

## Branching Consistency as a Syntactic OCP Constraint to Hakka Relative Construction

Yu-Ching Tseng

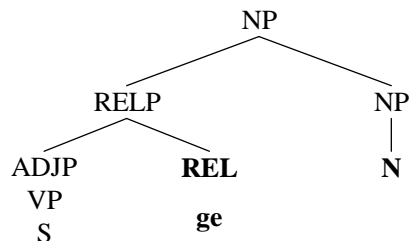
*State University of New York, University at Albany*

The OCP is traditionally recognized as a phonological constraint which bans the adjacency of two linguistic elements that share some identical phonological properties. This paper adopts an OT approach and shows a sub-type of OCP that is triggered by purely syntactic configurations. Knowing that the OCP can be used as a cooccurrence restriction prohibiting the multiple occurrences of some marked construction, this paper adopts the constraint to prevent the marked branching direction from occurring more than once if the basic word order shows the reversed direction. In Hakka, the basic word order of NP is head-final and left-branching, which gives way to right-branching as the marked configuration to indicate topicalization. It is argued that the syntactic OCP effect is observable when more than one nominal constituent within the so-called multiple embedded relative constructions receive a topicalized interpretation, and thus yielding an ungrammatical pattern where the marked right-branching order occurs at two levels of branching within the tree of NP, which is strictly disallowed by the OCP.

### 1. Hakka Relative Construction

According to Li & Thompson (1981), relative construction in Mandarin Chinese is constructed simply by placing a nominalized clause in front of a noun to modify it. Similarly in Hakka, a noun may be preceded by a relative modifier (REL<sub>P</sub>), which consists of a phrasal projection, could it be an ADJP, S, or a VP, in front of the modificational head morpheme *ge*. The structure is represented by the form (1):

(1)



According to (1), both NP and RELP are head-final. As a nominal modifier, the RELP functions as an adjunct adjoining to the left of NP, and within the RELP, the head relativizer selects either a phrasal (2a) or a clausal (2b, c) complement, which also precedes the head. A few examples are given in the following:

- (2) a. Ngai tok bun gi [NP[RELP[AP ka ho] ge] ng-e]  
 I select to him more good REL fish  
 ‘I picked the fish that is in better quality for him.’
- b. [NP[RELP[s Gi cong go-e] ge] sang] dong ho-tang  
 He sing song REL voice really harmonious  
 ‘The voice with which he sings is very good to listen to.’
- c. [NP[RELP[VP siit fan] ge] cien] ma-ngin oi cut  
 eat meal REL money who will pay  
 ‘Who will pay the money which is for the meal?’

By adopting an Optimality Theoretic (abbreviated as OT) (Pince and Smolensky 1993) approach, the predominately right-headed order can be generated by proposing a directional Generalized Alignment Constraint (McCarthy and Prince 1993), formally expressed by **ALIGN-R (Head, NP/RELP)**. This constraint aligns the head of NP or RELP to the right edge, ensuring that the structure of NP must be head-final and left-branching.

Tableau 1

| RC: [XP REL] N | <b>ALIGN-R</b> |
|----------------|----------------|
| ☞ [XP REL] N   |                |
| [REL XP] N     | *!             |
| N [XP REL]     | *!             |

As illustrated in tableau 1, in order to avoid violating ALIGN-R, the word order of NP must be as follows: Modifying Phrase-Relativizer-Head Noun, with this order the head of RELP and NP both occurs at the right edge.

## 2. Restrictive v.s. Non-Restrictive Relative Construction

Most languages divide relative clauses into two types, restrictive and non-restrictive. The restrictive RELP restricts the referent of the head noun it modifies to a subset of a larger domain, while the non-restrictive RELP simply add parenthetical information to the head noun.

As suggested by Tiee (1986), in Chinese languages the restrictive and non-restrictive distinction is made by the placement of relative clause with respect to the classifier

phrase (CLP) when they modify the same noun. Compare the two Hakka examples in (3) below:

(3) a. gi mai-tet [dong tai ge] [ge gien] vuk  
 He sell-off really big REL that CL house

b. gi mai-tet [ge gien] [dong tai ge] vuk  
 He sell-off that CL really big REL house

‘He sold that very big house.’

