P & *SUBJ/Ninitial,	*NSUBJ/P-A$_{-vol}$ & Spine-R	TOP-L	FOC-L	CANON$_{GF}$
a1. S$_{P-A/FOC}$O$_{P-P/TOP}$V			1	0	
a2. O$_{P-P/TOP}$S$_{P-A/FOC}$V			0	1	*
a3. S$_{P-A/FOC}$VO$_{P-P/TOP}$			2	0	
a4. O$_{P-P/TOP}$VS$_{P-A/FOC}$			0	2	*
a5. VS$_{P-A/FOC}$O$_{P-P/TOP}$			2	1	
a6. VO$_{P-P/TOP}$S$_{P-A/FOC}$			1	2	*
b1. O$_{P-A/FOC}$S$_{P-P/TOP}$V	*!		1	0	*
☞ b2. S$_{P-P/TOP}$O$_{P-A/FOC}$V			0	1	
b3. O$_{P-A/FOC}$VS$_{P-P/TOP}$	*!	*!	2	0	*
b4. S$_{P-P/TOP}$VO$_{P-A/FOC}$		*!	0	2	
b5. VO$_{P-A/FOC}$S$_{P-P/TOP}$	*!	*!	2	1	*
b6. VS$_{P-P/TOP}$O$_{P-A/FOC}$	*!	*!	1	2	

The violation patterns for the two top-ranked markedness constraints for (b) candidates in tableau 3 are just like the ones in tableau 2, and they

will remain the same for all other tableaux too, because the violations of these conjoined constraints are not sensitive to discourse context. Thus, the constraint ranking proposed here accounts for the fact that in a clause with a Proto-Patient subject and a nonvolitional Proto-Agent object word order is fixed as SOV for expressing the content in (58), capturing the basic generalization that highly marked argument types occur only in unmarked position.

In sum, I have argued that harmonic alignment in OT can fruitfully be applied in word order freezing found in Hindi: the 'worst of the worst' generalization follows naturally from the general model of harmonic alignment. I have also shown that local conjunction of markedness constraints is highly appropriate to dealing with the relative markedness of argument types and its interaction with word order.

3.4.2 Korean

Marked associations of prominence hierarchies also provide an important source of word order freezing in Korean. Before we turn to how harmonic alignment and local conjunction play a role, let us first consider honorification, with which word order freezing is associated in an interesting way.

Korean is a language which has a highly developed system of honorification. The honorification system can be classified into three types, according to who the target of honoring is—subject honorification, nonsubject honorification and addressee honorification. Here I focus only on subject honorification as it is the only type of honorification that is directly relevant to the discussion of word order freezing in Korean.

Subject NP forms and verb forms in Korean vary depending on the relative social relationship among the referent of the subject, the speaker and the addressee. (60) shows six possible contexts according to the relative relationship among the participants of subject honorification, and which subject NP form and which verb form are used in each context.[24] The relative relationship among the participants of subject honorification can be explicitly represented in the input by using the feature honorification LEVEL. Four levels of honorification are represented by the integers 1–4.[25]

[24]Subject honorification has been used as a classic test for subjecthood in Korean. Only a grammatical subject—not a logical subject (i.e. a passive agent)—is responsible for *si*-marking on the verb. The object does not trigger *si*-marking even if it is topicalized.

[25]In (60), CASE.HON and *-si* indicate the use of an honorific case form and an honorific subject verb form respectively while CASE represents the use of a nonhonorific (plain) case form and Ø a nonhonorific (plain) subject verb form.

(60) Subject honorification in Korean (based loosely on Han (1991))

Social relationship	Subject NP	Verb form	level
a. **subj** > speaker,addressee	N-case.hon	V-*si*	4
b. addressee > **subj** > speaker	N-case.hon/ N-case	V-*si*	3
c. speaker > **subj** > addressee	N-case	V-*si*/V-Ø	2
d. speaker, addressee = subj	N-case	V-Ø	1
e. speaker > **subj** = addressee	N-case	V-Ø	1
f. speaker, addressee > **subj**	N-case	V-Ø	1

level 4 in (60a) corresponds to the context in which the subject referent is socially superior to both the speaker and the addressee. This is the only situation in which the honorific case form and the honorific verbal suffix *-si* are both used, as shown in (61a). In the context (60b), corresponding to level 3, the honorific case marker is optionally used as in (61b), and in the context (60c) the use of the honorific verbal suffix is optional as in (61c). Word order freezing happens only when the honorific affix is present on the verb in the contexts (60a,b,c) (see examples (12) and (13)).

(61) a. Lee kyoswu**nim-kkeyse** hakkyo-ey ka-**si**-ess-ta.
Prof. Lee-nom.hon school-to go-hon-past-decl
'Prof. Lee (honorified) went to school.' (level 4)

b. Lee kyoswu**nim-i/-kkeyse** hakkyo-ey ka-**si**-ess-ta.
Prof. Lee-nom/nom.hon school-to go-hon-past-decl
'Prof. Lee (honorified) went to school.' (level 3)

c. Lee kyoswu**nim**-i hakkyo-ey ka-**si**/Ø-ess-ta.
Prof. Lee-nom school-to go-hon/Ø-past-decl
'Prof. Lee (optionally honorified) went to school.' (level 2)

d. Lee kyoswu-ka hakkyo-ey ka-Ø-ss-ta.
Prof. Lee-nom school-to go-Ø-past-decl
'Prof. Lee went to school.' (level 1)

Honorific information about the subject of the clause can be represented by the feature hon, and there are three sources of [hon+] information that need to be represented in the input: the honorific verbal suffix *-si*, the honorific form of case markers, and the nominal affix *-nim*. Typically, in the contexts of (60a) and (60b), all of these are used, indicating that the subject's referent is owed honor.

The honorific verb form affixed with the verbal honorific suffix *-si* and the nonhonorific verb form have the feature structure in (62a) and (62b), respectively. In (62a) [hon+] information is contributed morphologically by *-si*. The nonhonorific verb form in (62b) is unspecified for [hon].

(62) a. *-si* marked verb form (honorific subject verb form)

V
|
ka-si-ess-ta

$$\begin{bmatrix} \text{TNS} & \text{PAST} \\ \text{SUBJ} & \begin{bmatrix} \text{HON} & + \end{bmatrix} \\ \text{PRED} & \text{`go<P-A}_x\text{,(P-P}_y\text{)>'} \end{bmatrix}$$

b. *-si*-less verb form (non-honorific subject verb form)

V
|
ka-ss-ta

$$\begin{bmatrix} \text{TNS} & \text{PAST} \\ \text{PRED} & \text{`go<P-A}_x\text{,(P-P}_y\text{)>'} \end{bmatrix}$$

I hypothesize the respective representations in (63a) and (63b) for the subject NP form affixed with the honorific case marker and the one with the nonhonorific case marker. [HON+] information is contributed morphologically by the honorific case marker *-kkeyse* in (63a) and the nominal affix *-nim* in (63b). The nonhonorific noun form in (63c) is assumed to be unspecified for [HON].

(63) a. Honorific noun with honorific nominative case marker *-kkeyse*

NP
|
Lee kyoswunim-kkeyse

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{NOM} \\ \text{HON} & + \\ \text{PRED} & \text{`Prof. Lee'} \end{bmatrix} \end{bmatrix}$$

b. Honorific noun marked with plain nominative case marker *-i*

NP
|
Lee kyoswynim-i

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{NOM} \\ \text{HON} & + \\ \text{PRED} & \text{`Prof. Lee'} \end{bmatrix} \end{bmatrix}$$

c. Nonhonorific noun marked with plain nominative marker *-ka*

NP
|
Lee kyoswu-ka

$$\begin{bmatrix} \text{SUBJ} & \begin{bmatrix} \text{CASE} & \text{NOM} \\ \text{PRED} & \text{`Prof. Lee'} \end{bmatrix} \end{bmatrix}$$

Honorific verbal morphology and nominal morphology, which share the [HON+] information, are further classified as in (64) according to the honorification LEVEL information they express.

(64) a. Honorific morphology which is used in the context in which the subject's referent is socially superior to both the speaker and the addressee (LEVEL 4), or to the speaker (LEVEL 3): Honorific case forms.

b. Honorific morphology which is used in the context in which the subject's referent is socially superior to either the speaker or the addressee (LEVEL 4, 3, 2):
Honorific verbal affix and noun forms.

With this classification of honorific morphology we can now formulate three markedness constraints, which apply to the form-content pairings, shown in (65):[26]

(65) HARMONY:

a. $\text{HON}_{CASE} \Leftrightarrow$ LEVEL $3 \vee 4$: Honorific case forms express honorification LEVEL whose value is 3 or higher.

b. $\text{HON}_{V} \Leftrightarrow$ LEVEL $2 \vee 3 \vee 4$: Honorific verb forms express honorification LEVEL whose value is 2 or higher.

c. $\text{HON}_{N} \Leftrightarrow$ LEVEL $2 \vee 3 \vee 4$: Honorific noun forms express honorification LEVEL whose value is 2 or higher.

The harmony constraints in (65) check the correspondence relations (or harmony relations) between types of nominal and verbal forms and the LEVEL information they are associated with. These constraints are violated by a candidate representation whenever the LEVEL information in the f-structure and the honorific marking of the subject and that of the verb in the c-structure make conflicting indications about the facts of the social relationship among the participants of honorification. This point is illustrated in the table in (66) containing four sentences with the same sentential content 'Prof. Kim left'.

(66)

INPUT: [LEVEL 3]	Constraint violations
a. #Kim kyoswunim-i ttena-Ø-ss-ta. Prof. Kim-NOM leave-Ø-PAST-DECL 'Prof. Kim left.'	(65a,b)
b. #Kim kyoswu-ka ttena-si-ess-ta. Prof. Kim-NOM leave-HON-PAST-DECL	(65a,c)
c. #Kim kyoswunim-kkeyse ttena-Ø-ss-ta.	(65b)
d. #Kim kyoswu-ka ttena-Ø-ss-ta.	(65a,b,c)

(All outputs are LEVEL 3.)

[26] The idea on form-function harmony constraints here is drawn from Bresnan (2001a), who has demonstrated convincingly that the pronominal inventories of a language can be derived from interactions of harmony constraints on possible pairings of the pronominal forms and the functional content, faithfulness constraints and markedness constraints on pronominal forms.

Thus the harmony constraints in (65) mark disagreeing candidates like (66a,b,c) or candidates like (66d), in which both the honorific marking of the subject and that of the verb do not correctly indicate the facts of the social context. Candidates like (66d), in which no honorific expressions are present, also violate faithfulness constraints in (67), which require the input feature structures (here the features HON and LEVEL) to correspond to the candidate feature structures to ensure the expressibility of content.

(67) Faithfulness constraints
   a. IDENT-IO(HON): The value of the feature HON in the input is preserved in the output.
   b. IDENT-IO(LEVEL): The value of the feature LEVEL in the input is preserved in the output.

