# Pro-active Parsing of Korean Scrambling[1]

**Jieun Kiaer and Ruth Kempson**
**King's College London**

This paper argues that the processing of Korean is incremental and moreover 'pro-active'. Though the verb, which triggers the actual composition, comes at the end, the parser of Korean can build up a structure incrementally from the very beginning with the help of prosody and case. In particular, based on scrambling (e.g., local, long-distance, multiple long-distance scrambling), we shall show that on-line interaction between prosody and case specification along with structural abduction play a crucial role in incremental structure building in Korean. In this paper, we adopt the parsing-based framework, Dynamic Syntax (Kempson et al. 2001, Cann et al. 2005) and show that the characteristics of pro-active parsing of Korean are consequences of monotonic-structure growth processes from left to right.

## 1. Incremental Parsing

The incremental nature of head-final languages like Korean or Japanese is hard to capture, unlike languages like English. For instance, in English, a verb occurs early and thus helps the parser to guess the overall structure at the early stage of parsing, whereas in Korean, as the verb comes at the last position, this sort of guess seems to be hard. Because of its head-final nature, it is often assumed that in Korean or Japanese, hearers refrain from decisions of structure-building until they get to the end of the whole sentence or at least until any verb is parsed (See Pritchett 1992 among others for head-driven parsing approaches). However, recently, Inoue and Fodor (1995) among others suggested that the parsers of Japanese, just like English parsers, make structural decisions as they go along parsing without any delay. Though the parsers can be misled by making such decisions too early, Fodor and Hirose (2003:195) claimed that the parser prefers guessing to waiting. Crucial evidence of incremental parsing in Japanese has been given by Mazuka and Itoh (1995) and Miyamoto and Takahashi (2001) among others. The eye-tracking study of Mazuka and Itoh (1995) showed that even before reaching any verb, a series of three nominative NPs at the beginning of the sentence is read significantly slower than compared to those with a series with distinct case marking, as *NOM ACC DAT*. They take this as evidence that some syntactic processing is performed 'immediately', since the two groups should be read identically if no processing at all takes place until a verb is encountered. Miyamoto and Takahashi (2001) introduced a Typing Mismatch Effect (TME) as a diagnostic of where a *wh* phrase is interpreted. They showed slower reading time (TME) for verbs marked with *-to*, the declarative complementizer, rather than those with *-ka*, the question marker. Based on this, they suggested that when Japanese speakers parse a *wh* phrase, they expect a question particle to follow in the same clause. If such anticipation is contra-indicated, parsing is delayed and TME occurs. These works imply that parsing in Japanese is incremental and that structural decisions are made step-by-step rather than all at the end. Moreover, this work implies that the parser builds up a certain anticipation for the upcoming structure.

In this paper, we argue that parsing Korean is incremental regardless of its head-final nature, pro-drop property and relatively free NP ordering. This is possible because the parser incorporates case and prosody information directly from decoding the parse sequence. Furthermore, as we shall see, Korean parsing is pro-active: the parser may update the emergent structure by hypothesizing future structure, possibly far ahead of the actual morphological choice point (i.e.,the parsing of the verb, complementizer or relative clause marker, etc) at which the structure can be determined. In the following sections, we will show how this pro-activeness of Korean parsing can be successfully captured in the parsing-oriented framework of Dynamic Syntax.

[1]Thanks to Carlos Gussenhoven, Wilfried Meyer-Viol, Ad Neeleman, Mary Dalrymple, Devyani Sharma and the audience at WCCFL 24 for their useful comments.

 Proceedings of the 24th West Coast Conference on Formal Linguistics, ed. John Alderete et al., 209-217. Somerville, MA: Cascadilla Proceedings Project.

### *1.1. Pro-active Parsing via Case and Prosody*

According to Kim (1999), Korean parsers can incorporate case information directly into structure-building without any delay. However, case information alone is not enough to disambiguate possible parse sequences in on-line parsing because of various scrambling constructions in Korean. So, in hearing (1), the human Korean parser will know that *Yuna-rul* is an object of a clause. Yet, it is still quite unspecified as to whether *Yuna-rul* is an object of a simple clause or it is a scrambled object.

(1) *Yuna-rul...*
Yuna$_{ACC}$

Given the availability of even multiple long-distance scrambling, it might naturally be expected that the more NPs are parsed, the more possible parse sequences become available. However, contrary to any such expectation, in parsing (2), with two more NPs, even before reaching any verb, the parser achieves only one possible parse sequence, in which every NP is interpreted locally to each other.