The two sentences in (3) present the two different orders between CLP and RELP. In (3a) the RELP precedes the CLP, while in (3b) the word order is reverse. As I will argue, some slight semantic difference actually exists to distinguish between their meanings. The reading of (3a) implies that “he owns more than one house, and among them the one he sold is the specific one that has a unique property as a very big one, which distinguishes this house from all the others.” But such implication is much weaker in (3b). This contrast of semantic meaning can be obtained by providing the following question-answer test (4):

(4) Q: Ge sam gien vuk gi mai-tet nai gien?  
 That three CL house he sell-off which CL  
 ‘Which of those three houses is the one he sold?’

Ans 1: (✓ better) Gi mai tet [dong tai ge] [ge gien] vuk.....(3a)

Ans 2: (# worse) Gi mai tet [ge gien] [dong tai ge] vuk.....(3b)

An appropriate answer to (4) ought to precisely single out one from the three houses that both speakers have common knowledge about. The first answer successfully attains this goal by emphasizing the size of a specific house, but the reading of the second answer simply describes his selling that big house as an event. Therefore, the first answer is considered more adequate as an appropriate response to the proposed question. As shown in (5), if a given question requires some general description about what he has done, in this situation the question can be answered by describing an event as explanation, the second answer in (4) that corresponds to (3b) then becomes a good answer.

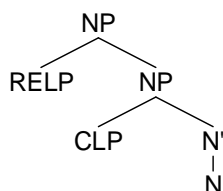
(5) Q: Gi ngiong-voi con an-do cien no  
 He how come make so much money PART  
 ‘How could he make so much money?’

**Ans:** (✓ good) yin-vi gi mai-tet [**ge gien**] [**dong tai ge**] vuk  
 Because he sell-off that CL really big REL house  
 ‘It is because he sold that very big house.’

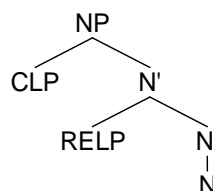
The semantic contrast described above can be observed in NP when the order between the modifier RELP and CLP varies. As maintained by Tsee (1986), the different order decides whether a relative clause is restrictive or non-restrictive. Generally speaking, a relative clause is restrictive when it precedes the CLP, and which is non-restrictive when it follows the CLP. As I will argue in this paper, the phrase structure for the two types of relative clause is fundamentally different, reflected in the following (6):

(6)

Restrictive Relative Clause



Non-Restrictive Relative Clause



The RELP adjoins to the higher NP in restrictive relative construction; while in non-restrictive construction, the RELP adjoins to the lower N'. The derived word order between RELP and CLP is thus opposite; that is, for restrictive clauses, the RELP precedes the CLP; while in non-restrictive clauses, the RELP follows the CLP.

### 3. Topicalization and Non-restrictive Interpretation

In Hakka another alternative to mark the non-restrictive reading for a restrictive relative construction is through topicalization. See the following examples (7). Given a restrictive relative construction (7a) as described by the left form in (6), the process takes place by fronting the topicalized N (7c, d) or the lower NP (7b) to the leftmost position. When the fronted element is recognized as the center of discussion, it turns out to be the most salient part in the entire NP, and the following modifiers merely offer additional information about the referring topic.

(7) a. ngai gau gi gong [**mi-guet ngin gong ge**] [**ge zung**] yin-vun  
 I teach he speak America people speak REL that CL English  
 ‘I taught him to speak that kind of English which the American uses.’

- b. *ngai gau gi gong ge zung yin-vun*, [mi-guet ngin gong ge]  
 I teach him speak that CL English America people speak REL  
 ‘I taught him that kind of English, which the American uses.’
- c. *ngai gau gi gong yin-vun*, [mi-guet ngin gong ge] [ge zung]  
 I teach him speak English America people speak REL that CL  
 ‘I taught him English, that kind which the American uses.’
- d. *ngai gau gi gong yin-vun*, [ge zung] [mi-guet ngin gong ge]  
 I teach him speak English that CL America people speak REL  
 ‘I taught him English, that kind which the American uses.’

