

Grammar and the Shortest Path

ANNA MARIA DI SCIULLO
Department of Linguistics
Université du Québec à Montréal
P.O. Box 8888, Station Centre-Ville
Montreal, Qc H3C 3P8
CANADA

di_sciullo.anne-marie@uqam.ca <http://www.interfaceasymmetry.uqam.ca>

Abstract – The application of the operations of the grammar in the computational space yields the geometrical basis (phrase marker) upon which paths in natural language configurations can be identified. A path between two nodes in a phrase marker provides a measure of the trajectory of a moved constituent and the distance between bound elements. I propose that the properties of paths in phrase markers are determined by interface conditions (Full Interpretation) and Economy. The Minimal Path reduces derivational complexity and contributes to make complex systems such as natural languages usable by humans.

Key-Words: - *Asymmetry Theory, asymmetric agree, proper subset relation, computational space, movement, binding, shortest path, minimal path*

1 Natural Languages

The identification of the shortest path between two points in a graph is a central issue in mathematics [1], [2]. It is also a central topic in linguistic theory, since the nodes of a linguistic graph (phrase marker) are related to one another, and the properties of the paths (chains) relating them are crucial for the legibility of linguistic expressions [3], [4].

I would like to suggest that interface legibility conditions and principles of Economy minimize paths in order to reduce the derivational complexity that would be brought about by the identification of the shortest amongst all the possible paths for a given relation in a phrase marker. Derivations are compared only if they must be, as a last resort strategy.

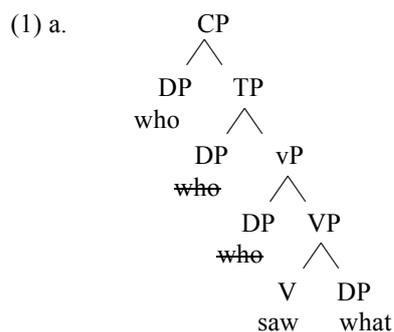
The organization of this paper is the following. First, I illustrate that paths are part of phrase markers, derived by the operations of the grammar of natural languages. Second, I consider the locality conditions that have been proposed on movement and binding. Finally, I propose the notion of Minimal Path, which length follows from Full Interpretation and Economy. Grammar determines paths to be minimal but never computes them.

2 Computational Space, Operations, and Phrase Markers

The architecture of the grammar of natural languages includes a computational space, where the operations of the grammar derive the structure of linguistic expressions. This structure cannot be reduced to mere

precedence relations between words, because words form constituents, and constituents are related to one another, displaced and bound.

The linguistic expressions are represented in terms of oriented graphs, where each node immediately dominates two children nodes [5]. Phrase markers, such as (1a), and their notational equivalents, such as (1b), allow for the identification of unambiguous paths between constituents. Paths that are not short enough bring about a decrease in the acceptability of linguistic expressions, as it is the case for (1c) compared to (1b), and in some cases lead to linguistic gibberish, as discussed in section 2.



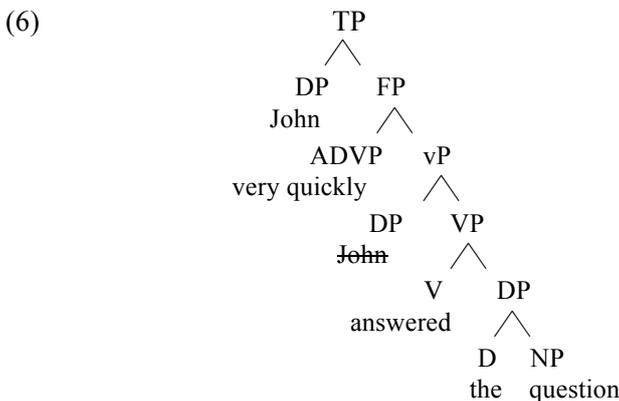
b. [CP who [TP who [vP who [VP saw what]]]]

c. ?[CP what [C did [TP who [vP who [VP saw what]]]]

I will start by assuming that a path is identified on the oriented graphs (phrase markers) derived by the operations of the grammar, such as the ones defined in Asymmetry Theory [6], (2), (3), which apply under asymmetric Agree, (4).