(2) *Yuna-rul Mina-ka Jina-ekey ... [ACC NOM DAT]*
Yuna$_{ACC}$ Mina$_{NOM}$ Jina$_{DAT}$

We assume that this is because the parser builds up a structure locally within a single structure unless (i) a new structure is initiated by a nominative case marker or (ii) an intonational break (i.e., Intonational Phrase boundary)[2] separates the structure currently being built from the previous one. The evidence of local structure-building can be found in self-paced reading of (3), which has two dative marked NPs.

(3) *Yuna-ekey Mina-rul Jina-ekey ... [DAT ACC DAT]*
Yuna$_{DAT}$ Mina$_{ACC}$ Jina$_{DAT}$

In a self-paced reading task,[3] delay is caused at the point of parsing the second dative NP *Jina-ekey*. Given that structure-building is local, unless it's debarred by prosody, delay is naturally expected as *Yuna-ekey* is already fixed in the current structure. Thus, *Jina-ekey* has no place at which it can be fitted in (Filled-Gap Effect, Stowe 1986). Delay at *Jina-ekey* is a direct consequence of reanalyzing the structure. However, no such delay occurs if there is an intonational break after *Yuna-ekey* in a self-paced listening task, as the intonational break signals the existence of two separate structures.

(4) *Jina-ka apu-si-n... [NOM VERB$_{HON,REL}$]*
Jina$_{NOM}$ sick$_{HON,REL}$

Similarly, in a self-paced reading of (4),[4] the parser will slow down at the point of parsing *-si*, the honorific marker, because of honorification mismatch between *Jina-ka* and *apu-si* 'to be sick', as *Jina* is not a form for which an honorific verb is suitable. And indeed, such a putative parse must be ruled out, as the relative marker *-n* confirms. Yet, there is no delay at *-si* if there is an intonational break after *Jina-ka*, because the parser is alerted to the necessity of two discrete structures from that moment, given the prosodic assurance that *Jina-ka* and *apu-si-n* are not in the same structure. This also shows that in Korean, prediction of relative clause structure can occur far earlier than the definitive choice point, which is the post-verbal relative clause marker or the following head of the relative clause.

All these pieces of evidence lead in the same direction, indicating that parsing Korean is pro-active, with hearers able to make structure-building decisions not only at the end, but also far-ahead of the actual choice points in a pro-active manner. As we have seen above, this is possible because, right from the beginning, the parser is making use of information not only from case but also prosody, and other possible default strategies. We call all such structural choice-making (structural) abduction (Section 3). Notice that in the examples (1) to (3), the parser can be seen to be making decisions even without knowing the upcoming verb. In the following section, we show how the framework can capture these various displays of incrementality in Korean parsing.

[2]The IP boundary in Korean is indicated by (i) lengthening of the final syllable of the IP, (ii) a boundary tone on the final syllable of the IP and (iii) the optional pause afterwards (see Jun 1993).

[3]See Kiaer (in prep) for detailed discussion of experimental results.

[4]HON means an honorific and REL means a relative clause marker.

## 2. Dynamic Syntax: Pro-active Approach

The motivation of Dynamic Syntax (DS) is to capture how the human parser builds up a syntactic structure incrementally from left-to right in real-time parsing. The main challenge of this framework is to reflect how human parsers can manipulate partial information at each step of parsing to draw a bigger picture of the meaning of the string as early as possible. The DS framework adopts representationalist assumptions about the nature of mind (Fodor 1983) and assumes that semantic interpretation is given as a structural representation of content, with trees representing predicate-argument structure in which the top node of a tree is decorated with a propositional formula and each dominated node is a subterm of that formula, with type-specifications indicating how the parts combine. The framework assumes furthermore that the human parser displays inferential or anticipatory abilities at each step of parsing until a full proposition is achieved. In DS, the unfolding of a structure, and the making of pro-active guesses for the upcoming structure, are driven by (i) a set of language-universal computational actions which induce structural options, (ii) lexical actions of each word, (iii) pragmatic actions (such as substituting a value from context for anaphoric expressions).