Of course, not all languages have an extensive set of honorific forms, so further constraints are required. The constraints in (68) impose marks against the realization of case markers, nouns and verbs are honorific forms.

(68) Structural markedness constraints
   a. *HON$_{CASE}$: Avoid an honorific form of a case marker.
   b. *HON$_{N}$: Avoid an honorific form of a noun.
   c. *HON$_{V}$: Avoid an honorific form of a verb.

The three families of constraints introduced above are ranked in Korean as shown in (69):

(69) Ranking for Korean:
   Harmony constraints in (65) $\gg$ Faithfulness constraints in (67) $\gg$ Markedness constraints in (68)

The relative ranking of the harmony constraints and the faithfulness constraints is based on the degree of acceptability caused when they are violated: violations of the harmony constraints are less acceptable than violations of faithfulness constraints. The ranking of the faithfulness constraints above the structural markedness constraints yields an honorific language like Korean, in which specifications for HON and LEVEL in the input are realized at the surface. Conversely, the ranking of the structural markedness constraints above the faithfulness constraints would yield a language like English limited in its expressibility by the absence of honorific expressions.

Now let us return to our original problem, i.e. word order freezing. Recall from section 3.2.1 that in Korean a dative subject triggering honorific agreement on the verb cannot be in a noninitial position. To derive this result we can apply harmonic alignment of the grammatical function scale (70a) with the case scale in (70b) and the structural position scale in (70c).

(70) Universal Scales:
  a. GF: SUBJ > NonSUBJ
  b. Case: NOM > OBL
  c. Position: Initial > Noninitial

The harmony scale in (71), derived through harmonic alignment, expresses the relative markedness of particular associations, for example, the fact that a nominative-marked subject is more harmonic (less marked) than an oblique-marked subject. The constraints that are derived by inverting the rankings of *SUBJ/Initial over *SUBJ/NInitial in (71) and prefixing the Avoid operator "*" are given in (72). For instance, the ranking of *SUBJ/NInitial over *SUBJ/Initial in (72b) means that in the absence of any relevant higher ranking constraint, a clause with a subject in a noninitial position will lose out in direct competition to a clause with a subject in the unmarked initial position.

(71) Harmony Scales derived through Harmonic Alignment
  a. H_5: SUBJ/NOM $\succ$ SUBJ/OBL
  b. H_6: SUBJ/Initial $\succ$ SUBJ/NInitial

(72) Constraint Subhierarchies:
  a. C_6: *SUBJ/OBL $\gg$ *SUBJ/NOM
  b. C_7: *SUBJ/NInitial $\gg$ *SUBJ/Initial

In order to capture the idea that if the oblique subject triggering honorific marking on the verb occurs in a noninitial position, it is the worst of the worst, again we can use local conjunction of existing constraints. Conjunction of the two subhierarchies in (72) results in two new constraint subhierarchies in (73).

(73) a. Conjoining *SUBJ/OBL with C_6:
  C_8: *SUBJ/OBJ & *SUBJ/NInitial $\gg$
  SUBJ/OBL & *SUBJ/Initial
  b. Conjoining *SUBJ/NOM with C_6:
  C_9: *SUBJ/NOM & *SUBJ/NInitial $\gg$
  SUBJ/NOM & *SUBJ/Initial

The high-ranked constraint in (73a) expresses the basic idea that if the subject is in a marked case (oblique case), it should not be in a marked non-initial position. But not every Korean construction with such a highly marked subject is excluded: such a construction is excluded only when the subject cooccurs with honorific agreement on the verb. What is left to do, then, is to make the constraints in (73) apply to the specific construction involving honorification. This is done through the conjunction of the markedness constraint *HON$_V$ in (68c) with the two subhierarchies in (73), as shown in (74).

(74) Conjunction of *Hon$_V$ with the two subhierarchies in (73)

a. *Hon$_V$ &
[*subj/obl &
*subj/NInitial]

b. *Hon$_V$ &
[*subj/nom &
*subj/NInitial]

c. *Hon$_V$ &
[*subj/obl &
*subj/Initial]

d. *Hon$_V$ &
[*subj/nom &
*subj/Initial]

A violation as in (a) represents the most marked situation, and (d) represents the least marked, with (b–c) as intermediate cases. For the current discussion, what is needed to rule out the ungrammatical sentences in (12b) (=(75a)) and (13b) is the topmost constraint in (74). This constraint expresses the idea that in case the verb is realized as an honorific form, the most marked configuration obtains when the subject is in oblique case and in the noninitial position of a clause. This is the only situation in which word order is frozen, and less marked configurations that are penalized by the low-ranked constraints are tolerated in the dative subject construction in Korean. This point is illustrated in (75) containing four sentences with the same sentential content 'Prof. Lee (honorified) needs a dog.'

(75)

	candidates and violations:		
a.	*Kay-ka	Lee kyoswunim-kkey	philyoha-si-ta.
	dog-nom	Prof. Lee-dat.hon	need-hon-decl
	Violation: (74a)		
b.	Kay-ka	Lee kyoswunim-kkeyse	philyoha-si-ta.
	dog-nom	Prof. Lee-nom.hon	need-hon-decl
	Violation: (74b)		
c.	Lee kyoswunim-kkey	kay-ka	philyoha-si-ta.
	Prof. Lee-dat.hon	dog-nom	need-hon-decl
	Violation: (74c)		
d.	Lee kyoswunim-kkeyse	kay-ka	philyoha-si-ta.
	Prof. Lee-nom.hon	dog-nom	need-hon-decl
	Violation: (74d)		

The constraints introduced so far are ranked in Korean as in (76).

(76) Ranking for Korean: *Hon$_V$ & [*subj/obl & *subj/NInitial] ≫ Harmony constraints ≫ Ident-IO(hon), Ident-IO(level), Top-L ≫ Canon$_{GF}$

The higher ranking of the topmost conjoined constraint in (74) over the information structuring constraint Top-L has the effect of restricting the word order freedom motivated by the topicality of arguments. There is no crucial ranking between the two faithfulness constraints and Top-L. What this ranking predicts is that in a clause with a highly marked honorific oblique subject in a noninitial position, scrambling of object over subject will not be possible as it results in violation of the top-ranked constraint *Hon$_V$ & [*subj/obl & *subj/NInitial], thus capturing the generalization that marked subject types must occur in unmarked position in the clause.

In a discourse context in which the theme is a topic, that is, [prom+, new–] and the experiencer is nonsalient background information, the input is as (77). This results in the optimal output (78a) with canonical sov order. (78) schematically represents only candidates which contain at least one element that expresses [hon+] and [level] and hence satisfy Ident-IO(hon) and Ident-IO(level). Also, only candidates with dative subjects are represented. Though they compete against each other in the universal candidate set, sentences with dative subjects and those with nominative subjects differ in the information status of their arguments, and each is more faithful to a different input. Violations of the top-ranked markedness constraint, which penalizes a highly marked subject, cause ungrammaticality in a strong sense, as indicated in (78e,g) (which are never optimal in any kind of context). On the other hand, examples involving violations of the Harmony constraints are generally infelicitous: they could be uttered in certain contexts.[27] Among the candidates violating the Harmony constraints, the candidates (b), (c), (f) and (g) are instances of 'disagreement' because the honorific markers are present only on the subject NP or on the verb, not on both. Hence, the candidate (a) is the best; it satisfies more higher-ranking constraints than any other candidates. As was the case of Hindi, the violation pattern for the top-ranked conjoined constraint for each candidate remains the same for all the other tableaux provided by different inputs, because this constraint is not sensitive to discourse context.

[27]Some of them may be appropriate in a context where the social status of the subject referent is placed between that of the speaker and that of the addressee.

(77) Input:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Prof. Lee'} \\ \text{HON} & + \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'dog'} \\ \text{PROM} & + \\ \text{NEW} & - \end{bmatrix}_y \\ \text{TNS} & \text{PRES} \\ \text{LEVEL} & 4 \\ \text{PRED} & \text{'need<P-A}_x\text{, P-P}_y\text{ >'} \end{bmatrix}$$

(P-A: discourse-neutral, P-P: topic)

(78) Tableau 4. Word order freezing in Korean

	CANDIDATES:	*Hon$_V$ & [*Subj/Obl & *Subj/NInitial]	Harmony	Top-L	Canon$_{GF}$
☞	a. S-dat.hon O-nom V-hon			*	
	b. S-dat.hon O-nom V-Ø		*	*	
	c. S-dat O-nom V-hon		*	*	
	d. S-dat O-nom V-Ø		**	*	
	e. O-nom S-dat.hon V-hon	!*			*
	f. O-nom S-dat.hon V-Ø		*		*
	g. O-nom S-dat V-hon	!*	*		*
	h. O-nom S-dat V-Ø		**		*

In summary, markedness constraints derived through harmonic alignment and local conjunction in OT provide interesting analyses of word order freezing in Hindi and Korean, a phenomenon which has remained unexplained in generative approaches to syntax. In the current OT account the generalizations about constituent order emerge naturally from interactions among the markedness constraints derived through harmonic alignment and local conjunction, and constraints on the realization of information structure and canonical ordering.

3.4.3 Typological and Theoretical Implications

Word order freezing in the situation of the 'worst of the worst' is not an idiosyncracy of Hindi and Korean. Clauses involving highly marked argument types consistently show restricted ordering possibilities in a variety of languages. For instance, in Kekchi, a Mayan language with canonical VOS order (Pinkerton 1976; Tomlin 1986), all six possible permutations of subject, object and verb are attested. However, when both the subject and object are animate and human, those noncanonical orders in which the object precedes the subject are precluded. When the subject is inanimate and the object is animate or human, those marked orders in which the first NP is the object and the second one is the subject are again precluded.[28]

Some Bantu languages are remarkable in precluding marked order in multiple object constructions. Logooli, a Luhya language spoken in Kenya, has rather free order of a beneficiary NP and a theme/patient NP if no ambiguity results (Morolong and Hyman 1977; Hyman and Duranti 1982; Siewierska 1988). When ambiguity might otherwise arise, for example, if both objects are animate, the thematically preferred beneficiary-theme order is selected. If, on the other hand, the prominence relations between animacy and thematic roles are reversed as beneficiary-nonhuman and theme-human, the animacy hierarchy determines word order: the order of the two is fixed with the theme preceding the beneficiary. Thus the marked association between the thematic role hierarchy and the animacy hierarchy is excluded in the marked order in which the less animate object precedes the more animate one, and surfaces only in the canonical order which conforms to the animacy hierarchy. Shona (Hawkinson and Hyman 1974) is another example which seems to belong in this class.

The effects of morphological case on word order that we have seen in Korean also show up in other languages with rich case morphology. For instance, in Icelandic with canonical SVO order marked orders such as OVS and VSO are permitted. However, not all marked orders are permitted in every case. In cases where the subject is dative and the object is nominative, OVS order is not possible: the dative argument has to precede the nominative one (Holmberg 1998). The impact of case marking on the relative order of arguments is particularly obvious in languages with dative shift (Primus 1998). While there exists considerable variation in factors influencing word order freezing, in all of the languages mentioned here, asymmetries in the word order patterns of arguments reflect exactly the same pattern: highly marked argument types involving the most marked associations of the prominence hierarchies can occur only in unmarked position.