- (2) *Shift* (α, β)
Given two objects α and β , *Link* (α, β) creates a new object projected from α .
- (3) *Link* (α, β)
Given two objects α and β , *Link* (α, β) creates a new object where α and β are featurally related.
- (4) *Agree* (φ_1, φ_2)
Given two sets of features φ_1 and φ_2 , *Agree* holds between φ_1 and φ_2 iff φ_1 properly includes φ_2 , and the node dominating φ_1 asymmetrically c-commands the node dominating φ_2 .
- (5) a. C-command (constituent command)
 α c-commands β , iff each γ that dominates α dominates β , and α does not dominate β .
b. *Asymmetric c-command*
 α asymmetrically c-commands β , if α c-commands β , and β does not c-command α .

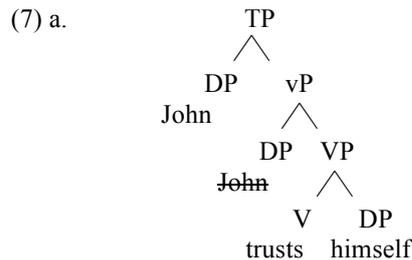
The head-complement and adjunct-adjoined relations could be defined in term of paths. The relation of a head to its complement would be the path relating for example the V *answered* in (6) to its DP complement *the question*, and the relation of an modifier to its modified constituent would be in the structure in (6) the path between the adverbial phrase (ADVP) *very quickly* to verbal constituent VP *answered the question*.



However, the head-complement and the modifier-modified relations relate different constituents, a V and a DP in the first case, and an ADVP and a VP in the second case. Moreover, no displacement or anaphoric relation holds between a head and its complement, or between an adverbial and the verbal projection it modifies at the points of the derivation where these constituents merge together in the construction of the phrase marker. I focus here on the properties of paths of singular constituents in Movement and Binding.

It is generally assumed that a path is derived by Movement, i.e., the displacement of constituents [4]. I will extend the notion of path to Binding, i.e., the relation of anaphors and pronouns to their antecedents. The operation *Link* in (3) covers both.¹

For example, in the simplified phrase marker in (7a), there is a path derived by the movement of the Determiner Phrase (DP) *John* from the subject internal verb phrase (vP) position to the specifier of Tense Phrase (TP), the external subject position. There is also another Path derived by Binding of the reflexive pronoun *himself* by the DP *John* in the subject position. (7b) and (7c) illustrate further that path size counts for legibility.



- b. [_{TP} John [_{VP} ~~John~~ [_{VP} trusts himself]]]
- c. *[[_{TP} John's mother] [_{VP} ~~John's mother~~ [_{VP} trusts himself]]]

In the case of movement, a displaced constituent must asymmetrically c-command its copy; likewise for the binding of reflexive anaphors, the antecedent must asymmetrically c-command its anaphor. In (7a,b) the constituent in subject position asymmetrically c-commands its lexical anaphor *himself*; whereas this is not the case in (7c). Thus, it is generally assumed that the constitutive elements of a path (chain) must be in asymmetric c-command relation [4], [3]. This is the case for Binding as well as for Movement.

In the Asymmetry Theory, Movement and Binding are derived by the operation *Link* (α, β). The Paths derived by Movement and Binding relations can be given a unified account in terms of this operation.

Considering Binding relations, the difference between (7b) and (7c) illustrates that Binding is based on asymmetric c-command. In (7b) the DP *John* is the antecedent of the reflexive anaphor *himself*, and *John* asymmetrically c-commands the reflexive. However, the DP *John* cannot be the antecedent of the reflexive

¹ The idea that movement and binding are instances of a more general principle of which these are just instances has been discussed in [7], [8], [9]. My proposal differs from the preceding ones, as I do not take binding to be movement, but different instances of the generic operation *Link*, as defined in (3). One difference between movement and binding being that binding does not lead to the displacement of constituents, whereas movement does.

anaphor in (7c), since it does not asymmetrically c-command the anaphor.

Considering Movement, a displaced XP constituent must asymmetrically c-command its copy (strikethrough), which is the origin of the movement path. Thus, in (7a) the DP John in the specifier of TP asymmetrically c-commands its copy in the specifier of vP.