### *2.1. Underspecification of Content and its Update*

Because DS is committed to building representations of content rather than structure inhabited by words, it is able to characterize the flexible word order and pro-drop nature of head-final languages (e.g., Korean and Japanese) by assuming that the verb projects a full propositional structure, whose arguments are semantically underspecified (as though projected from anaphoric expressions). Such actions are part of the specification of each verb. For instance, *iki*-'beat' projects the partial tree in Fig 1, with two argument nodes decorated with place-holding metavariables *U* and *V* as below (the pointer, $\diamondsuit$ , indicates the node next to be developed).

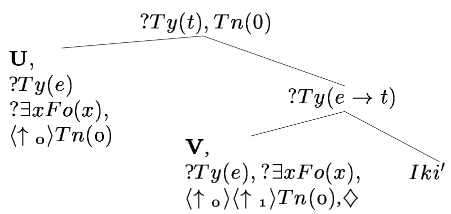



**Fig 1. Propositional structure projected by** $Iki'$

The tree gives no indication of order, as this is a representation of content, inhabited by concepts, not by words of the string. $?Ty(t)$ indicates the overall goal of establishing a proposition of $Ty(t)$ with the sub-goal of a predicate, $?Ty(e \rightarrow t)$. The concept of requirement,?X for any X, is the central concept of the framework: all requirements, whether of content, structure, or type, have to be met by the end of a sequence of parse actions – notice that the argument nodes in Fig 1 equally have a requirement for update of content ($?\exists x Fo(x)$), which can be met either by update from within the structure-building process or from context. Thus, by defining words as projecting procedures for building structural representations of content (and not merely some concept-denoting term), an account of pro-drop is naturally made available, and one that is notably more successful than the multiply-typed assignment of categories to words as suggested in other constraint-based frameworks (e.g., Combinatory Categorial Grammar).[5]

### *2.2. Structural Underspecification and Update*

Central to this building up of tree-structure representations of content is the articulation of different forms of structural underspecification and update within the construction process. Such updates can be motivated or driven by computational, lexical or pragmatic actions. Though such actions may construct

[5]Though Multi-modal CCG (Baldridge 2002) captures the ordering flexibility by specifying an ordered set of arguments, optionality of arguments is not captured to date.

'fixed' tree-relations, as in other frameworks, an important property of the tree-growth system is that nodes in the partial tree may be introduced without having any particular relation initially fixed; and their syntactic position may be resolved later on in on-line parsing. This is the basis of long-distance dependency, where the concept of movement is replaced by one of positional underspecification, analogous to anaphora as an alternative form of underspecification.[6] In virtue of having adopted a tree-logic for describing the trees, such underspecified relations can be described in terms of $\langle\uparrow *\rangle$, 'X holds at some dominating node'(see the LFG concept of functional uncertainty, Kaplan and Zaenan 1989).[7] The process of seeking a fixed position for the unfixed node is then presumed to take place across an arbitrary sequence of daughter relations. There are three types of Adjunction, reflecting different locality variations in the domain within which update of the underspecified relation is to take place (again analogous to anaphoric expressions). First of all, such underspecifications may require local update within an individual predicate-argument structure: 'Local * Adjunction'. This Local * Adjunction is defined to explain local scrambling in Korean as it can capture flexible ordering of NPs within its local domain, each being taken to decorate some locally underspecified node which is then immediately updated by its attendant case relation (which we return to shortly). Secondly, to model the updating procedure that can be presumed to be construed across the clause boundaries, as displayed by regular instances of long-distance dependency, '* Adjunction' is defined as a process introducing an unfixed node whose hierarchical position in the emergent structure must be resolved within an individual tree structure. Lastly, structural update can take place within some overall construction process, possibly across a sequence of trees. This loosest form of update is termed 'Generalised Adjunction'. Intuitively, what it does is to posit a very weak and arbitrarily dependent relation between two local structures. These processes may, furthermore, feed each other, so a step of *Adjunction may feed a sequence of steps of Local *Adjunction, as we shall see in the analysis of multiple long-distance scrambling.