[28] Other Mayan languages such as K'iche' (Mondloch 1981) and Kaqchikel (Broadwell 2000) show interesting interactions between word order and the hierarchies of grammatical function and definiteness. See Broadwell (2000) for an analysis of word order variation in Kaqchikel within the OT theory of markedness.

Thus under the markedness approach to word order taken here, word order freezing in languages with fairly free word order is not the exception, but can be derived from more general properties of the grammatical system, namely the markedness of argument types. Further, word order freezing can be seen as an instantiation of the general phenomenon of contextual neutralization often observed in phonology. For example, in some languages nasality is contrastive in vowels, except in the context before a nasal consonant, where all vowels are nasal (Kager 1999). This situation is produced by ranking (79), where the contextual markedness constraint *V$_{ORAL}$N dominates the faithfulness constraint Ident-IO(nasal), which in turn dominates the context-free markedness constraint *V$_{NASAL}$. This ranking states that nasal realization of vowels before nasal consonants takes priority over preservation of input values of [nasal], which in turn takes priority over the total orality of vowels. In sum, the nasality contrast in vowels is avoided in the marked position, i.e. before a nasal consonant.

(79) Contextual neutralization of nasality in vowels
(taken from Kager 1999:38–39)
Contextual markedness ≫ Faithfulness ≫ Context-free markedness
*V$_{ORAL}$N ≫ Ident-IO(nasal) ≫ *V$_{NASAL}$

This situation of neutralization of phonological contrasts in the marked position finds an interesting parallel in the syntactic domain of word order freezing. In languages with fairly free word order, noncanonical orderings are preferred options to mark a special information structure. However, as we have seen in the previous sections, under the special circumstances of markedness, they are replaced by the less marked, canonical order. This has been shown to be due to the ranking of the markedness constraints banning marked argument types in the marked positions and the information structuring constraints, which favor realization of contrasting prominence of arguments. The overriding of the information structuring constraints such as Top-L gives rise to contextual neutralization in word order: realization of contrasts in discourse prominence, even if otherwise preferred in the language, is avoided in the most marked argument types. Under the uniform framework of constraint interaction both in phonology and syntax, this striking parallel can be given a unified markedness explanation.

3.5 Bidirectional Optimization and Word Order Freezing

As mentioned in section 3.2, 'the worst of the worst' is not the only source of word order freezing. Arguments bearing identical case markings can be restricted to unmarked word order position in Hindi and Korean, and if their ordering is reversed, the meaning of the sentence cannot be maintained. In this section, I show that by extending optimization to comprehension as well as to production (Smolensky 1996b, 1998), the bidirectional model of optimization, based on the same constraint ranking, correctly predicts

that the type of scrambling leading to irrecoverability of the input content will not occur (section 3.5.1). Furthermore, I show how recovery of marked interpretation is dealt with (section 3.5.2) and address how the existence of ungrammatical and ambiguous strings can be captured (section 3.5.2).

3.5.1 Bidirectional Optimization and Recoverability

As already discussed for the Hindi double nominative examples, repeated below in (80), these sentences can only be unambiguously interpreted with the first NP as the agent, not as the patient (without strong contextual licensing of scrambled word order).

(80) a. Patt^{h}ar botal todegaa.
stone-NOM bottle-NOM break-FUT
(i) 'The stone will break the bottle.'
(ii) *'The bottle will break the stone.'

b. Botal patt^{h}ar todegaa.
bottle-NOM stone-NOM break-FUT
(i) 'The bottle will break the stone.'
(ii) *'The stone will break the stone.'

However, this type of word order freezing does not yet follow from the standard generation-based OT grammar sketched above. The obvious problem is overgeneration of ungrammatical scrambling in sentences with ambiguous case marking, more generally generation of structures from which the (original) meaning is not recoverable:[29] the standard OT grammar freely generates outputs like (80b.ii), where the interpretation of the arguments does not conform to the canonical word order. As illustrated in (82), the constraint ranking for Hindi proposed in sections 3.3 and 3.4.1 yields candidate (a2) with the starred reading (=(80b.ii)) as the optimal output from the input in (81) (the [NEW−] is omitted; [PROM+] is abbreviated [P+]).

(81) Input:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'stone'} \\ \text{P-ROLE} & \text{PROP} \quad \begin{bmatrix} \text{CAUS} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'bottle'} \\ \text{PROM} & + \end{bmatrix}_y \\ \text{TNS} & \text{FUT} \\ \text{PRED} & \text{'break } <\text{P-A}_x, \text{P-P}_y > \text{'} \end{bmatrix}$$

(P-A: discourse-neutral, P-P: topic)

[29] Motivations for developing bidirectional OT in other domains of linguistics will not be discussed here. A growing body of work shows that many of the motivations for the bidirectional approach to syntax are paralleled in phonology and semantics. Especially important among works developing a bidirectional OT for phonology and semantics in recognition of the problems of ambiguity and interpretational preferences are Boersma (1999), Beaver (2000) and Blutner (1999); see also Aissa and Fong (2000).

(82) Tableau 5. Double nominative construction in Hindi

		*SUBJ/P-P	DEP-IO(PROM)	TOP-L	CANON$_{GF}$
	CANDIDATES:				
	a1. botal$_{S/P-P[P+]}$ patt^{h}ar$_{O/P-A}$ todegaa	*!			
☞	a2. botal$_{O/P-P[P+]}$ patt^{h}ar$_{S/P-A}$ todegaa				*
	a3. botal$_{S/P-P}$ patt^{h}ar$_{O/P-A}$ todegaa	*!	*		
	a4. botal$_{O/P-P}$ patt^{h}ar$_{S/P-A}$ todegaa		*		*
	b1. patt^{h}ar$_{S/P-A}$ botal$_{O/P-P[P+]}$ todegaa			*	
	b2. patt^{h}ar$_{O/P-A}$ botal$_{S/P-P[P+]}$ todegaa	*!		*	*
	b3. patt^{h}ar$_{S/P-A}$ botal$_{O/P-P}$ todegaa		*		
	b4. patt^{h}ar$_{O/P-A}$ botal$_{S/P-P}$ todegaa	*!	*		*

Intuitively, if we are going to rule out winners in standard generation-based optimization associated with the interpretation that does not match the preferred reading of the string, then we need to have a formal method for allowing the output of generation-based optimization to be checked against the string corresponding to the syntactic parse. This can be achieved by extending optimization to comprehension (or parsing) as well as production (or generation) directions (Smolensky 1996b, 1998). I will call this extended model of optimization the *bidirectional optimization* model.[30] The relation between production-directed optimization and comprehension-directed optimization proposed in Smolensky (1998) is summarized in (83).

(83) Production/Comprehension Chain of /I/:
/I/—Prod → [S] — overt part ⟶ "O" — Comp →/I′/

/I/ is recoverable from "O", where /I/— Prod → [S], if "O" — Comp →/I/

Production-directed optimization starts with an input /I/ and gives the set of structural descriptions [S] of the input /I/ as an output, whereas comprehension-directed optimization starts out from an overt form "O", and gives all structural descriptions the overt part of which is "O" as an output. So, in syntax we can think of production as a function that takes an underspecified f-structure input to a set of fully specified (c- and f-structure) analyses, and comprehension as a function that takes the string

[30]The term 'bidirectional optimization' was coined by Wilson (1996, 1997, 2001), who proposes an OT account of anaphora binding, based on 'interpretive' optimization and 'expressive' optimization. Here I use the term bidirectional optimization model to refer to the production/comprehension model proposed by Smolensky.

part of the production output and gives a semantic content as an output (i.e. an underspecified f-structure).

This extended model of optimization has two important properties which play a key role: first, production-directed optimization is based on a candidate set with a common input content, whereas comprehension-directed optimization is based on a candidate set with a common string input. Second, the same grammar can be used for both comprehension and production (Boersma 1999; Hale and Reiss 1998).[31] I will now demonstrate how this extended model of comprehension, based on the same constraint ranking, resolves the problem of generation of structures from which the meaning is not recoverable, thus explaining the emergence of the unmarked effect in word order.[32]

The tableaux in (84) and (85) represent the production and comprehension half of the bidirectional model of optimization respectively. Observe the differences in competitor sets in production and comprehension. As shown in (84), in production, what competes are candidates that share the particular semantic form or input. Hence the candidates where *botal* 'bottle' is the agent and *patt^h ar* 'stone' is the patient are not competitors in production as they are not faithful to the argument structure semantics. Those which are excluded by the high-ranked constraint Ident-IO(p-role), which requires the value of the proto-role features in the input is preserved in the output feature structure (here [caus+]), are shaded out in (84).

In contrast, in comprehension of the string *botal patt^h ar todegaa* (i.e. the overt part of the production output), only the candidates which conform to the input string are competing structures. Such candidates that do not share the same string are shaded out again in the tableau in (85).

[31] At first glance, it appears that bidirectional OT might be more complex than unidirectional OT. While the complexity and decidability of bidirectional OT and an OT system as a whole are still open issues, several significant results about these issues have been reported. For a detailed discussion of the formal properties and the issues of the complexity and decidability of OT (both unidirectional and bidirectional), the reader is referred to Jäger (2000) and Kuhn (1999, 2000a,b, 2001). In particular, Jäger (2000) has argued that bidirectional optimization can be modeled by means of finite state techniques.

[32] Kuhn (1999, 2001) independently establishes bidirectional optimization as a formal explanation for word order freezing in German. For recent works applying bidirectional optimization to other ambiguity-related phenomena in syntax, see Asudeh (2001), Donohue (1999), Kuhn (2000a), among others.