The asymmetric c-command restricts the properties of Movement and Binding, as well as it restricts the properties of paths. This follows from the Asymmetry Theory, since in this theory the operation Link (α, β) applies under asymmetric Agree, see (4).

3 Movement and Binding

Principles and conditions have been proposed to limit the domain of application of Movement, as well as the domain of Binding of pronouns and anaphors to their antecedents, as described in what follows.

3.1 Movement

Several principles and conditions on Movement have been proposed to account for the fact that shorter movements are preferred over longer ones, see (1b) vs. (1c), (5b) vs. (5c). The following paragraphs describe some of them.

Given the Principle of the strict cycle [10], movement applies first to the smallest domain where it can apply. For example, in (8a), the movement of the wh-DP constituent *which book* from the complement of the embedded clause to the subject position of that clause takes place before that constituent moves to the left periphery of the matrix clause. The example in (8b) is similar.

(8) a. [_{CP} What book did [_{TP} Mary think [_{CP} that [_{TP} ~~what book~~ was read ~~what book~~]]]]

b. [_{CP} Which students seem [_{CP} to [_{TP} ~~which students~~ have been contacted ~~which students~~ for the job]]]]

Cyclic domains are referred to in terms of bounding nodes, and languages vary with respect to bounding categories. For example, the bounding nodes in English are TP and DP, whereas CP and DP are bounding nodes in Italian, [12]. The example in (9) shows that in Italian TP is not a bounding node, whereas the example in (10) shows that CP and DP are bounding nodes.

(9) Tuo fratello [_{CP} a cui [_{TP} mi domando [_{CP} che storia [_{TP} abbiamo raccontato ~~tuo fratello~~]]] era molto preoccupato (It)

‘Your brother to whom I wondered what story they have told was very worried.’

(10)*Tuo fratello [_{CP} a cui [_{DP} la possibilita [_{CP} che abbiamo raccontato tutto a ~~tuo fratello~~ ...]]] (It)
‘Your brother to whom the possibility that they have told everything ...’

The Subjacency Condition [11], (11), expresses the generalization that movement cannot cross two bounding nodes. Subjacency accounts for the difference between cases such as (12b,c) and (12d). In (12d), the wh-DP *what* has crossed two TPs, thus violating Subjacency. This condition is not violated in the examples in (12b,c), where wh-movement crosses only one TP.

(11) Subjacency

No rule can relate x and y in the structure

... x... [α ... [β ... y ...] ...] ... x ...

iff α and β are bounding nodes.

(12) a. I wonder [_{CP} [_{TP} you will see who]]

b. I wonder [_{CP} who [_{TP} you will see ~~who~~]]

c. Bill wonders [_{CP} who [_{TP} ~~who~~ saw what]]

d. ?*What does [_{TP} Bill wonder [who [_{TP} ~~who~~ saw ~~what~~]]]

The generalization that shorter movements are preferred to longer ones is expressed in terms of the Minimality condition in [12], (13).

(13) Minimality

A movement operation cannot involve X_1 and X_3 over an X_2 which is identical to X_1 .

... X_1 ... X_2 ... X_3 ...

Minimality holds only if X_2 asymmetrically c-commands X_1 . Minimality covers cases such as (14), where X_2 asymmetrically c-commands X_1 . In (15), there is no asymmetric c-command relation between the relevant constituents, and the movement of a DP over a non-asymmetrically c-commanding DP is possible.

(14) a. John wondered [_{CP} who [_{TP} ~~who~~ saw what]]

b. *John wondered [_{CP} what [_{CP} who [_{TP} ~~who~~ saw ~~what~~]]]

c. *John wondered [_{CP} who [_{CP} what [_{TP} ~~who~~ saw ~~what~~]]]

(15) a. John wondered [_{CP} what [_{DP} whose sister ~~of who~~ said ~~what~~]]

b. John wondered [_{CP} who [_{TP} stories about what] might interest ~~who~~]

Furthermore, the movement of a head to another head position, for example, the movement of the auxiliary *will* in (16a) over the subject, is subject to the Head Movement Constraint [13], (17).

- (16) a. [_{CP} What [_C will] [_{TP} John [_T ~~will~~] buy]]?
 b. * [_{CP} What [_C buy] [_{TP} John [_T will] [_{VP} buy]]?