The concept of case in DS interacts with the structural growth process, constraining updates of underspecified structure projected by NPs. For instance, in a local scrambling sequence as in (5), when the parser processes *Mina*, with no other specification, *Mina* may turn out to decorate a subject, an object or a dative. In DS terms, if such a NP is to decorate an unfixed node, it will be decorating a structure whose relative position is merely $\langle\uparrow *\rangle Tn(0)$ ('the root node dominates this node'). However, given that *Mina* is nominative case-marked, this requirement can be very much narrowed down, since it imposes a requirement that the node is immediate daughter of a $Ty(t)$ requiring node: $?\langle\uparrow_0\rangle Ty(t)$. In the case of Local *Adjunction, such a narrowing will indeed determine a fixed result. If accusative case-marked, the requirement is that the node be immediate daughter of a predicate-requiring node: $?\langle\uparrow_0\rangle Ty(e \rightarrow t)$. And so on. Derivations involving Local *Adjunction, the local scrambling phenomenon, that is, follow the pattern displayed in Fig.2 below.

(5) *Mina-ka* *Jina-rul* *iki-ess-e* *[S O V]*
Mina$_{NOM}$ Jina$_{ACC}$ beat$_{PAST,DECL}$

(6) *Jina-rul* *Mina-ka* *iki-ess-e* *[O S V]*
Jina$_{ACC}$ Mina$_{NOM}$ beat$_{PAST,DECL}$
'Mina beat Jina'

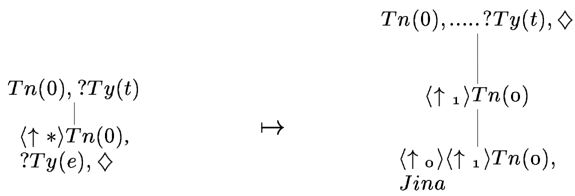
As Fig 2 displays, each NP is parsed by presuming on a step of Local *Adjunction inducing a weakly specified tree relation; and this relation is enriched into the particular argument node within the emergent structure by the particular case specification. The effect is that argument nodes can be introduced into a structure independent of the projection of the predicate itself, allowing incremental projection of structure from the words, despite the verb not occurring until the end of the string. Given this, the account thus anticipates no difference in complexity of parsing (5) or (6). [8]

[6] To express island constraints, relative clause modification and co-ordination is associated with the building of paired quasi-independent structures, as the underspecified tree-relation in these cases is not to able to be updated across such discrete paired structures (see Kempson et al. 2001, Cann et al. 2005).

[7] The tree itself is described by a modal tree logic (LOFT: Blackburn and Meyer-Viol 1994) in which relations are described between nodes in a tree, with arrow $\langle\uparrow\rangle X$ indicating that X holds at the mother of some current node, $\langle\downarrow\rangle X$ indicating that X holds at its daughter, etc. Two distinct daughter relations are introduced according as $\langle\downarrow_0\rangle$ the daughter is an argument, $\langle\downarrow_1\rangle$ the daughter is a functor

[8] Yamashita (1997) found no difference either in a self-paced reading experiment or in a lexical decision task, though there are also some contradictory results as in Mazuka, Itoh and Kondo (2002).

(i) Parsing *Jina-rul* :*-rul* updates $\langle\uparrow *\rangle Tn(0)$ into $\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(o)$ via $?\langle\uparrow_0\rangle Ty(e\rightarrow t)$




(ii) Parsing *Mina-ka*:*-ka* updates $\langle\uparrow *\rangle Tn(0)$ into $\langle\uparrow_0\rangle Tn(o)$ via $?\langle\uparrow_0\rangle Ty(t)$

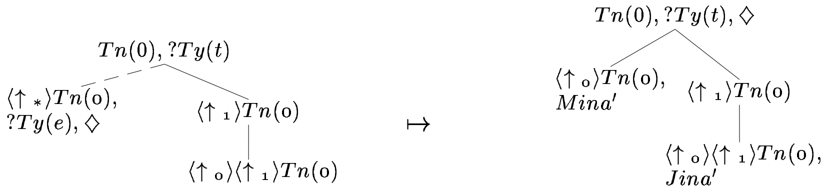



**Fig 2. Local * Adjunction and parsing the NP sequences in (6)**

## *2.3. Local and Non-local Update via Adjunctions*

The availability of the three processes of *Adjunction will of course mean that there will be more than one alternative for any initial step of parsing long-distance scrambling construction in Korean, and, as we shall see, the role of intonational break is to eliminate the disjunctions which these alternatives set up. More specifically, the parser will have two choices in parsing a case-marked NP as to whether to fix it immediately to its on-going local-structure via Local * Adjunction and immediately enrich it in virtue of case,[9] or to assume through application of * Adjunction that its position will be resolved at some later juncture in the parse process. With no intonational indication to the contrary, the parser will fix the node locally to the on-going structure via Local * Adjunction as in local scrambling shown in Fig 2. However, if a case-marked NP is followed by an intonational break, we take this as indication to the parser that such a sequence of steps of Local *Adjunction is contra-indicated, indicating instead that the expression just parsed is to decorate a node introduced by *Adjunction, forcing the relative independence and later resolution of the underspecified relation. This might seem to immediately lead to an overly flexible system, which contradicts incrementally restricted (from left to right) structure-building processes in Korean. Yet, updating structures in Dynamic Syntax is not unrestricted, as we shall now turn to.