(84) Tableau 6. Production-directed optimization

INPUT = (81)	Ident-IO(p-role)	*subj/P-P	Dep-IO(prom)	Top-L	Canon$_{GF}$
a1.	*!				
a2.	*!	*!			*
a3.		*!			
☞ a4.					*
b1.	*!		*		
b2.	*!	*!	*		*
b3.		*!	*		
b4.			*		*
c1.				*	
c2.		*!		*	*
c3.	*!	*!		*	
c4.	*!			*	*
d1.			*		
d2.		*!	*		*
d3.	*!	*!	*		
d4.	*!		*		*

a1. = botal$_{S/P-A[P+]}$ patt^{h}ar$_{O/P-P}$ todegaa
a2. = botal$_{O/P-A[P+]}$ patt^{h}ar$_{S/P-P}$ todegaa
a3. = botal$_{S/P-P[P+]}$ patt^{h}ar$_{O/P-A}$ todegaa
a4. = botal$_{O/P-P[P+]}$ patt^{h}ar$_{S/P-A}$ todegaa
b1. = botal$_{S/P-A}$ patt^{h}ar$_{O/P-P}$ todegaa
b2. = botal$_{O/P-A}$ patt^{h}ar$_{S/P-P}$ todegaa
b3. = botal$_{S/P-P}$ patt^{h}ar$_{O/P-A}$ todegaa
b4. = botal$_{O/P-P}$ patt^{h}ar$_{S/P-A}$ todegaa
c1. = patt^{h}ar$_{S/P-A}$ botal$_{O/P-P[P+]}$ todegaa
c2. = patt^{h}ar$_{O/P-A}$ botal$_{S/P-P[P+]}$ todegaa
c3. = patt^{h}ar$_{S/P-P}$ botal$_{O/P-A[P+]}$ todegaa
c4. = patt^{h}ar$_{O/P-P}$ botal$_{S/P-A[P+]}$ todegaa
d1. = patt^{h}ar$_{S/P-A}$ botal$_{S/P-P}$ todegaa
d2. = patt^{h}ar$_{O/P-A}$ botal$_{S/P-P}$ todegaa
d3. = patt^{h}ar$_{S/P-P}$ botal$_{O/P-A}$ todegaa
d4. = patt^{h}ar$_{O/P-P}$ botal$_{S/P-A}$ todegaa

(85) Tableau 7. Comprehension-directed optimization

	INPUT = (84b)	IDENT-IO(P-ROLE)	*SUBJ/P-P	DEP-IO(PROM)	TOP-L	CANON$_{GF}$
	a1.			*		
	a2.		*!	*		*
	a3.		*!	*		
	a4.			*		*
☞	b1.					
	b2.		*!			*
	b3.		*!			
	b4.					*
	c1.			*	*	
	c2.		*!	*	*	*
	c3.		*!	*	*	
	c4.			*	*	*
	d1.					
	d2.		*!			*
	d3.		*!			
	d4.					*

a1. = botal$_{S/P-A[P+]}$ patt^{h}ar$_{O/P-P}$ todegaa
a2. = botal$_{O/P-A[P+]}$ patt^{h}ar$_{S/P-P}$ todegaa
a3. = botal$_{S/P-P[P+]}$ patt^{h}ar$_{O/P-A}$ todegaa
a4. = botal$_{O/P-P[P+]}$ patt^{h}ar$_{S/P-A}$ todegaa
b1. = botal$_{S/P-A}$ patt^{h}ar$_{O/P-P}$ todegaa
b2. = botal$_{O/P-A}$ patt^{h}ar$_{S/P-P}$ todegaa
b3. = botal$_{S/P-P}$ patt^{h}ar$_{O/P-A}$ todegaa
b4. = botal$_{O/P-P}$ patt^{h}ar$_{S/P-A}$ todegaa
c1. = patt^{h}ar$_{S/P-A}$ botal$_{O/P-P[P+]}$ todegaa
c2. = patt^{h}ar$_{O/P-A}$ botal$_{S/P-P[P+]}$ todegaa
c3. = patt^{h}ar$_{S/P-P}$ botal$_{O/P-A[P+]}$ todegaa
c4. = patt^{h}ar$_{O/P-P}$ botal$_{S/P-A[P+]}$ todegaa
d1. = patt^{h}ar$_{S/P-A}$ botal$_{S/P-P}$ todegaa
d2. = patt^{h}ar$_{O/P-A}$ botal$_{S/P-P}$ todegaa
d3. = patt^{h}ar$_{S/P-P}$ botal$_{O/P-A}$ todegaa
d4. = patt^{h}ar$_{O/P-P}$ botal$_{S/P-A}$ todegaa

An additional difference between production and comprehension is that the candidates where [PROM+] is present on *botal* 'botal' are eliminated by DEP-IO(PROM), as this information is not present in the input (now the string, not a morphosyntactic and semantic content). Note that a different candidate from (84) (i.e. the candidate (b1) here), violating none of the constraints under consideration, is selected as the winner in the comprehension direction. Thus by bidirectional optimization we correctly derive the emergence of the unmarked effect (McCarthy and Prince 1994) in word order: the canonical SOV order emerges as the unmarked case in a null or neutral context, where there is no motivation for noncanonical orderings.

The optimal candidate (b1) has the meaning shown in (86), which is different from the input f-structure (81).[33] This means that the input f-structure we started from is not the most harmonic meaning for the string corresponding to the winning candidate in the production direction, and hence we can consider that candidate ungrammatical under bidirectional optimization as it does not provide recoverability for the original input.

(86) Recovered semantic content:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'bottle'} \\ \text{P-ROLE} & \text{PROP} \quad \begin{bmatrix} \text{CAUS} & + \end{bmatrix} \end{bmatrix}_y \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'stone'} \\ \text{PROM} & + \end{bmatrix}_x \\ \text{TNS} & \text{FUT} \\ \text{PRED} & \text{'break<P-A}_y\text{,P-P}_x\text{ >'} \end{bmatrix}$$

Similar results can be shown for the double nominative construction in Korean and the constructions containing multiple nominals with *-ko* and *-se* in Hindi (see section 3.2.2): in these cases too, the meaning recovered in the comprehension direction by the output is not identical to the input for the production grammar. Specifically in the case of Hindi, CANON_{θ}, the constraint ranked lower than TOP-L (but higher than CANON_{GF}), emerges as decisive in comprehension-directed optimization, forcing the choice of the canonical order (Agent-Source-Theme-V) determined by the thematic role hierarchy in sentences containing multiple nominals with *-ko* and *-se*. I will not go through these other results here.

[33] Several steps are involved in getting (86) from the output c- and f-structures of production-directed optimization. First, the function *yield* (Kuhn, 2001) applies to a c-structure and returns the string of terminal symbols; from the string input to comprehension-directed optimization, the GEN yields a set of candidate c- and f-structure analyses; the underspecified f-structure like the one in (86) is obtained by applying to the candidate f-structure of comprehension the filtering function *F*, which takes a fully specified f-structure to an underspecified f-structure. See Kuhn (1999, 2001) for a more detailed discussion of bidirectional competition and parsing analysis in OT-LFG.

3.5.2 Further Consequences of Bidirectional Optimization

In this section we will examine how contextual effects on recovery of grammatical relations and the ambiguity of a string can be captured in the bidirectional optimization model.

Recovery of Marked Interpretation

In Hindi (and Korean) the order of the two arguments bearing identical case markings is fixed in SOV order in a null or neutral context. Word order can be said to have a disambiguating function: it is the fixed word order that determines which phrase is the subject and which is the object.

This freezing effect operative in sentences with ambiguous case marking, unlike the 'worst of the worst' type examined earlier, is in fact overridden by contextualization. A further consequence of the bidirectional approach is that it makes room for context-directed disambiguation (Joan Bresnan, p.c., March 1999). That is, if discourse context, verb meaning or use of intonation, stress, morphological materials distinct from case markers, etc. disambiguate the grammatical functions or thematic roles of nominals, the recoverability problem is predicted not to arise, and the relevant OT constraints (e.g., Top-L) come into play in selecting the best candidate. This section examines how the effects of an increase of such information (available clues that activate faithfulness constraints and information structuring constraints) on the construction of grammatical analyses can be integrated into bidirectional optimization and how this integration captures recovery of the marked interpretation of an ambiguous string.

Suppose the following sentence in (87) was uttered in a context where *botal* 'bottle' is highly topical and *patt^{h}ar* 'stone' is a causer. Although case forms on the two arguments are identical, the preceding discourse context clearly overrules the ordering preferences (e.g., basic SOV): the initial nominative argument is interpreted as the topicalized theme object.

(87) Botal patt^{h}ar todegaa.
bottle-NOM stone-NOM break-FUT
'The stone will break the bottle.'

The input that results in object initial candidates like (87) in production is identical to (81) and EVAL proceeds as in (84). Given the high ranking of the faithfulness constraints and Top-L above Canon$_{GF}$, the optimal output is always the candidate which correctly parses the argument structure semantics and [PROM+] present on the theme argument in the input.

However, in order for a candidate with noncanonical order (e.g., OSV) to be selected as a winner in the comprehension grammar as well, one modification of Smolensky's model is needed, because bidirectional optimization in Smolensky's model gives only the unmarked order for any string involving ambiguous case marking. The source of the problem is the impoverishment of the input: the input to comprehension, taken to consist

of bare strings in Smolensky's production/comprehension model, lacks all information (other than strings) available to language users that activates the constraints on the realization of information structure and the proto-role properties of arguments, and as a result, these constraints can never be active in comprehension, if the input consists solely of bare strings. To solve this problem, we need to assume that the input can no longer be regarded as consisting solely of the string. A representation of the contextual and proto-role information is now also part of this input, and it can be formally modeled as an underspecified feature structure indicating the information status of the string referents. In other words, when context that brings out the marked or scrambled interpretation is supplied to comprehension-directed optimization, the string input to comprehension is enriched with interpretational features, just as the f-structure input to the production direction contains information about each element's discourse status and proto-role information. This additional information plays a role in selecting the optimal analysis of the string, by activating the Ident-IO(p-role) constraint and the Dep-IO(prom) constraint.

Another way of treating context-defeating word order freezing would be to say that the SOV restriction on nominal arguments bearing identical case marking may not be syntactic at all[34] and that other overriding factors are also beyond syntactic considerations. However, I believe that there are deep theoretical and empirical motivations for why a theory of linguistic competence needs to model the integration of "nonlinguistic" information.[35] Here I will mention three interrelated but distinct reasons.

First, our refined model of bidirectional optimization provides a highly flexible framework for having syntactic and other extra-syntactic factors interact in constructing grammatical analyses. Research on language processing and learning has suggested that language users and learners are extremely sensitive to information from discourse context, frequency biases and plausibility information and that knowledge of discourse context

[34] An argument against the view that freezing effects lie outside the realm of syntax and are merely the result of performance was detailed in Bloom (1999). He demonstrated that in Russian clauses with two syncretized nominal arguments are still frozen into SVO (in non-emotive speech), even when we have enough information from context or verbal morphology to resolve the grammatical roles of the syncretized arguments.

[35] E. Prince (2000) makes a distinction between two kinds of linguistic intuition: 1) the unconscious, inaccessible ability to process utterances of one's language; 2) meta-linguistic intuition, which is conscious, accessible meta-intuition about linguistic intuitions. The latter type includes speakers' knowledge of social/situational and discourse contexts in which an utterance may be used felicitously. The traditional terms 'linguistic' and 'nonlinguistic' information I use here correspond to the kind of information accessed by our linguistic and meta-linguistic intuitions respectively. It should be emphasized that our refined model of bidirectional optimization is not trying to model performance or perceptual strategies. Nor is it proposing to blur the distinction between the two kinds of linguistic knowledge. Rather what it is trying to model are the effects of various kinds of linguistic information, broadly including discourse context, on the construction of grammatical analyses, which interact in parallel.

and the statistical properties of input is part of linguistic knowledge. So if a grammar is to be compatible with models of sentence processing and learning, then it needs to formally model linguistic knowledge in a way that allows for the dynamic interaction between syntactic and extra-syntactic factors in the construction of grammatical analyses.