- (17) Head Movement Constraint (HMC)
 A head can move only to the next superior c-commanding head position.

HMC places a limit on the distance over which a head may be moved.

In the Minimalist Program [14], Subjacency, Minimality, and the HMC are subsumed under the Shortest Move condition, (18).

- (18) Shortest Move
 A constituent must move to the first superior, available, compatible position from its source position.

This accounts for the super raising violation, illustrated in (19a), where the subject *John* moved over the expletive subject *it*, Superiority effects, as in (19b), where *what* moved over *who*, as well as HMC violation, as in (19c), where *have* moved over the auxiliary *will*.

- (19) a. *John is likely [for it to seem [~~John~~-to have left]]
 b. *what [did you persuade [who to buy ~~what~~]]
 c. *have [John will ~~have~~ left by the time we get here]

As formulated in (18), Shortest Move imposes a locality condition on Movement operations themselves.

In recent Minimalist Theory [15], [3], the only conditions of the grammar are interface legibility conditions and principles of Economy. Thus, conditions such as (18) are no longer available. In this framework, the derivations proceed by phases (cyclic domains of computation and interpretation such as propositions (CP)), and movement is triggered by uninterpretable feature checking. For example, in the case of wh-movement, an uninterpretable wh-feature in the complementizer (C) position must be checked (eliminated) by the movement of a wh-expression in the specifier position of CP (Complementizer Phrase). A notion of path defined in terms of phasal feature checking is formulated in section 4.

3.2 Binding

Locality conditions also cover the relation between anaphors and pronouns and their antecedents. The Binding Theory [16] (20) applies in local domains, and requires that an anaphor be bound by an antecedent in its local Binding Domain (BD).

- (20) Binding Theory
 A. An anaphor is bound in its BD.
 B. A pronominal is free in its BD, such as a proposition.

- (21) a. α is bound by β iff
 α and β are co-indexed, and
 β asymmetrically c-commands
 b. α is free iff α is not bound.

The Binding Theory correctly predicts that an anaphor must be bound in the same domain in which a pronoun must be free. This can be seen with simple cases such as the ones in (22). The Binding Domain of the examples in (22) is the propositional domain, and within this domain the anaphor *himself* is bound by the DP *John* in (22a), whereas in (22b) the pronoun *him* is free.

- (22) a. John trusts himself.
 b. John trusts him.

The examples in (23)-(25) show that asymmetric c-command is a necessary condition for Binding, but not co-reference. Thus in (23), the pronoun *he* is co-referential with a non-asymmetrically c-commanding antecedent *John*, which is in a relative clause. The example in (24) shows that a pronoun *he* cannot bind a constituent in its asymmetric c-command domain. The example in (25) shows that even though *John* precedes the pronoun *he*, it does not asymmetrically c-command *he*, and co-reference is possible between these two constituents.

- (23) [Who that John₁ knows]₂ does [he₁ like t₂]
 (24) *He₁ likes [everyone that John₁ knows]
 (25) [Everyone that John₁ knows]₂ [he₁ likes t₂]

Interestingly, the conditions that block movement also apply to lexical anaphors. This is the case for reflexive pronouns, as the examples in (26) illustrate, as well as for reciprocals, such as *each other*, (27). In (26a), because the embedded clause is not tensed, the BD of the lexical anaphor *himself* is the whole proposition, and within this domain, the reflexive pronoun is bound by an asymmetrically c-commanding antecedent DP *John* located in the matrix subject position. In (26b), the BD of the reflexive anaphor

himself is the embedded clause. The principle B of the Binding Theory correctly predicts that (26b) is not well-formed. In (26c), the subject of the embedded clause *Mary* is not a possible antecedent for the reflexive anaphor *himself*; moreover this subject blocks the binding of the reflexive by the DP subject *John* in the matrix clause.

- (26) a. John believes [himself to be a genius]
 b. *John believes [[that himself is a genius]]
 c. *John believes [Mary to like himself]

- (27) a. The boys believed [each other to be geniuses]
 b. *The boys believed [[that each other are geniuses]]
 c. *The boys believed [Mary to like each other]

The examples in (28) illustrate that pronouns are in complementary distribution with reflexive/reciprocals.