## *2.4. Restrictions on Structure-Building: One unfixed node of a type at a time*

As speakers will confirm, updating is not totally free: indeed in (7), the parser gets stuck while processing the second dative *Kiho-ekey*, given the indicated intonation, which, by assumption, shows that *Jina-ekey* and *Pen-ul* form one constituent and thus need to be interpreted in the same local structure. Given that the constituent built by *Jina-ekey* and *pen-ul* is to be interpreted within the embedded clause, as only the embedded verb can take a concrete object, the parser is expected to register some slow-down effect in facing the second dative NP *Kiho-ekey* found in the same clause (Filled-Gap Effect).

(7) *Jina-ekey Pen-ul BREAK Mina-ka Kiho-ekey cwuesstako haysse.*
Jina-DAT Pen-ACC BREAK Mina-NOM Kiho-DAT gave said
Intended off-line reading ‘Mina said to Jina that she gave a pen to Kiho.’

[9]Given the restriction that only one unfixed node of a type can be available in any partial tree (see section 2.4), the only successful sequences of NPs construed as co-arguments must lead to such immediate fixing.

One of the constraints imposed by DS is that though all three adjunctions may be used in a single structure-building process, nevertheless the framework enforces the assumption that only one unfixed node via any adjunction of any one type can be built at a time, since such underspecified relations are only distinguishable through the tree node address arising from the particular combinatorial action. This is self-evident with the familiar fixed tree relations, but it applies equally with this underspecified relations. And it constitutes the basis of the slow-down effect (TME) in processing of (7) with the indicated intonation.

## 3. Sample Parsing

Bearing this in mind, we now turn to the more detailed modelling of the various interpretations and the way they are indicated by intonation in Korean. We regard an intonational break (=BREAK) as in effect a lexical action. The core action of the BREAK is to trigger a non-local update so that the already parsed lexical item cannot be construed as providing a term in the same structure as that associated with the following lexical item. Putting this assumption together with the assumed trio of adjunction processes, we are now in a position to show how DS with its structural growth dynamics can reflect the pro-active parsing of (8). (8) has a single long distance reading when there is an intonational break after *Yuna-rul* as in (8a), but multiple long-distance reading as in (8b) when the break appears after *Mina-ka*.

(8) *Yuna-rul Mina-ka Jina-ka ikiessta-ko haysse*
Yuna$_{ACC}$ Mina$_{NOM}$ Jina$_{NOM}$ beat$_{COMP}$ said
(a) 'As for Yuna - Mina said that Jina beat.':single long scrambling
(b) 'Jina said that Mina beat Yuna': multiple long-distance scrambling

### *3.1. Single Long-distance Scrambling*

Firstly, we will show how the parser can establish the single-long distance reading as in (8a). As the opening step of a parse, the parser has two options of structure-building for *Yuna-rul*, the operation of *Adjunction or Local * Adjunction, either one of which will enable the first NP to be processed. However, when the intonational break is parsed, the parser can hold the object NP *Yuna-rul* as decorating a node introduced by * Adjunction as in Fig 3, discarding any possible parse sequence involving Local *Adjunction, leaving the subject-marked *Mina-ka* alone to be parsed using Local *Adjunction and case-driven update:

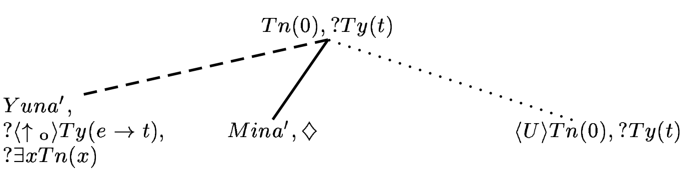



**Fig 3. Parsing *Yuna-rul BREAK Mina-ka***

With the parser then encountering a sequence of two nominative NPs *Mina-ka Jina-ka*, two discrete structures must be assumed, one at a lower level of embedding than the other. The first nominative marked NP will have induced a step of Local *Adjunction, immediately identified as subject of the matrix clause, updating that underspecified relation to a fixed daughter relation as in Fig 3. But to parse the second nominative-marked sequence, the parser has only one remaining option if the constraint of only one unfixed node of any one type is to be respected (See 2.4), and that is to introduce the required subordinate relation by Generalised(=U) Adjunction.[10] So, by this move, an indefinitely embeddable