The Above three sentences (7b-d) correspond to (7a). They contain a topicalized information expressed by a fronted nominal, which converts the following modifier into non-restrictive. Note that when the topic is presented simply by a head noun, the order between its modifier RELP and CLP is syntactically free, as contrasted by (7c) and (7d). To make the process of topicalization a possible solution for restrictive relative clauses to acquire the non-restrictive reading, another generalized alignment constraint **ALIGN-L (TOP, NP)** must be proposed, which functions to locate the topicalized element to the leftmost position of NP. As shown by the following tableaux 2 and 3, this constraint must outrank the **ALIGN-R (Head, NP/RELP)** proposed in the previous section.

Tableau 2

| R.RC: RELP [CLP N]         | ALIGN-L (TOP, NP) | ALIGN-R (Head, NP/RELP) |
|----------------------------|-------------------|-------------------------|
| (a) RELP <b>CLP</b> N      | *!                |                         |
| (b) <b>CLP</b> RELP N      | *!                |                         |
| (c) [N <b>CLP</b> ] RELP   |                   | **!                     |
| (d) N RELP <b>CLP</b>      | *!                | **                      |
| (e) RELP N <b>CLP</b>      | *!                | *                       |
| ☞ (f) [ <b>CLP</b> N] RELP |                   | *                       |

Provided with an input of restrictive construction, in tableau 2 the lower NP is presented as the topic (**marked by bold text**) that carries some kind of pragmatic prominence. Since the constraint ALIGN-L outranks another constraint ALIGN-R, knowing that the topicalized constituent in candidate (c) and (f) has been fronted to the leftmost position of NP, without violating ALIGN-L, the two candidates win over all the other candidates. Candidate (f) beats candidate (c) in candidate competition because by preserving the head-final structure for the topicalized NP, candidate (f) incurs only one violation on ALIGN-R, while in (c) the head noun stands at the left of both CLP and

RELP, thus, it incurs two violations on ALIGN-R. Since Alignment constraints are accumulative, (c) triggers more serious violation on ALIGN-R than (f) does.

Tableau 3

| R.RC: RELP CLP N | ALIGN-L (TOP, NP) | ALIGN-R (Head, NP/RELP) |
|------------------|-------------------|-------------------------|
| (a) RELP CLP N   | *!                |                         |
| (b) CLP RELP N   | *!                |                         |
| ☞ (c) N CLP RELP |                   | **                      |
| ☞ (d) N RELP CLP |                   | **                      |
| (f) RELP N CLP   | *!                | *                       |
| (g) CLP N RELP   | *!                | *                       |

In tableau 3 only the head noun receives the topic reading, by assigning high ranking to ALING-L, the head noun may not stay at phrase final when it is highlighted as the center of interest in the phrase. In this tableau, candidate (c) and (d) are the winning candidates even though each of them gets double violations on ALIGN-R, a constraint that requires the head noun to appear at the far right of NP. The violations are tolerated since all the other candidates place the topicalized noun in position other than the left edge of NP, which leads to a worse violation on the higher ranked ALIGN-L.

#### 4. Multiple Embedded Relative Construction

A multiple embedded relative clause refers to a complicated construction where a noun within a relative clause is further modified by another relative clause. An example is shown in the following (8):

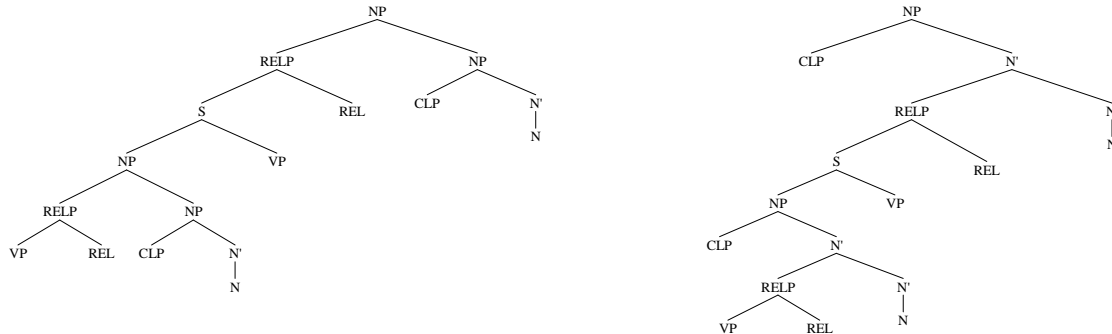
(8) a. gi bun ngai [[hi mi-guet ge] [ge zhak] sei-lai kon go ge] [ge bun] su  
 He give I go America REL that CL boy read ASP REL that CL book  
 ‘He gave me the book which had been read by the boy who went to the U.S..’