Second, another advantage of including a context representation as part of the input is that it allows us to capture the symmetry between the production grammar and comprehension grammar. Research on language processing has shown us that linguistic knowledge is process-independent and that the kind of linguistic description that linguistic theory provides should therefore be a process-neutral grammar.[36] Note that in our model of bidirectional optimization the representation of contextual and proto-role information is part of the input both in production-directed and in comprehension-directed optimization. This refinement of the model has the effect of allowing faithfulness constraints sensitive to contextual and semantic information as well as markedness constraints to apply simultaneously to representations of linguistic structures in both production and comprehension. Thus, in our model of grammar, both production and comprehension can be viewed in terms of a process of satisfaction of the same set of competing constraints. The process-neutrality of our linguistic description seems very suggestive, given that linguistic knowledge is process-independent.

Third, in addition to modeling the symmetry between the production and comprehension grammar, the present model of extended optimization offers the formal integration to syntax and phonology. As Boersma (1999) independently observes, contextual/semantic information is also needed to account for phonological acquisition and the interaction between phonology and semantics. As an example from phonology parallel to the case discussed above, consider the case of final devoicing in Dutch, which causes the two words *rad* 'wheel' and *rat* 'rat' to merge on the surface. Smolensky's comprehension model cannot account for the fact that if the semantic context is 'turn', the recognition of 'wheel' is favored over that of 'rat', as far as the lexicon is concerned—i.e. the fact that infrequent and contextually disfavored items are hard to access. Boersma (1999) therefore proposes to include the semantic context as part of the input to the recognition grammar (e.g., input: [rat], context = 'turn') and to include in the recognition grammar lexical-access constraints whose rankings depend on the semantic context and on frequency of occurrence. So by including a formal representation of the extra-sentential context as part of the input both in comprehension-directed and production-directed optimization, the present

[36]For further discussion of how a declarative system of constraints can best ensure process-neutrality, see Bresnan and Kaplan (1982), Halvorsen (1983), Pollard and Sag (1994), Smolensky (1996b) and Sag and Wasow (1999).

model of bidirectional optimization can capture the symmetry between phonology-semantics interactions and syntax-semantics interactions.[37]

Now, let us examine how recovery of marked structure and interpretation is captured in our refined model of bidirectional optimization. Suppose that the argument *botal* 'bottle' of the verb *todegaa* 'break', for instance, is presented prominently in the context where *patt^{h}ar* 'stone' causes the breaking event. In this context the input to comprehension is like (88).

(88) Input:

Botal	patt^{h}ar	todegaa
[PROM+] (TOP)	[P-ROLE PROP [CAUS+]] (P-A)	

The input in (88) feeds the EVAL process, shown in (89) (here and in subsequent tableaux, I show only candidates which satisfy the linking constraints, for simplicity). The [PROM+] and [CAUS+] information in the string input then will activate the constraints on faithfulness to proto-role information (IDENT-IO(P-ROLE)) and on discourse information (DEP-IO(PROM)), and the information structuring constraints, which are not in effect when no such information is present in the input string, as we saw above.

(89) Tableau 8. Comprehension-directed optimization in the double nominative construction in a topical context

	CANDIDATES:	IDENT-IO(P-ROLE)	DEP-IO(PROM)	TOP-L	CANON$_{GF}$
	a1. botal$_{S/P-A[P+]}$ patt^{h}ar$_{O/P-P}$ todegaa	*!			
☞	a2. botal$_{O/P-P[P+]}$ patt^{h}ar$_{S/P-A}$ todegaa				*
	b1. botal$_{S/P-A}$ patt^{h}ar$_{O/P-P}$ todegaa	*!	*		
	b2. botal$_{O/P-P}$ patt^{h}ar$_{S/P-A}$ todegaa		*		*
	c1. patt^{h}ar$_{S/P-A}$ botal$_{O/P-P[P+]}$ todegaa			*	
	c2. patt^{h}ar$_{O/P-P}$ botal$_{S/P-A[P+]}$ todegaa	*!		*	*
	d1. patt^{h}ar$_{S/P-A}$ botal$_{S/P-P}$ todegaa		*		
	d2. patt^{h}ar$_{O/P-P}$ botal$_{S/P-A}$ todegaa	*!	*		*

[37] Enrichment of the input in comprehension is also justified by computational considerations. For parsing/comprehension tasks, a given string is parsed to arrive at possible structures (to which production-directed optimization can apply). To ensure decidability of the parsing/comprehension task and to make constraint violations more detectable, the bidirectional model needs to be further constrained, and one way to do this is to take a context representation into account (Jonas Kuhn, p.c., March 2000).

The crucial constraints here are the faithfulness constraints: any candidate that has *botal* 'bottle' as agent and *patt^{h}ar* 'stone' as patient eliminated by Ident-IO(p-role). Candidates (b1) and (b2) show that the [prom+] information, if present in the input, must be realized in the output, or else offending candidates are eliminated by Dep-IO(prom). Note here that the same candidates have no violation of Dep-IO(prom) in (85) since there is no [prom] feature present in the input that must be realized in a null context. Lastly, candidates (a1) and (a2), ruled out by Dep-IO(prom) in a null context, incur no violation of Dep-IO(prom) as they correctly realize the [prom+] information in the string input on the patient. So, candidate (a2) is the optimal output, under the ranking shown. This recovers the semantic content in (90), which is identical to the input in (81). The reader can also see in tableau (89) that the winner in the comprehension grammar and the winner in the production grammar ((84)) are identical.

(90) Recovered semantic content:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'stone'} \\ \text{P-ROLE} & \text{PROP} \quad \begin{bmatrix} \text{CAUS} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'bottle'} \\ \text{PROM} & + \end{bmatrix}_y \\ \text{TNS} & \text{FUT} \\ \text{PRED} & \text{'break <P-A}_x\text{,P-P}_y\text{ >'} \end{bmatrix}$$

We have so far paid attention only to sentences which can have both SOV and OSV readings, depending on the context. In some cases, however, only the OSV reading is available. For instance, in the Korean sentences (91) and (92), which involve nominative object scrambling, the bracketed part can only be interpreted with the first nominative NP as theme (mapped onto object) and the second one as experiencer (mapped onto subject).

(91) Computer-ka philyoha-n-tey, [**VCR-i** aitul-i
computer-NOM need-PRES-but VCR-NOM kids-NOM
philyoha-ta-ko] hay-se VCR-ul sa-ss-ta.
need-DECL-COMP say-therefore VCR-ACC buy-PAST-DECL
'I need a computer, but a VCR, my kids said that they need, and therefore I bought a VCR.'

(92) Cinan cwu say coffee shop-i mwun-ul yel-ess-nuntey,
last week new coffee shop-NOM door-ACC open-PAST-and
ku coffee shop-uy [**coffee-ka** chinkwu-ka coh-ta-ko]
that coffee shop-GEN coffee-NOM friend-NOM like-DECL-COMP
hay-ss-ta.
say-PAST-DECL
'Last week a new coffee shop opened. Coffee at that coffee shop, my friend said that he/she likes.'

It is important to note that the same constraint system works in these cases. In particular, the constraint ranking Ident-IO(p-role) $\gg$ Canon$_{GF}$ correctly yields the osv order as the optimal output in both production and comprehension: due to the high-ranking constraint Ident-IO(p-role), candidates with the canonical sov (or experiencer-theme-V) order will be eliminated, since they have a non-sentient argument (i.e. 'VCR' in (91) and 'coffee' in (92)) as the subject and sentient argument (i.e. 'kids' in (91) and 'my friend' in (92)) as the object. Only candidates with the noncanonical osv (or theme-experiencer-V) order will correctly parse the proto-role property of [sentience±] associated with each argument of the predicates used in (91) and (92), satisfying the high-ranked constraint Ident-IO(p-role) (and also the constraints on the proto-role properties-to-grammatical function linking).

Thus, by enriching the input string with semantic and contextual information, the present ot account captures the fact that with an increase of information that activates the constraints on the realization of arguments and information structure, the recoverability problem observed in a neutral context does not arise when the arguments differ in information status or proto-role properties.

Lastly, let us address the question of how bidirectional optimization can be correctly applied to the examples with unambiguous case marking, and has an effect only with the ambiguous ones. In contrast to the double nominative construction considered above, the recoverability problem does not occur in sentences with unambiguous case marking (e.g., ergative-nominative or ergative-accusative in Hindi) even without taking contextual information into consideration. For example, any permutation of the three elements in (93a) does not change the meaning of the sentence.

(93) a. Anuu-ne caand dek^{h}aa. (=(1a))
Anu-erg moon-nom see/look at-perf
'Anu saw the moon.'

b. Caand Anuu-ne dek^{h}aa. (=(1b))
moon-nom Anu-erg see/look at-perf
'Anu saw the moon.'

Consider the production half first, illustrated in (94) and (95) below. Tableau (95) shows that candidate (a2) (=(93b)) with osv order becomes optimal for expressing the topical status of the theme *caand* 'moon' and the volitionality of the experiencer 'Anu'.

(94) Input:

$$\begin{bmatrix} \text{GF1} & \begin{bmatrix} \text{PRED} & \text{'Anu'} \\ \text{P-ROLE} & \text{PROP} \quad \begin{bmatrix} \text{VOL} & + \end{bmatrix} \end{bmatrix}_x \\ \text{GF2} & \begin{bmatrix} \text{PRED} & \text{'moon'} \\ \text{PROM} & + \end{bmatrix}_y \\ \text{ASP} & \text{PERF} \\ \text{PRED} & \text{'see <P-A}_x\text{,P-P}_y\text{>'} \end{bmatrix}$$

(P-A: discourse neutral, P-P: topic)

(95) Tableau 9. Production-directed optimization in standard case frame

		Ident-IO(p-role)	Dep-IO(prom)	Top-L	Canon$_{GF}$
	CANDIDATES:				
	a1. caand$_{S/P-A[P+]}$ Anuu-ne$_{O/P-P}$ dek^{h}aa	*!			
☞	a2. caand$_{O/P-P[P+]}$ Anuu-ne$_{S/P-A}$ dek^{h}aa				*
	b1. caand$_{S/P-A}$ Anuu-ne$_{O/P-P}$ dek^{h}aa	*!	*		
	b2. caand$_{O/P-P}$ Anuu-ne$_{S/P-A}$ dek^{h}aa		*		*
	c1. Anuu-ne$_{S/P-A}$ caand$_{O/P-P[P+]}$ dek^{h}aa			*	
	c2. Anuu-ne$_{O/P-P}$ caand$_{S/P-A[P+]}$ dek^{h}aa	*!		*	*
	d1. Anuu-ne$_{S/P-A}$ caand$_{S/P-P}$ dek^{h}aa		*		
	d2. Anuu-ne$_{O/P-P}$ caand$_{S/P-A}$ dek^{h}aa	*!	*		*

When the string *caand Anuu-ne dek^{h} aa* is uttered in a context in which *caand* 'moon' is topical, the input is like (96). The [VOL+] information can be inferred from the context, but even without any contextual clue, 'Anuu' must be interpreted as the volitional experiencer and *caand* 'moon' as the theme. This is because the ergative case marker *-ne* independently carries volitionality information about the subject of the clause.