[10] $\langle U\rangle Tn(0)$ indicates that the root node is arbitrarily higher in the tree than the node so decorated.

structure is induced as a dotted line in Fig 3 shows. Now in order to proceed and provide an enrichment of the initially underspecified tree relation, with its object-marked node, the hearer has to guess ("abduce") a fixed relation, and the parse duly reflects this (see Fig 4), determining a relation of immediate propositional subordination from the matrix structure for the predicate-argument array projected by *-iki* later:

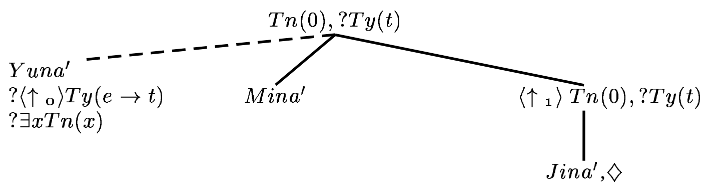



**Fig 4. Parsing *Yuna-rul BREAK Mina-ka Jina-ka***

With this move, and the second step of Local *Adjunction fixing a subject relation within that embedded structure, the verb *iki-* can then be parsed to project a full template structure with two argument nodes as in Fig 5.

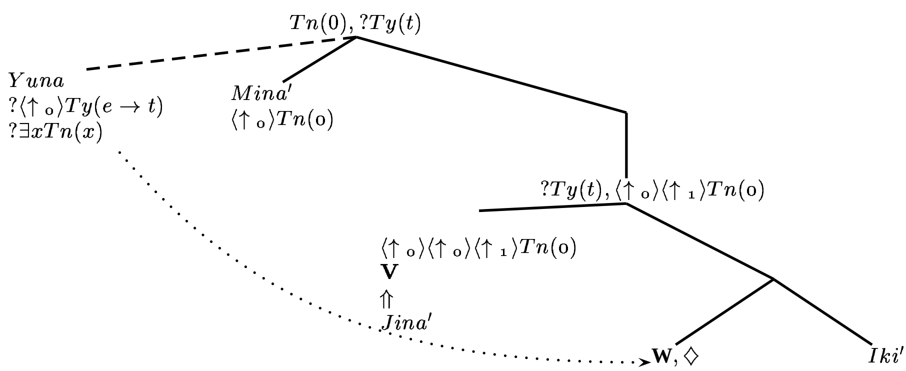



**Fig 5. Parsing *Yuna-rul BREAK Mina-ka Jina-ka Iki-***

As Fig 5 shows, once the full template of structure is projected by the parsing of *iki-*, both its argument nodes can be established. V is immediately updated into *Jina-ka* as in Fig 5 by the lexical action of nominative marker *-ka*. The two apparently discrete subject nodes then immediately collapses. With this unfolded structure, the relation of the unfixed node decorated by *Yuna-rul* is updated to become the object of *iki-* as in Fig 5. Though our system doesn't specify where the underspecified node has to be merged as all general processes are optional, it is notable that a second, twinned, break commonly occurs after the complementizer *ko-*. What we suggest is that whereas the first break is a signal to the parser that the just-parsed lexical item needs to wait to find its position, the second break signals the twinned resolution process that the update associated with such an initial underspecification must have been resolved (indeed it must have been in order for the actions of *-ko* to take place. See Kiaer (in prep) for detailed discussion). The parser, that is, unifies the dangling unfixed node initially introduced with the object argument node decorated with W as this step satisfies the open requirements on these two nodes. On the one hand, the unfixed node decorated with *Yuna'* gets updated to a fixed structural position: on the other hand, the object-denoting node of *iki'-* is provided with the formula *Yuna'*. This hypothesized update is confirmed at the parse of the complementizer *-ko* (notice that this is not the point when the abduction step is made, but confirmed). At this point, before completing one local structure, we can see how, by using the twinned concepts of underspecification plus update, structural update and anaphoric update can be shown to happen together in real-time parsing. Lastly, when the matrix verb is parsed, the whole structure building is completed.