b. gi bun ngai [ge bun] [[ge zhak] [hi mi-guet ge] sei-lai kon go ge] su  
 He give I that CL that CL go America REL boy read ASP REL book  
 ‘He gave me that book, which that boy, who went to the U.S., had read before.’

In (8a), a restrictive RELP is embedded into another restrictive RELP, while in (8b) the two RELPs are both non-restrictive. The distinction is made by the different order between CLP and RELP. (8a) exhibits the order in which the RELP precedes the CLP, while in (8b) the order is opposite.

The following diagrams (9) illustrate the structure for the two object NPs in (8). The left diagram corresponds to the NP in (8a), and the right diagram corresponds to the NP in (8b).

(9)



As indicated in the preceding section, topicalization provides an alternative to express the non-restrictive interpretation. All sentences presented in (10) below are examples of multiple-embedded construction. Each of them contains a non-restrictive modifier that follows the topicalized N/NP. Notice that (10a, a') actually sound more natural than (10b, b') even though both of them are accepted as grammatical. The reason is that for an NP that contains such heavy-loaded information, it takes more time for a Hakka speaker to understand the meaning conveyed by sentences (10b, b'), since in such cases the topic and the head of the object NP are not coherent. As to the first pair of sentences (10a, a'), the topicalized information that states the key point of the sentence is in coherence with the head of NP, which makes it easier to understand what the complex NP is about.

- (10) a. gi bun ngai [**ge bun su**, [hi mi-guet ge ge zhak sei-lai] kon go ge]  
 He give I that CL book go America REL that CL boy read ASP REL  
 'He gave me that book, which had been read by the boy who went to the US.'
- a'. gi bun ngai [**su**, [hi mi-guet ge ge zhak sei-lai] kon go ge ge bun]  
 He give I book go America REL that CL boy read ASP REL that CL  
 'He gave me the book, that one which had been read by the boy who went to the US.'
- b. gi bun ngai [[**ge zhak sei-lai**, hi mi-guet ge] kon go ge ge bun su]  
 He give I that CL boy go America REL read ASP REL that CL book  
 'He gave me the book which that boy, who went to the US, had read before.'
- b'. gi bun ngai [[**sei-lai**, hi mi-guet ge ge zhak] kon go ge ge bun su]  
 He give I boy go America REL that CL read ASP REL that CL book  
 'He gave me the book which the boy, that one who went to the US, had read before.'

However, if we need both embedded RELPs to be non-restrictive, the process of topicalization may not occur to both RELPs, this can be shown by the ungrammatical sentence in (11):

- (11) \*gi bun ngai [ge bun su, [ge zhak sei-lai, hi mi-guet ge] kon go ge]  
 He give I that CL book that CL boy go America REL read ASP REL  
 ‘He gave me that book, which that boy, who went to the U.S., had read before.’

To account for why is the ungrammaticality of (11), an OT-based proposal will be developed in the next section.

## 5. Toward a New Type of Syntactic OCP

The Obligatory Contour Principle (OCP) (Leben 1973, Goldsmith 1976, McCarthy 1981, 1986), a well-known linguistic constraint which prohibits the juxtaposition of two identical elements, was originated around 1970s and used predominately in the field of phonology. Even though linguists including Mohanan (1994), Golston (1995), Yip (1995, 1998), Anttila and Fong (2000) and others have implemented the principle in their research to deal with morph-syntactic phenomena, the concept of OCP is still phonological in nature, which started from a sense of disfavor of phonological identity of some kind, and the trigger of effects are elements that share identical phonological property. In the following I will introduce a new type of syntactic OCP which differs from all the previous approaches in that it places identity restriction on purely syntactic configurations.