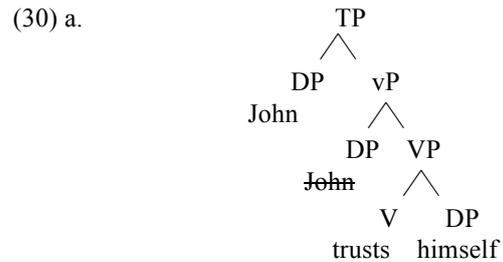
- (28) a. *John believed [him to be a genius]
 b. John believed [[that he is a genius]]
 c. John believed [Mary to like him]

Even though Binding relations are derived by Movement, and thus there is no uninterpretable feature that must be checked and erased, there is however uninterpretable feature valuing (feature sharing) between an anaphor and its antecedent. I assume that an anaphor is an element that has no independent reference, and that the binding of the anaphor to an antecedent values the anaphor. I take Binding to hold under asymmetric Agree, which does not lead to feature erasure, but to feature valuing, [17].²

- (29) Agree-erase : Agree leading to the elimination of uninterpretable features.

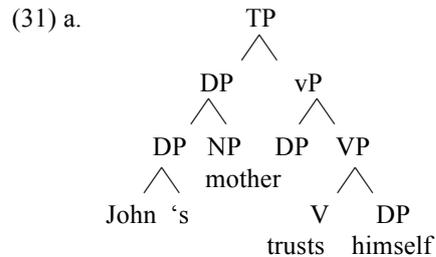
Agree-value : Agree leading to the valuing of uninterpretable features.

Given the distinction between Agree-delete and Agree-value in (29), paths can be identified for Binding relations, which are derived from the operation Link (α, β). Thus, considering the example in (7) above, repeated here in (30), the Path for the reflexive anaphor *himself* in (30a,b) is (30c).



- b. [_{TP} John [_{vP} ~~John~~ [_{vP} trusts himself]]]
 c. Path(*himself*)= {VP, vP, TP}

Considering the example in (31a,b), where the DP *John* does not asymmetrically c-command the reflexive anaphor, and where the latter cannot take the former as its antecedent, the path for the reflexive is (31c), and it is longer than in (30c). The difference between (30) and (31) could be seen as following from the view that shorter paths are preferred to longer ones, see also [4], [5], [12].



- b. *[[_{TP} [_{DP} John's mother]] [_{vP} ~~John's mother~~ [_{vP} trusts himself]]]
 c. Path(*himself*)= {DP, DP, TP, vP, TP}

Moreover, the Path(*herself*) in the example in (32b) where the reflexive anaphor *herself* agrees with the morphological features of the DP *John's mother* is shorter than the Path *himself* in (31c).

- (32) a. [_{TP} [_{DP} John's mother]] [_{vP} ~~John's mother~~ [_{vP} trusts herself]]]
 b. Path (*herself*)= {DP, TP, vP, TP}

Thus, given the difference between Agree-delete and Agree-value, the notion of Path(x) is derived by Link (α, β), and covers Movement and Binding. As it is the case for Movement, with Binding, short paths are preferred to long ones. Long paths give rise to semantic gibberish in the case of Binding, and to syntactic gibberish in the case of Movement.

4 Path, Shortest or Minimal?

In [4], a path is defined as in (33), and a path length as in (34). Paths are compared via the sub-set relations.

² See [18],[19] for discourse pronominal anaphora resolution using asymmetric Agree. The system is based on the identification of proper sub-set relations between the sets of formal and semantic features of antecedents and pronouns in minimal discourse domains.

(33) Path: a Path is the union of the sets of maximal projections (XPs) that dominate the target and the launch site.

(34) Path length: Path A is shorter than Path B, iff Path A is a proper sub-set of Path B.

This covers the cases of XP movement such as the one discussed above in (1), repeated here in (35). Considering (35a) and (35b), the Path(*who*) is a proper sub-set of the Path(*what*). Assuming that in linguistic expressions, path length should be minimized, it follows that (35a) is well-formed, but not (35b).