Further evidence of the incrementality of the decision-making process by the help of intonational break comes from the availability of intermediate readings when there is more than two level of embedding as in (9).

(9) *Amwu-ekey-to BREAK Jina-ka pen-ul cwuessta-ko an hayssta-ko BREAK haysse*
Anyone$_{DAT,even}$ BREAK Jina$_{NOM}$ pen$_{ACC}$ gave$_{COMP}$ NEG said$_{COMP}$ BREAK said
Preferred reading: Jina said that she didn't say to anyone that she gave a pen to someone.'

In (9), *amwu-ekey-to*'to anyone' is a negative polarity item, which seeks a negative trigger. Given that structure-building decision in Korean is made locally (=embedded clause reading preferred) or non-locally (=matrix clause reading preferred), such intermediate reading cannot be determined in advance, though Korean listeners can get the intended reading without difficulty compared to matrix reading or the more deeply embedded reading, as long as a paired sequence of intonational breaks occurs, after *amwu-ekey-to* and *hayssta-ko*. With such intonational help, incremental build-up of interpretation is straightforward.

## *3.2. Multiple Long-distance Scrambling*

Now, we can see how the DS parser can anticipate multiple long-distance scrambled readings as in (8b).

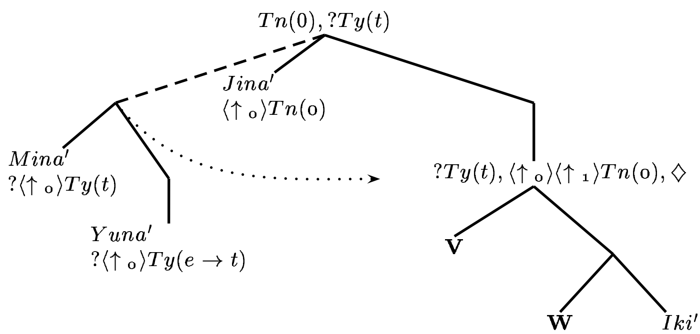



**Fig 6. Parsing *Yuna-rul Mina-ka BREAK Jina-ka in (8b)***

The only difference between two modes of assertion for (8) lies in the location of the intonational break. At the beginning, the parser unfolds the same propositional structure with the same options of *Adjunction and Local *Adjunction. But this time, the former option is taken to feed the latter. That is, from an intermediate but unfixed node requiring a propositional formula (introduced by *Adjunction and decorated with $?Ty(t)$), two successive steps of Local *Adjunction and update will enable the two NPs (*Yuna-rul Mina-ka*, which indicated by the lack of any intonational break to be local to each other, to be construed as decorating two argument nodes (subject and object respectively) in a local structure. This structure however lacks a predicate value. The intonational break immediately follows this local structure and confirms that the partial structure that has been just built is not to be interpreted as local to what follows. Then, what follows is *Jina-ka*. This node can be updated via Local * Adjunction, induced from the root node ($Tn(0)$). The option of * Adjunction is not available anymore as it is already used in building a partial structure of unfixed node for *Yuna-rul Mina-ka*. But Generalised Adjunction can apply, just as before, for which a step of abduction will be needed to fix this very weak relation as one of immediate subordination ($\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)$). With this sequence of actions, all as before, the unfixed node initially introduced can be resolved by merging with a fixed node in the emergent structure. The second break, which confirms that update of the unfixed node will have already taken place, is also observed after *-ko*. But this time the resolution of this unfixed node will take place at the $Ty(t)$ requiring node, and will immediately provide both arguments for the predicate *iki'-*. And then, the parser will unify the unfixed sub-tree with the subordinate structure as the arrow shows in Fig 6. By doing so, the structure projected by *iki-* will be updated so that its object argument is now specified as *Yuna'*, and *Mina'* is identified as its subject.

Notice the naturalness of this analysis as against the lack of any account of incrementality of multiple long-distance dependency in most accounts. The flexibly-ordered sequence of NPs can induce a partial structure building via Local * Adjunction (immediately following a step of *Adjunction) well before reaching any verb, reflecting incrementality of interpretation. Our current argument is close to Phillips (1996) in terms of assuming incremental structure building and pursuing unified system for grammar and processing. However, the crucial difference between his and our framework is that, by adopting an explicitly parsing-based perspective realized in terms of monotonic tree-growth, we do not destroy temporarily built structure during the course of parsing as he argued but just update the initially underspecified partial structure.