### 5.1. Branching Consistency and the Markedness OCP

Linguists generally agree on the hypothesis which argues for a systematic correlation between the basic word order and the ordering of other phrasal categories. The Greenbergian word order typology is termed the “Head-Dependent Theory (HDT)” by Dryer (1992), who proposed an alternative account to the HDT, which is termed the “Branching-Direction Theory (BDT)”. This paper will not go over the detail of these two theories, all relevant discussions can be found in their original research.

Based on these typological theories, it is generally accepted that there exists a tendency for right-branching languages to have recursive branching on its right; while in left-branching languages, the recursive side normally occurs on the left. This is also suggested by Broadwell (2002), according to him, syntactic structures tend to have consistent branching direction. Therefore, relating the idea of branching consistency with the markedness theory, it is considered marked for a right-branching language to have left-branching constructions in the scope; in contrast, for a left-branching language, right-branching constructions are considered the marked configuration. The typology of branching direction is recapitulated in the following (12):



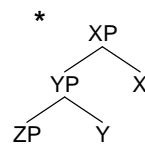
(12) Correlations of Branching Direction and Markedness Configuration

|   | <b>Marked</b>                                 | <b>Unmarked</b>                               |
|---|---|---|
| <b>Right-branching<br/>(head-initial)</b> | Phrasal elements precede non-phrasal elements | Phrasal elements follow non-phrasal elements  |
| <b>Left-branching<br/>(head-final)</b>    | Phrasal elements follow non-phrasal elements  | Phrasal elements precede non-phrasal elements |

This paper adopts the concept of OCP developed by Alderete (1996, 1997) as well as Itô and Mester (1996, 1998). Following their suggestion, the OCP can be used as a cooccurrence restriction which bans the multiple occurrences of some marked constructions, enhanced by the idea of self-conjunction of markedness constraints. The issue of structural consistency therefore has potential to be identified as an OCP-related phenomenon. The basic idea is that the multiple occurrences of marked configuration should be recognized as ill-formed in the language where the basic word order shows opposite branching direction. That is, the issue related to structural consistency may be identified as an OCP-triggered effect, which prevents the branching direction that is considered marked in the target language from occurring multiple times if its basic word order shows the reversed direction, an illustration of this idea is provided in (13):

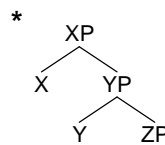
(13) Branching Consistency and OCP Violation