(96) Input:

Caand	Anuu-ne	dek^{h}aa
[PROM+]	[P-ROLE PROP [VOL+]]	
(TOP)	(P-A)	

As in the examples involving ambiguous case-marking discussed above, the same constraint system works in comprehension of the string in (96). Observe again the key role played by the high-ranked faithfulness constraints in selecting candidate (a2) with OSV order in (97) as the optimal output.

It correctly realizes the Proto-Agent property of volitionality contributed by the case morphology as well as the topicality of *caand* 'moon'.

(97) Tableau 10. Comprehension-directed optimization in standard case frame

		Ident-IO(p-role)	Dep-IO(prom)	Top-L	Canon$_{GF}$
	candidates:				
	a1. caand$_{S/P-A[P+]}$ Anuu-ne$_{O/P-P}$ dek^{h}aa	*!			
☞	a2. caand$_{O/P-P[P+]}$ Anuu-ne$_{S/P-A}$ dek^{h}aa				*
	b1. caand$_{S/P-A}$ Anuu-ne$_{O/P-P}$ dek^{h}aa	*!	*		
	b2. caand$_{O/P-P}$ Anuu-ne$_{S/P-A}$ dek^{h}aa		*		*
	c1. Anuu-ne$_{S/P-A}$ caand$_{O/P-P[P+]}$ dek^{h}aa			*	
	c2. Anuu-ne$_{O/P-P}$ caand$_{S/P-A[P+]}$ dek^{h}aa	*!		*	*
	d1. Anuu-ne$_{S/P-A}$ caand$_{S/P-P}$ dek^{h}aa		*		
	d2. Anuu-ne$_{O/P-P}$ caand$_{S/P-A}$ dek^{h}aa	*!	*		*

Again the production and comprehension processes yield an identical winner, and hence it is predicted that the recoverability problem will not occur in sentences with unambiguous case marking, since the winners in the two processes recover the identical semantic content. Therefore this analysis correctly accounts for the fact that the interpretation of the sentences with unambiguous case marking in terms of thematic roles does not rely on word order, but is driven by the overt case marking.

In summary, I have argued that word order freezing in sentences with ambiguous case marking in Hindi and Korean can be explained when the constraint system allows an extension to bidirectional competition. I have also shown that by enriching the string input to comprehension with semantic and contextual information, the bidirectional approach, based on the same set of constraints, predicts that preferences for canonical ordering are overridden by faithfulness constraints on discourse prominence and proto-role information, which outrank markedness constraints against non-canonical GF order in both the production and comprehension grammars.

Capturing Ungrammatical and Ambiguous Strings

In the current OT constraint system, ungrammaticality arises from a failure to recover the input from the output. Further, the constraint system predicts whether a given string in the language is ambiguous or not. Recall from section 3.2.1 that Hindi avoids a possible (parsing) ambiguity in

order to uniquely identify the subject and object of a nonvolitional transitive and that the ungrammaticality of the string in (98b), associated with the marked structure, is part of this larger generalization about ambiguity avoidance. In this section, I discuss how the constraint system can capture this generalization from a bidirectional perspective.

(98) Grammaticality of a string with theme-experiencer-V order

a. Niinaa Anuu-ko uskii bastii-mẽ dik^{h}ii. (SOV)
Nina-NOM Anu-DAT PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{i/*j}$ neighborhood.'

b. *Niinaa Anuu-ko uskii bastii-mẽ dik^{h}ii. (OSV)
Nina-NOM Anu-DAT PRON-GEN neighborhood-LOC appear-PERF
'Anu$_i$ saw Nina$_j$ in her$_{j/*i}$ neighborhood.'

Let us first consider the the production-directed optimization in (99). The input is identical to the f-structure in (53) (the two candidates (a) and (b) shown in (99) correspond to (b2) and (a2) in ((54) and (59) respectively)). In this tableau, the high-ranking constraints do not distinguish the candidates[38], and it is the lower-ranking constraint Canon$_{GF}$ that breaks the tie and favors the less marked candidate with SOV order.

(99) Tableau 11. Production-directed optimization in Hindi nonvolitional transitives

INPUT=(53)	*SUBJ/P-P & *SUBJ/Nimitial	*NSUBJ/P-A$_{-vol}$ & Spine-R	TOP-L	CANON$_{GF}$
☞ a. Niinaa$_{S/P-P}$Anuu-ko$_{O/P-A}$dik^{h}ii				
b. *Niinaa$_{O/P-P}$Anuu-ko$_{S/P-A}$dik^{h}ii				*

[38]One of the key features of grammatical function selection in Hindi (and Marathi) is subject/object alternation in verb classes that do not take a volitional argument. If there is no volitional argument (more precisely, an argument lacking the property of 'conscious choice' (Mohanan 1994a)) and there is more than one argument that has a proto-patient property (e.g., argument not entailed to possess sentience), there will be a subject/object alternation. Here, I am assuming that it is variable constraint rankings as proposed by Asudeh (2001) which yield the subject/object alternation for nonvolitional transitives. I have omitted the floating linking constraints having variable ranking values (Boersma 1997; Asudeh, 2001) in the tableaux in (99) and (100) since they are too complicated to be treated here. It is sufficient to note that these floating constraints on the semantic role–grammatical function linking also equally (dis)favor the two candidates under consideration in the tableaux in (99) and (100).

The tableau in (100) illustrates how the same constraint ranking yields the candidate with SOV order as the optimal output in comprehension. Once again, the $\textsc{Canon}_{GF}$ constraint operates decisively to correctly identify the first argument as the subject and hence serves to rule out the marked structure in the absence of any difference in the volitionality of the two arguments.

(100) Tableau 12. Comprehension-directed optimization in Hindi nonvolitional transitives

INPUT: *Niinaa Anuu-ko dik^{h}ii*	*SUBJ/P-P & *SUBJ/Ninitial	*NSUBJ/P-A$_{-vol}$ & Spine-R	TOP-L	CANON$_{GF}$
☞ a. Niinaa$_{S/P-P}$Anuu-ko$_{O/P-A}$dik^{h}ii				
b. *Niinaa$_{O/P-P}$Anuu-ko$_{S/P-A}$dik^{h}ii				*

Thus, in short, in the tableaux here, we can see that of the available structural alternatives sharing the same string expression of a Hindi nonvolitional transitive, what emerges as the optimal output is the unmarked structure that conforms to canonical SOV order; the ungrammaticality of the OSV analysis of the string in (100) follows from its relative markedness compared to the other parsing alternative that competes with it.

However, while word order, morphology, verb meaning and context can eliminate some ambiguity in the identification of the grammatical role of arguments, they do not remove all ambiguity. Given that all of the cases that we have been examining so far involve only one winning candidate, the existence of ambiguous strings becomes of interest. Before I conclude this section, I exemplify some of the cases where ambiguity still remains with certain verbs.

Passives of ditransitives in Hindi are another example of a verb class that shows an alternation in which the argument can be realized as the subject or object. These verbs are similar to nonvolitional transitive verbs that we have been examining in that the grammatical function realization of their (internal) arguments is also determined by word order in the theme-initial clauses like (101a) (Mohanan 1992; Mohanan and Mohanan 1994). Unlike nonvolitional transitives, however, goal-initial passive sentences as in (101b) do receive more than one grammatical function analysis. That is, they are ambiguous regarding which (internal) argument is assigned to the

subject function and which is assigned to the object function, though the sentential meaning is constant.

(101) Parsing ambiguity in passives of ditransitives in Hindi (Mohanan 1992)

a. Baccaa Mohan-ko diyaa gayaa.
child-NOM Mohan-DAT give-PERF go-PERF
(i) 'The child was given to Mohan.' (SUBJ: theme, OBJ: goal)
(ii) *'Mohan was given a child.' (*SUBJ: goal, OBJ: theme)

b. Mohan-ko baccaa diyaa gayaa.
Mohan-DAT child-NOM give-PERF go-PERF
(i) 'Mohan was given a child.' (SUBJ: goal, OBJ: theme)
(ii) 'The child was given to Mohan.' (SUBJ: theme, OBJ: goal)

In some languages, true semantic ambiguity is observed. For example, in the Dogon language Donno Sɔ (Culy 1995), which is both head- and dependent-marking, ambiguity arises in ditransitives under special circumstances. In Donno Sɔ only objects are case-marked but not subjects, with which finite verbs in matrix clauses agree in person and number. Culy (1995) discovered that object case marking in Donno Sɔ is subject to various ordered conditions such as animacy, semantic role, pronominality, definiteness and ambiguity. The strongest condition among these is the animacy restriction: if the goal/recipient and theme objects of a ditransitive verb are both human, it is the goal/recipient that must occur with case marking. Thus, there is potentially some ambiguity if the subject and the theme object of a ditransitive are equal in animacy and both refer to arguments of the same number. In these instances, the clause will be ambiguous as to which arguments are subject and object, as seen in (102) (I am glossing the case OM ('object marker'), following Culy (1995)). In each case, either the first or second NP can be interpreted as the subject. What is somewhat unusual is that such sentences are ambiguous only when the subject and theme object are not separated by the goal/recipient.

(102) Ambiguity in ditransitives in Donno Sɔ (Culy 1995:57)

a. Yaana I wojinɛ anna pay-ñ tagaa be.
female child stranger male old-OM showed AUX
'A girl showed a stranger to an old man.' or
'A stranger showed a girl to an old man.'

b. Wojinɛ yaana I anna pay-ñ tagaa be.
stranger female child male old-OM showed AUX
'A stranger showed a girl to an old man.' or
'A girl showed a stranger to an old man.'

Sesotho, a Bantu language spoken in Lesotho and adjacent areas in South Africa (Morolong and Hyman 1977; Hyman and Duranti 1982) is another interesting example of a language that shows semantic ambiguity. In this language, when both objects of a ditransitive verb (or a beneficiary

applied verb) are human, both word orders are possible with potential ambiguity, as shown in (103). Both sentences in (103) have two meanings which involve the exact reversal of the thematic role interpretation of the two objects (beneficiary and theme).

(103) Ambiguity in ditransitives in Sesotho
(Morolong and Hyman 1977:203)
a. Ke-bítselítsé morena baná.
I-called chief children
'I called the children for the chief.' or 'I called the chief for the children.'
b. Ke-bítselítsé baná morena.
I-called children chief
'I called the chief for the children.' or 'I called the children for the chief.'

The effect of animacy on the interpretation of arguments shows up in many other languages. Here I briefly discuss the case of Russian, in which the word order is fixed as SVO, as in Hindi and Korean, if the case markings on both the subject and object NPs are identical. The classic examples, taken from Jakobson (1963), are shown in (104).