- (35) a. Who saw what?
Path (who)= {TP, CP}
b. *What did who saw?
Path (what)= {TP, Cp, vP, VP}

I would like to take a different view of the properties of paths in grammar, and propose the following: paths are not generally compared, thus there is generally no shortest path, however the grammar determines paths to be minimal. One argument in favor of this view is that obstacles (intervening constituents) cannot be avoided by following alternative paths. These cases include situations where a subject, a wh-constituent, or another head in the case of Head-movement intervenes between the origin and the putative endpoint of a path, see (19) above, repeated here in (36) for convenience.

- (36) a. *John is likely [for it to seem [~~John~~ to have left]]
b. *what [did you persuade [who to buy ~~what~~]]
c. *have [John will ~~have~~ left by the time we get here]

Under the proposed view, cases such as (36a) are excluded because the uninterpretable features they include are not checked in their next higher phase. Thus, they fail to meet Full Interpretation (FI), which requires that every element have only interpretable features at the Interfaces, phase by phase. Cases such as (36b,c) are excluded by Economy: the constituents have no relevant uninterpretable feature to check and delete by movement out of their phase. The movement is not necessary, and thus by Economy consideration it must not take place.

I propose that the length of a Path is determined by the operations of the grammar applying phase by phase, under Agree-erase and Agree-value, and define the notion of Minimal Path as follows:

(37) Minimal Path(x)= The union of the set of maximal projections dominating x and the constituents with which it asymmetrically Agrees.

Given that the derivations proceed by phases (CP, DP) and that phases are subject to the Impenetrability Condition, according to which only the edge (specifier) and the head of the phase are accessible from outside, paths are bounded. Minimal paths are identified phase by phase, and can be extended to an adjacent phase only from the peripheral position of a phase.

Given the Minimal Path, determining the optimal path for a constituent is not a problem that is NP-complete, since the properties of a path are determined by the grammar. Thus, in the proposed view, the grammar does not compute the length of a path. Path length follows from the operations and the conditions of the grammar.

5 Natural Language Processing

Notwithstanding their complexity, natural languages are easily tractable by humans. Why is this so? A natural answer to this question is that the Minimal Path contributes to reduce the computational complexity that may arise. Thus comparing derivations and measuring paths in order to determine the shortest one amongst the set of potential ones is generally not necessary. By Economy, if it is not necessary, it does not generally occur. Paths are compared only when needed, that is, when a decision point arises in the derivation, or at the interfaces. If paths are compared, derivational complexity arises, and this leads to a decrease of human performance.

One example of derivational complexity is the case of the Delay of Principle B Effect (DPBE). The DPBE has been discussed in various studies, including [20], that show that Children around age 5 seem to violate Principle B of the Binding Theory, (20), when the antecedent of the pronoun is a name, but not when the antecedent is a quantifier ([20], [21] for English). Thus, for sentences such as (38a) the children interpret the pronoun *her* as bound to the antecedent *Mama Bear* 50% of the time, thus performing at chance level; whereas, they show an adult-like performance (85%) in sentences such as (38b), where the antecedent is a quantifier. Moreover, children speaking a Romance Language have a 64% level of performance for sentences such as (39), [22].

- (38) a. Mama Bear touches her.
b. Every bear touches her.

(39) Gianni lo achiuga. (It)
Gianni him-dries off
'Gianni dries him off.'

In [23], the DPBE is accounted for by assuming that the phenomenon arises when children have to compare two alternative derivations (binding or coreference) for equivalence. In both cases the result is similar: the children take guesses in the process of interpreting the anaphoric dependency, thereby performing at chance level.

7 Conclusion

The notion of path is part of the grammar and is determinant for the well-formedness and the interpretation of linguistic expressions. The proposed notion of Minimal Path covers both Movement and Binding and avoids the complexity brought about by notions such as shortest path which require comparing derivations and measuring path length. Given that Minimal Paths are defined in terms of asymmetric Agree, mere linear precedence between words is not at stake, as in [24], [25]. Derivations are not generally compared, and when they are, the comparison leads to an increase of the derivational complexity and a concomitant non-optimal human linguistic performance.

6 Acknowledgments

I thank N. Hornstein, C. Aguero-Bautista, and two anonymous reviewers for comments on earlier versions of this paper. This work is supported in part by funding from the SSHRC of Canada to the Interface Project 214-2003-1003, and by a grant from the FQRSC to the Dynamic Interfaces Project 103690.