## 4. Conclusion

In this paper, we have shown how structure building in Korean not only can, but must, be construed as incremental, in this respect just like English; and this incrementality has been reflected directly by DS analyses involving structural underspecification and update. Local and long-distance scrambling were used as evidence for this parsing-oriented perspective, the different forms of initial structural underspecification which underpin them predicting multiple long-distance scrambling as an immediate consequence.

## References


Baldridge, Jason. 2002. *Lexically Specified Derivational Control in Combinatory Categorial Grammar*. Ph. D. thesis, University of Edinburgh.

Blackburn, Patrick and Wilfried Meyer-Viol. 1994. Linguistics,logic, and finite trees. *Bulletin of Interest Group of Pure and Applied Logics.* Vol 2, 2-39.

Cann, Ronnie, Kempson, Ruth and Lutz Marten. 2005. *Dynamics of Language*. Elsevier.

Fodor, Jerry. 1983. *Modularity of Mind*. Cambridge. MIT Press.

Fodor, Janet Dean and Yuki Hirose. 2003. What Japanese Parsing Tells us about Parsing. *Japanese/Korean Linguistics*. In William McClure (Ed.) Vol 12, 192–205. Stanford, CA: CSLI

Inoue, Atsu and Janet Fodor 1995. Information-paced parsing of Japanese *Japanese Sentence Processing*. In R. Mazuka and N. Nagai (Eds.) 9–64. Hillsdale: Lawrence Erlbaum Associates.

Jun, S. A. 1993. *The Phonetics and Phonology of Korean Prosody,* Ph. D. thesis, Ohio State University.

Kaplan, Ronald and Annie Zaenen. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In *Alternative Conceptions of Phrase Structure* M.Baltin and A.Kroch, (Eds.) 17-42. University of Chicago Press.

Kempson, Ruth., W. Meyer-Viol, and Dov Gabbay. 2001. *Dynamic Syntax: The Flow of Language Understanding*. Blackwell.

Kiaer, Jieun. in prep. How does prosody help parsing Korean?. ms. KCL.

Kim, Yongjin. 1999. The effects of case marking information on Korean sentence processing. *Language and Cognitive Processes* Vol 14. 687–714.

Miyamoto, Edson and Shoich Takahashi. 2001. The processing of wh-phrases and interrogative complementizers in Japanese. *Japanese/Korean Linguistics*, Vol 10, In N. Akatuka and S. Strauss, (Eds.) 62-75. Stanford, CA: CSLI

Mazuka, Reiko and K, Itoh. 1995. Can Japanese Speakers be Led Down the Garden Path? *Japanese Sentence Processing.* In R. Mazuka and N. Nagai (Eds.) 295–329. Hillsdale: Lawrence Erlbaum Associates.

Mazuka, Reiko, K. Itoh and T, Kondo. 2002. Cost of scrambling in Japanese sentence processing. *Sentence Processing in East-Asian Languaegs* In M. Nakayama (Ed.) 131–166. Stanford, CA: CSLI.

Phillips, Colin. 1996. *Order and Constituency,* PhD thesis. MIT.

Pritchett, Bradley. 1992. *Grammatical Competence and Parsing Performance,* Chicago, The University of Chicago Press.

Stowe, Laurie. 1986. Parsing Wh Constructions: Evidence for on-line gap location *Language and Cognitive Process* Vol 1, p. 227–245.

Yamashita, Hiroko. 1997. The effects of word-order and case marking information onj the processing of Japanese. *Journal of Psycholinguistic Research*. Vol 26, 163–188.

# Proceedings of the 24th West Coast Conference on Formal Linguistics

edited by John Alderete, Chung-hye Han, and Alexei Kochetov

Cascadilla Proceedings Project Somerville, MA 2005

## Copyright information



ISBN 1-57473-407-5 library binding



## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

## Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Kiaer, Jieun and Ruth Kempson. 2005. Pro-active Parsing of Korean Scrambling. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. John Alderete et al., 209-217. Somerville, MA: Cascadilla Proceedings Project.

or:

Kiaer, Jieun and Ruth Kempson. 2005. Pro-active Parsing of Korean Scrambling. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. John Alderete et al., 209-217. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1225.