(104) Case syncretism and word order freezing in Russian
a. Mat' ljubit doč'.
mother-NOM/ACC loves daughter-NOM/ACC
'The mother loves the daughter.'
b. Doč' ljubit mat'.
daughter-NOM/ACC loves mother-NOM/ACC
'The daughter loves the mother.'

Interesting speaker variation in argument interpretation has been reported regarding Russian sentences with syncretized arguments. For some speakers consulted, sentences like (104a,b) with two human arguments are not ambiguous; they can only be unambiguously interpreted with the first NP as the agent, not as the patient. But these speakers accepted sentences like (105) as ambiguous, where both arguments are inanimate, with both SVO and OVS readings.[39] Yet other speakers accepted all the sentences in (104) and (105) with both SVO and OVS readings.[40]

(105) Avtobus pereexal trolejbus
bus-NOM/ACC hit trolley-NOM/ACC
'The bus hit the trolley.' or 'The trolley hit the bus.'

Building on Asudeh's (2001) insight that ambiguity can be characterized as a situation in which more than one optimal output is selected in the com-

[39]Thanks to Roger Levy for providing the Russian examples and reporting some Russian speakers' judgements to me.

[40]I am grateful to the anonymous reviewer for making this observation and reporting some Russian speakers' judgements to me.

prehension grammar, Lee (2000a, 2001) has shown that bidirectional optimization, coupled with the notion of probabilistically ranked constraints (Boersma 1997, 1999; Boersma and Hayes 2001), provides a simple explanation for cross-linguistic variation in the resolution of ambiguity as to the identity of grammatical relations. Although at present numerous questions raised by the recoverability phenomenon cannot be fully answered, a careful investigation of this phenomenon will show, I believe, that various types of systematic crosslinguistic differences in the treatment of ambiguity and recoverability must lie within the domain of linguistic competence, and are not simply attributable to performance effects.

3.6 Conclusion

This paper has presented an OT-LFG account of word order freezing in Hindi and Korean. Marked associations of morphosyntactic hierarchies, which provide an important source of the 'worst of the worst' type of freezing, have been formally modeled as harmonic alignment: the most marked associations of grammatical function with other prominence hierarchies are expressed in the unmarked word order. I have also shown that word order freezing in sentences with ambiguous case marking can be explained when the constraint system allows an extension to bidirectional competition. Yet these results cannot be achieved in most current formal syntactic frameworks, because they give no theoretical role to markedness, as opposed to purely structural aspects of grammar (e.g., transformational derivations), and the production- or generation-oriented perspective alone is insufficient for explaining the phenomena of recoverability and ambiguity. Furthermore, word order freezing effects in Hindi and Korean show that concepts that have been successfully modelled in phonology—markedness hierarchies, harmonic alignment, contextual neutralization, etc.—also play a key role in the syntactic domain of constituent ordering. These preliminary results suggest that word order freezing phenomena can be subsumed under the universal theory of markedness, although further work is required.

References

Ahn, Hee-Don. 1991. *Light verbs, VP-movement, negation and clausal architecture in Korean and English*. Ph.D. thesis, University of Wisconsin-Madison.

Aissen, Judith. 1997. Person and Subject Choice in Optimality Theory. Paper presented at the Hopkins Optimality Workshop, Baltimore, May 9-12, 1997.

Aissen, Judith. 1999. Markedness and Subject Choice in Optimality Theory. *Natural Language and Linguistic Theory* 17:673–711.

Aissen, Judith. 2000. Differential Object Marking: Iconocity vs. Economy. MS., University of California, Santa Cruz.

Anttila, Arto and Vivienne Fong. 2000. The Partitive Constraint in Optimality Theory. MS., National University of Singapore, Singapore. Online, Rutgers Optimality Archive. ROA-410-0900. http://ruccs.rutgers.edu/roa.html.

Artstein, Ron. 1998. Hierarchies. Online, Rutgers Optimality Archive: `http://www.eden.rutgers.edu/~artstein`.

Asudeh, Ash. 2001. Linking, Optionality, and Ambiguity in Marathi. This volume.

Bahl, Kali Charan. 1967. *A Reference Grammar of Hindi (A Study of Some Selected Topics in Hindi Grammar)*. Chicago, Illinois: University of Chicago Press.

Beaver, David. 2000. Centering and the Optimization of Discourse. MS., Stanford University. Online: `http://www.stanford.edu/~dib`.

Bloom, Douglas. 1999. Case Syncretism and Word Order Freezing in the Russian Language. M.A. Thesis, Stanford University. Online: `http://www-lfg.stanford.edu/lfg/archive`.

Blutner, Reinhard. 1999. Some Aspects of Optimality in Natural Language Interpretation. In *Papers in Optimality Theoretic Semantics*, eds. Helen de Hoop and Henriëtte de Swart. Utrecht University, Utrecht Institute of Linguistics OTS, 1–21.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. In *IFA Proceedings 21*. University of Amsterdam: Institute of Phonetic Sciences, 43–58. ROA-221-109.

Boersma, Paul. 1999. Phonology-Semantics Interaction in OT and Its Acquisition. Online, Rutgers Optimality Archive: ROA-369-1299, `http://ruccs.rutgers.edu/roa.html`.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Bresnan, Joan. 1994. Linear Order vs. Syntactic Rank: Evidence from Weak Crossover. In *CLS 30*. Reprinted in *Formal Issues in Lexical-Functional Grammar*, ed. by Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III and Annie Zaenen, 241–274. Stanford, California: CSLI Publications.

Bresnan, Joan. 2000a. Explaining Morphosyntactic Competition. In *Handbook of Contemporary Syntactic Theory*, eds. Mark Baltin and Chris Collins. Oxford: Blackwell Publishers, 11–44.

Bresnan, Joan. 2000b. Optimal Syntax. In *Optimality Theory: Phonology, Syntax and Acquisition*, eds. Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer. Oxford: Oxford University Press, 334–385.

Bresnan, Joan. 2001a. The Emergence of the Unmarked Pronoun. In *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, Massachusetts: The MIT Press, 113–142.

Bresnan, Joan. 2001b. *Lexical Functional Syntax*. Oxford: Blackwell.

Bresnan, Joan and Jonni Kanerva. 1989. Locative Inversion in Chichewa: A Case Study of Factorization in Grammar. *Linguistic Inquiry* 20:1–50.

Bresnan, Joan and Ronald M. Kaplan. 1982. Introduction: Grammars as Mental Representations of Language. In *The Mental Representation of Grammatical Relations*, ed. Joan Bresnan. Cambridge, Massachusetts: The MIT Press, xvii–lii.

Bresnan, Joan and Annie Zaenen. 1990. Deep unaccusativity in LFG. In *Grammatical relations: A cross-theoretical perspective*, eds. Katarzyna Dziwirek,

Patrick Farrell, and Errapel Mejias-Bikandi. Stanford, California: CSLI Publications, 45–57.

Broadwell, George Aaron. 2000. Word Order and Markedness in Kaqchikel. In *Proceedings of the LFG00 Conference*, eds. Miriam Butt and Tracy H. King. Stanford, California: CSLI Publications Online: `http://csli-publications.stanford.edu/`.

Butt, Miriam and Tracy H. King. 1996. Structural Topic and Focus without Movement. In *Proceedings of the LFG96 Conference*, eds. Miriam Butt and Tracy H. King. Stanford, California: CSLI Publications Online: `http://csli-publications.stanford.edu/`.

Cho, Young-mee Yu and Peter Sells. 1995. A Lexical Account of Inflectional Suffixes in Korean. *Journal of East Asian Linguistics* 4:119–174.

Choi, Hye-Won. 1996. *Optimizing Structure in Context: Scrambling and Information Structure*. Ph.D. thesis, Stanford University.

Choi, Hye-Won. 1999. *Optimizing Structure in Context: Scrambling and Information Structure*. Stanford, California: CSLI Publications.

Choi, Hye-Won. 2001. Phrase Structure, Information Structure, and Resolution of Mismatch. This volume.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam. 1993. A Minimalist Program for linguistic theory. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, eds. Samule J. Keyser and Kenneth Hale. Cambridge, Massachusetts: The MIT Press, 1–52.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Massachusetts: The MIT Press.

Costa, João. 1998. *Word Order Variation: A Constraint-Based Approach*. Ph.D. thesis, Leiden University.

Culy, Christopher. 1995. Ambiguity and Case Marking in Donnɔ So (Dogon). In *Theoretical Approaches to African Languages*, ed. Akinbiyi Akinlabi. Trenton: Africa World Press, 47–58.

Davidson, Alice. 1969. Reflexivization and Movement Rules in Relation to a Class of Hindi Psychological Predicates. In *CLS 5*. Chicago, Illinois: Chicago Linguistic Society, 37–51.

de Hoop, Helen. 1996. *Case Configuration and Noun Phrase Interpretation*. New York, New York: Garland.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: The MIT Press.

Donohue, Cathryn. 1999. Optimizing Fore Case and Word Order. ms., Stanford University. Online: `http://www-csli.stanford.edu/~donohue`.

Dowty, David. 1991. Thematic Proto-roles and Argument Selection. *Language* 67:547–619.

Dviwedi, Veena Dhar. 1994. *Syntactic Dependencies and Relative Phrases in Hindi*. Ph.D. thesis, University of Massachusetts at Amherst.

England, Nora C. 1983. *A Grammar of Mam, a Mayan Language*. Austin: University of Texas Press.

Foley, William. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Gambhir, Vijay. 1981. *Syntactic Restrictions and Discourse Functions of Word Order in Standard Hindi*. Ph.D. thesis, University of Pennsylvania.

Gerdts, Donna and Cheong Youn. 1988. Korean Psych Constructions: Advancement or Retreat? In *CLS 25-I*. Chicago, Illinois: Chicago Linguistic Society, 155–175.

Grimshaw, Jane. 1991. Extended Projection. ms., Dept. of Linguistics and Center for Cognitive Science, Rutgers University.

Grimshaw, Jane. 1997. Projection, Heads, and Optimality. *Linguistic Inquiry* 28:73–422.

Gurtu, Madhu. 1985. *Anaphoric Relations in Hindi and English*. Ph.D. thesis, Central Institute of English and Foreign Languages, Hyderabad, India.

Hale, Mark and Charles Reiss. 1998. Formal and Empirical Arguments Concerning Phonological Acquisition. *Linguistic Inquiry* 29:656–683.

Halvorsen, Per-Kristian. 1983. Semantics for Lexical-Functional Grammar. *Linguistic Inquiry* 14:567–615.

Han, Eunjoo. 1991. Honorification in Korean. ms., Stanford University.

Hawkinson, Annie and Larry Hyman. 1974. Hierarchies of Natural Topic in Shona. *Studies in African Linguistics* 5:147–170.

Holmberg, Anders. 1998. Word Order Variation in Some European SVO Languages: A Parametric Approach. In Siewierska, Rijkhoff and Bakker (eds.), 553–598.

Hong, Ki-Sun. 1991. *Argument Selection and Case Marking in Korean*. Ph.D. thesis, Stanford University.