References

- [1] Dijkstra, E. W., A Note on Two Problems in Connexion with Graphs, *Numerische Mathematik* 1, 1959, pp. 269–271.
- [2] Seda, M., Geometric Data Structure and Their Selected Applications, *Computer Science* Vol. 11, 2006, pp. 61-66.
- [3] Chomsky, N., On Phases. To appear in Carlos Otero et al. eds. *Foundational Issues in Linguistic Theory*, Cambridge, Mass.: The MIT Press, 2006.
- [4] Hornstein, N., Minimality, C-command and Labels, ms. University of Maryland, 2005.
- [5] Kayne, R., *Connectedness and Binary Branching*, Kluwer Academic Press, Amsterdam, 1984.
- [6] Di Sciullo, A. M., *Asymmetry in Morphology*, Cambridge, Mass.: The MIT Press, 2005a.
- [7] Hornstein, N. *Move! A Minimalist Theory of Construal*, Blackwell, Oxford, 2000.
- [8] Lidz, J. & W. Idsardi, Chains and Phono-Logical Form, *Penn Working Papers in Linguistics*, Vol. 5, No. 1, 1997, pp. 109-126.
- [9] Swart, J.-W. Issues Relating to a Derivational Theory of Binding, in S. D. Epstein & T. D. Seely, eds., *Derivation and Explanation in the Minimalist Program*, 2002, pp. 269-304.
- [10] Chomsky, N., *Syntactic Structures*. Mouton, 1957.
- [11] Chomsky, N., *Reflections on Language*, New York: Pantheon, 1975.
- [12] Rizzi, L., *Relativized Minimality*, Cambridge, Mass.: The MIT Press, 1990.
- [13] Travis, L., Parameters and Effects of Word Order Variation, Ph.D. Dissertation, MIT, distributed by MITWPL, 1984.
- [14] Chomsky, N., *The Minimalist Program*, Cambridge, Mass.: The MIT Press, 1995.
- [15] Chomsky, N., Derivation by Phase, in M. Kenstowicz, ed., *Ken Hale: A Life in Language*, Cambridge, Mass.: The MIT Press, 2001, pp. 1-52.
- [16] Chomsky, N., *Lectures on Government and Binding*, Dordrecht: Foris, 1981.
- [17] Di Sciullo, A. M., Affixal Scope and Precedence Relations, in K. Choi & C. Yim, eds., *Proceedings of the 7th Seoul International Conference on Generative Grammar*, Korean Generative Grammar Circle, Konkuk University, 2005b, pp. 135-154.
- [18] Di Sciullo, A. M., Domains of Argument Structure Asymmetries,” in N. Callaos, W. Lesso, & K. Horimoto, eds. *Proceedings of the 7th Multiconference on Systemics, Cybernetics, and Informatics (SCI 2005)*, Orlando, Florida, 2005c, pp. 316-320.
- [19] Di Sciullo, A. M., Processing Bound Pronouns, *Computer Science*, Vol. 11, 2006, pp. 74-79.
- [20] Chien, Y.-C. & K. Wexler, Children's Knowledge of Locality Conditions in Bindings as Evidence for the Modularity of Syntax and Pragmatics, *Language Acquisition*, Vol. 1, 1990, pp. 225-295.
- [21] Grodzinsky, J. & T. Reinhart, The Innateness of Binding and Coreference, *Linguistic Inquiry*, Vol. 24, 1993, pp. 69-102.
- [22] McKee, C. A Comparison of Pronouns and Anaphors in Italian and English Acquisition, *Language Acquisition* Vol. 2, 1992, pp.21-54.
- [23] Di Sciullo, A. M. & C. Aguero, The Delay of Principle B Effect and its Absence in Some Languages, to appear in *Language and Speech*, 2006.
- [24] Bunescu, R. C. & R. J. Mooney, A Shortest Path Dependency Kernel for Relation Extraction, in *Proceedings of the Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, Vancouver B.C., 2005, pp. 724-731.
- [25] Solé, R., B.C. Murtra, S. Valverde, & L. Steels, Language Networks: Their Structure, Function and Evolution, in *Santa Fe Institute Working Paper*, 2005.