Right-branching languages



OCP violation

Left-branching languages

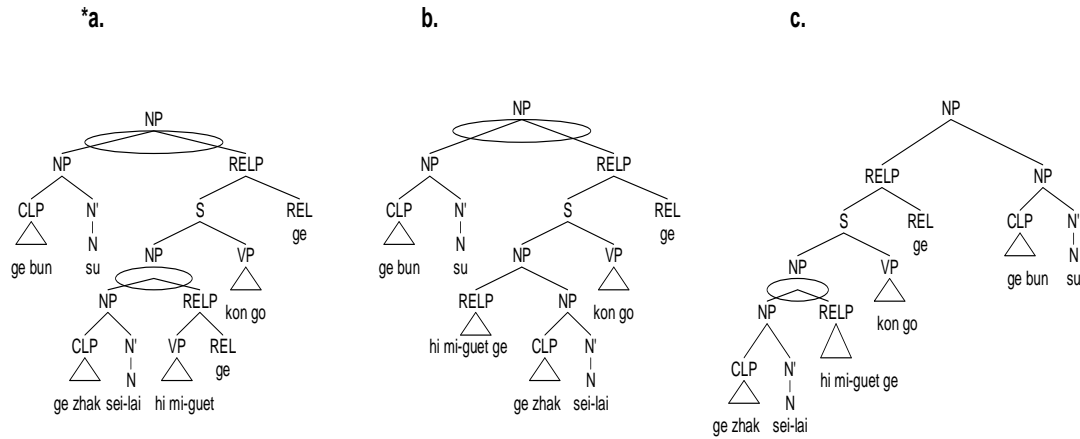


OCP violation

**5.2. The Case of Hakka**

We have described in earlier sections that the Hakka noun phrases are characterized by the “head-final” and “left-branching” rule. Following the inferences drawn from the prior paragraph, inside an NP, a branching modifier is supposed to precede the head noun to conform to the unmarked left-branching rule. If the modifier phrase follows the head noun, the construction is regarded as marked by having the recursive side on the right of the head. Now compare the following three diagrams:

(14)



The forms in (14) display the structure for the three cases of multiple-embedded clauses. (14a) corresponds to the previous example (11); (14b) and (14c) correspond respectively to (10a) and (10b). In (14b) and (14c), only one of the NPs in the tree, the upper one in tree (b) and the lower one in tree (c), undergoes topicalization, and they are both well-formed. In contrast, (14a) is ungrammatical because this diagram contains two topicalized NPs at different layers in the tree. That is to say, since “left-branching” is considered the unmarked rule for the Hakka NP, when topicalization applies to an NP, the syntactic process triggers right-branching configuration by fronting the head noun to the left of its modifier RELP, the resulting marked construction is allowed if it only appears once within an NP, but the construction becomes ill-formed if two of them cooccur in the same NP.

Going into the analysis, a sub-constraint of the OCP must be proposed, the constraint is written as **OCP-RightBranching (NP)**, which penalizes the repeated occurrence of the marked right-branching configuration in two depths of branch within the same NP. This OCP manifestation must rank above the previously proposed **ALIGN-L (TOP, NP)**, which demands topicalization to take place at the left edge of NP. The constraint interaction is illustrated in the following tableau 4, which show the established constraint ranking, **OCP-RB (NP) >> ALIGN-L (TOP, NP) >> ALIGN-R (Head, NP/RELP)**:

Tableau 4

| [[hi mi-guet ge] [ge zhak sei-lai] kon go ge] [ge bun] su<br>Go America REL that CL boy read ASP REL that CL book<br>'that book, which had been read by that boy, who went to the U.S.' | OCP-RB<br>(NP) | ALIGN-L<br>(TOP) | ALIGN-R<br>(Head) |
|---|----------------|------------------|-------------------|
| ge bun su, ge zhak sei-lai, hi mi-guet ge kon go ge   | *!             |                  | **                |
| ☞ ge bun su, hi mi-guet ge ge zhak sei-lai kon go ge  |                | *                | *                 |
| ☞ ge zhak sei-lai, hi mi-guet ge, kon go ge ge bun su   |                | *                | *                 |
| hi mi-guet ge ge zhak sei-lai kon go ge ge bun su   |                | **!              |                   |

When (14a, b, c) are candidates for the examination of multiple-embedded relative construction, with the two nominal constituents in both NPs marked by some topicalized salience, the evaluation is presented in tableau 4. As it shows, the first candidate, corresponding to (14a), is constantly ruled out as ungrammatical due to its violation of the high-ranking OCP; and the last candidate, in which both topicalized nominals stay in situ, is also ruled out since it collects double violation on ALIGN-L.

From the analysis we see that syntactically we should be able to topicalize the most salient element in both NPs when one of them is embedded into another; however, as demonstrated in this tableau, the OCP constraint proposed in this section against inconsistent branching direction crucially disallows multiple nominal constituents being topicalized within an NP domain. We can choose only one element that is emphatically stronger in discourse prominence to be aligned leftward preceding all the other constituents in the NP, at the risk of ALIGN-R, a constraint that assures the head-final configuration for NP and RELP; as well as ALIGN-L, which requires all the topicalized elements to be preposed in the leftmost position.

## 6. Conclusion

While the OCP is traditionally recognized as a universal phonological constraint, in which the trigger of violation are elements containing some identical phonological properties. This paper shows another type of OCP in which the triggers are purely syntactic configurations. The OCP is argued to be bound with the markedness theory, given the fact that marked features usually incur more serious OCP violation than unmarked features; if it is true that syntactic structures tend to be consistent in branching direction, phrases that disobey the tendency would be identified as marked configuration which, when applies to the OCP theory, may cause ungrammaticality if it occurs repeatedly in some defined domain.