Hyman, Larry and Alessandro Duranti. 1982. On the Object Relation in Bantu. In *Studies in Transitivity (Syntax and Semantics 15)*, eds. P. Hopper and S. Thompson. New York, New York: Academic Press, 217–239.

Jäger, Gerhard. 2000. Some Notes on the Formal Properties of Bidirectional Optimality Theory. Online, Rutgers Optimality Archive: ROA-414-09100, `http://ruccs.rutgers.edu/roa.html`.

Jakobson, Roman. 1963. Implications of Language Universals for Linguistics. In *Universals of Language*, ed. Joseph H. Greenberg. Cambridge, Massachusetts: The MIT Press.

Joshi, Smita. 1993. *Selection of Grammatical and Logical Functions in Marathi*. Ph.D. thesis, Stanford University.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kim, Jong-Bok. 1999. *The Grammar of Negation: A Lexicalist, Constraint-Based Perspective*. Stanford, California: CSLI Publications.

Kim, Young-joo. 1990. *The Syntax and Semantics of Korean Case: The Interaction between Lexical and Syntactic Levels of Representation*. Ph.D. thesis, Harvard University.

King, Tracy Holloway. 1995. *Configuring Topic and Focus in Russian*. Stanford, California: CSLI Publications.

Kiparsky, Paul. 1998. Structural Case. ms., Stanford University.

Kuhn, Jonas. 1999. Two Ways of Formalizing OT Syntax in the LFG Framework. ms., Universität Stuttgart. Online: `http://www.ims.uni-stuttgart.de/~jonas`.

Kuhn, Jonas. 2000a. Faithfulness violations and Bidirectional Optimization. In *Proceedings of the LFG00 Conference*, eds. Miriam Butt and Tracy H. King. Stanford, California: CSLI Publications Online: `http://csli-publications.stanford.edu/`.

Kuhn, Jonas. 2000b. Issues in the Formalization of OT Syntax. Paper presented at the Syntax Workshop, Stanford University, March 14, 2000.

Kuhn, Jonas. 2001. Generation and Parsing in Optimality Theoretic Syntax – Issues in the Formalization in OT-LFG. This volume.

Kuno, Susumu. 1980. A Further Note on Tonoike's Intra-subjectivization Hypothesis. In *MIT Working Papers in Linguistics 2: Theoretical Issues in Japanese Linguistics*, eds. Yukio Otsu and Ann Farmer. Cambridge, Massachusetts: The MIT Press, 171–184.

Lasnik, Howard and Mamoru Saito. 1992. *Move α*. Cambridge: Cambridge University Press.

Lee, Hanjung. 1999. The Emergence of the Unmarked Order. Online, Rutgers Optimality Archive: ROA-323-0699, `http://ruccs.rutgers.edu/roa.html`.

Lee, Hanjung. 2000a. Bidirectional Optimality and Ambiguity in Argument Expression. Paper presented at the LFG00 Conference, University of California, Berkeley, July 19-20. ms., Stanford University. Online: `http://www.stanford.edu/~hanjung`.

Lee, Hanjung. 2000b. The Emergence of the Unmarked Order in Hindi. In *Proceedings of NELS 30*, eds. Masako Hirotani, Andries Coetzee, Nancy Hall, and Ji-Yung Kim. Amherst, Massachusetts: GLSA, 469–483.

Lee, Hanjung. 2000c. Markedness and Pronoun Incorporation. To appear in the *Proceedings of BLS* 26. Berkeley, California: Berkeley Linguistics Society.

Lee, Hanjung. 2001. *Optimization in Argument Expression and Interpretation: A Unified Approach*. Ph.D. thesis, Stanford University.

Lee, Young-Suk. 1993. *Scrambling as Case-Driven Obligatory Movement*. Ph.D. thesis, University of Pennsylvania.

Legendre, Géraldine, Colin Wilson, Paul Smolensky, Kristin Homer, and William Raymond. 1995. Optimality and *Wh*-Extraction. In *Papers in Optimality Theory*, eds. Jill N. Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk. Amherst, Massachusetts: GLSA, University of Massachusetts, 607–636.

Lenerz, Jürgen. 1977. *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen: Gunter Narr Verlag.

Mahajan, Anoop. 1990. *The A/A-bar Distinction and Movement Theory*. Ph.D. thesis, MIT.

McCarthy, John and Alan Prince. 1994. The Emergence of the Unmarked: Optimality in Prosodic Morphology. In *NELS 24*, ed. M. Gonzàlez. GLSA, University of Massachusetts at Amherst, 333–379.

Miyagawa, Shigeru. 1997. Against Optional Scrambling. *Linguistic Inquiry* 28:1–25.

Mohanan, K.P. and Tara Mohanan. 1994. Issues in Word Order. In *Theoretical Perspectives on Word Order in South Asian Languages*, eds. Miriam Butt, Tracy Holloway King, and Gillian Ramchand. Stanford, California: CSLI Publications, 153–184.

Mohanan, Tara. 1992. Word Order in Hindi. Paper presented at the Syntax Workshop. Stanford University.

Mohanan, Tara. 1994a. *Argument Structure in Hindi*. Stanford, California: CSLI Publications.

Mohanan, Tara. 1994b. Case OCP: A Constraint on Word Order in Hindi. In *Theoretical Perspectives on Word Order in South Asian Languages*, eds. Miriam Butt, Tracy Holloway King, and Gillian Ramchand. Stanford, California: CSLI Publications, 185–216.

Mondloch, James. 1981. *Voice in Quiche-Maya*. Ph.D. thesis, State University of New York at Albany.

Morolong, 'Malillo and Larry Hyman. 1977. Animacy, Objects and Clitics in Sesotho. *Studies in African Linguistics* 8:199–218.

Müller, Gereon. 1998. German Word Order and Optimality Theory. In *Arbeitspapiere des Sonderforschungsbereichs 340, #126*. University of Tübingen.

Neeleman, Ad and Tanya Reinhart. 1997. The Syntax and Interpretation of Scrambling in Dutch. MS., University of Utrecht.

Pinkerton, Sandra. 1976. *Studies in Kekchi. Texas Linguistics Forum 3*. Austin: University of Texas.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago, Illinois and Stanford, California: The University of Chicago Press and CSLI Publications.

Primus, Beatrice. 1998. The Relative Order of Recipient and Patient in the Languages of Europe. In Siewierska, Rijkhoff and Bakker (eds.), 421–473.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report RuCCS Technical Report #2, Center for Cognitive Science, Rutgers University, Piscataway, New Jersey. To be published by the MIT Press.

Prince, Ellen. 2000. Linguistic and Meta-linguistic Intuition. To appear in the *Proceedings of BLS* 26. Berkeley, California: Berkeley Linguistics Society.

Rudin, Catherine. 1985. *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*. Columbus, Ohio: Slavica.

Sag, Ivan and Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. Stanford, California: CSLI Publications.

Saito, Mamoru. 1992. Long Distance Scrambling in Japanese. *Journal of East Asian Linguistics* 1:69–118.

Samek-Lodovici, Vieri. 1998. OT-interactions between Focus and Canonical Word Order: Deriving the crosslinguistic Typology of Structural Contrastive Focus. Online, Rutgers Optimality Archive: ROA-257-0498, `http://ruccs.rutgers.edu/roa.html`.

Sells, Peter. 1995. Korean and Japanese Morphology from a Lexical Perspective. *Linguistic Inquiry* 26:277–325.

Sells, Peter. 1999. Constituent Ordering as Alignment. In *Harvard Studies in Korean Linguistics VIII*, ed. Susumu Kuno et al. Seoul: Hanshin, 546–560.

Sells, Peter. 2000. Alignment Constraints in Swedish Clausal Syntax. MS., Stanford University. Online: `http://www-csli.stanford.edu/~sells`.

Sells, Peter. 2001. *Structure, Alignment and Optimality in Swedish*. To appear, Stanford, California: CSLI Publications.

Sharma, Devyani. 1999. Sentential Negation and Focus in Hindi. ms., Stanford University.

Siewierska, Anna. 1988. *Word Order Rules*. London: Croom Helm.

Siewierska, Anna, Jan Rijkhoff, and Dik Bakker (eds.). 1998. *Constituent Order in the Languages of Europe*. Berlin: Mouton de Gruyter.

Siewierska, Anna and Ludmila Uhlirova. 1998. An Overview of Word Order in Slavic Languages. In Siewierska, Rijkhoff and Bakker (eds.), 105–149.

Smolensky, Paul. 1995. On the Internal Structure of the Constraint Component Con of UG. ROA-86-0000. Rutgers Optimality Archive. `http://ruccs.rutgers.edu/roa.html`.

Smolensky, Paul. 1996a. The Initial State and "Richness of the Base" in Optimality Theory. Technical Report JHU-CogSci-96-4, Department of Cognitive Science, Johns Hopkins University.

Smolensky, Paul. 1996b. On the Comprehension/Production Dilemma in Child Language. *Linguistic Inquiry* 27:720–731.

Smolensky, Paul. 1998. Why Syntax is Different (but not Really): Ineffability, Violability and Recoverability in Syntax and Phonology. Stanford University Workshop: Is Syntax Different? (December 12–13, 1998).

Speas, Margaret J. 1990. *Phrase Structure in Natural Language*. Dordrecht: Kluwer Academic.

Srivastav, Veneeta. 1991. *WH Dependencies in Hindi and the Theory of Grammar*. Ph.D. thesis, Cornell University.

Tomlin, Russell. 1986. *Basic Word Order. Functional Principles*. Dordrecht: Foris.

Vallduví, Enric. 1992. *The Informational Component*. New York, New York: Garland.

Verma, Manindra K. and K. P. Mohanan (eds.). 1990. *Experiencer Subjects in South Asian Languages*. Stanford, California: CSLI Publications.

Webelhuth, Gert. 1989. *Syntactic Saturation Phenomena and the Modern Germanic Languages*. Ph.D. thesis, University of Massachusetts at Amherst.

Whitman, John. 1989. Topic, Modality, and IP Structure. In *Harvard Studies in Korean Linguistics III*, ed. Susumu Kuno et al. Seoul: Hanshin, 341–356.

Wilson, Colin. 1996. Alignment and Anaphora. Paper presented at the Stanford/CSLI Workshop on Optimality Theory and Its Implications for Cognitive Theory. Stanford University, December 6-8, 1996.

Wilson, Colin. 1997. Absolute Harmony, Relativized Minimality and Blocking. Paper presented at the Hopkins Optimality Workshop, Baltimore, May 9-12, 1997.

Wilson, Colin. 2001. Bidirectional Optimization and the Theory of Anaphora. In *Optimality-Theoretic Syntax*, eds. Géraldine Legendre, Jane Grimshaw, and Sten Vikner. Cambridge, Massachusetts: The MIT Press, 465–507.

Yoon, James Hye-Suk. 1996. Ambiguity of Government and the Chain Condition. *Natural Language and Linguistic Theory* 14:105–162.