In Hakka, since right-headed and left-branching is considered the unmarked rule to configure the structure of NP. The marked right-branching configuration triggered by the process of topicalization is prohibited to occur more than once within the domain that is defined to be the top NP in a multiple embedded relative construction. This Hakka case study is presented in this paper as a demonstration for the new type of syntactic OCP effect based on the idea of structural consistency and markedness theory.

REFERENCES

- ALDERETE, JOHN. 1996. A Formal Theory of Dissimilation. Handout presented at Rum J. Clam II, University of Massachusetts, Amhurst, MA. ROA-175.
- ALDERETE, JOHN. 1997. Dissimilation as Local Conjunction. *Proceedings of the NorthEast Linguistics Society* 27, 17-32. GLSA, Amhurst, MA.
- ANTTILA, ARTO, and VIVIENNE FONG. 2000. The Partitive Constraint in Optimality Theory, (ms.). Boston University, MA; National University of Singapore, Singapore. ROA-416.
- BROADWELL, GEORGE AARON. 2002. Branching Consistency as a Constraint on Zapotec Syntax. Paper presented at Chicago Linguistic Society (CIL) 2002, Chicago, IL.
- DRYER, MATTHEW S. 1992. The Greenbergian Word Order Correlations. *Language* 68-1, 81-138.
- GOLDSMITH, A. JOHN. 1976. *Autosegmental Phonology*. Doctoral dissertation, MIT, Cambridge, MA.
- GOLSTON, CHRIS. 1995. Syntax Outranks Phonology: Evidence from Ancient Greek. *Phonology* 12-3, 343-368.
- ITÔ, JUNKO, and R. ARMIN MESTER. 1996. Rendaku I: Constraint Conjunction and the OCP. Paper presented at Kobe Phonology Forum, Kobe, Japan. ROA-144.
- ITÔ, JUNKO, and R. ARMIN MESTER. 1998. Markedness and Word Structure: OCP Effects in Japanese, (ms.). UC Santa Cruz, CA. ROA-255.
- LEBEN, WILLIAM. 1973. *Suprasegmental Phonology*. Doctoral dissertation, MIT, Cambridge, MA.
- LI, CHARLES N., and SANDRA A. THOMPSON. 1981. *Mandarin Chinese: A Functional Reference Grammar*. U. of California Press, Berkeley, CA.
- MCCARTHY, JOHN. 1981. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry* 12, 373-418.
- MCCARTHY, JOHN. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17.2, 207-263.
- MCCARTHY, JOHN, and ALAN PRINCE. 1993. Generalized Alignment. In G. E. BOOIJ and J. VAN MARLE, (eds.). *Yearbook of Morphology*, 79-153. Dordrecht, Kluwer.
- MOHANAN, TARA. 1994. Case OCP: A Constraint on Word Order in Hindi. In MIRIAM BUTT, TRACY HOLLOWAY KING and GILLIAN RAMCHAND, (eds.). *Theoretical Perspectives on Word Order in South Asian Languages*, 185-216. CSLI, Stanford, CA.
- PRINCE, ALAN, and PAUL SMOLENSKY. 1993. Optimality Theory: Constraint Interaction in Generative Grammar, (ms.). Rutgers University, New Brunswick, NJ; University of Colorado, Boulder.
- TIEE, HENRY HUNG-YEH. 1986. *A Reference Grammar of Chinese Sentences with Exercises*. The University of Arizona Press.

- YIP, MORIA. 1995. Repetition and Its avoidance: the case of Javanese, (ms.). University of California, Irvine, CA. ROA-83.
- YIP, MORIA. 1998. Identity Avoidance in Phonology and Morphology. In STEPHEN G. LAPOINTE, DIANE K. BRENTARI and PATRICK M. FARRELL, (eds.). *Morphology and its relation to phonology and syntax*, 216-246. CSLI, Stanford, CA.

