Multiple Foci in Japanese Clefts and the Growth of Semantic Representation

Tohru Seraku[1]

[1] St. Catherine's College, Manor Road, Oxford, UK. OX1 3UJ
tohru.seraku@stcatz.ox.ac.uk

Abstract. There are two types of Japanese clefts, depending on the presence of a case particle attached to a focus; multiple foci are possible only in clefts with particles. This paper proposes that this distinct behavior with respect to multiple foci emerges as a consequence of the incremental growth of semantic structure within Dynamic Syntax; multiple foci are licensed when an underspecification for each item is resolved by each case particle. From this analysis also follows the partially case-specified foci and the so-called clause-mate condition.

Keywords: incrementality, structural underspecification, Dynamic Syntax

1 Introduction

Japanese clefts are divided into two types, depending on whether a focus item has a case particle or not [5], as illustrated in (1). In this paper, clefts with a particle are called "clefts$_{+P}$", and clefts without a particle are called "clefts$_{-P}$".

(1)	[*Tom-ga*	*kat-ta*	*no*]-*wa*	*ringo*(-*o*)	*da*.
	[T.-NOM	buy-PAST	NO]-TOP	apple(-ACC)	COP
	'It is an apple that Tom bought.'				

As observed in [12], multiple foci are possible only in clefts$_{+P}$.

(2)	[e_i e_j	*purezento-o*	*age-ta*	*no*]-*wa*
	[	present-ACC	give-PAST	NOM]-TOP
	Tom$_i$*(-*ga*)	*Mary*$_j$*(-*ni*)	*da*.	
	T.-NOM	M.-DAT	COP	
	Lit. 'It is Tom$_i$ to Mary$_j$ that e_i gave a present to e_j.'			

Within Principles-and-Parameters theory, the distinct behavior of clefts$_{-P}$/clefts$_{+P}$ with respect to multiple foci has been a challenge to a uniform analysis [3, 4, 11, 12, 13, 16]; the tendency is to assume that the particle *no* in clefts$_{+P}$ is a complementizer, while *no* in clefts$_{-P}$ is a pronominal. Kizu [10] is exceptional in presenting a unitary account, but her account is not applicable to multiple foci, as admitted in [10, p.54].

There is also an empirical puzzle that has been overlooked in the literature. Though case particles have been presumed to be obligatory in multiple foci, a second focus, but not a first focus, may be case-less, as shown in the contrast between (3) and (4).

(3) [e_i e_j *purezento-o* *age-ta* *no*]-*wa*
[present-ACC give-PAST NO]-TOP
Tom$_i$-*ga* *Mary*$_j$ *da*.
T.-NOM M. COP
Lit. 'It is Tom$_i$ to Mary$_j$ that e_i gave a present to e_j.'

(4) *[e_i e_j *purezento-o* *age-ta* *no*]-*wa*
[present-ACC give-PAST NO]-TOP
Tom$_i$ *Mary*$_j$-*ni* *da*.
T. M.-DAT COP

There is a passing remark on the partially case-specified foci data in [4], but no analysis is given.

This paper develops a uniform analysis of clefts in connection with multiple foci, from the perspective of the incremental growth of semantic structure, as modeled in Dynamic Syntax. After Section 2 introduces the framework, Section 3 articulates a unified analysis of clefts, and Section 4 examines the issues in multiple foci. Finally, the analysis is extended to long-distance clefts in Section 5, where it is argued that the clause-mate condition [11] naturally follows from the analysis presented.

2 Framework: Dynamic Syntax

Dynamic Syntax is a grammar formalism that models "Knowledge of Language", which is defined as a set of constraints on the building-up of semantic structure [1, 9]. With such constraints, a parser processes a string of words in order, and builds up semantic structure gradually. In this view, a string is directly mapped onto semantic structure; thus, there is no syntactic structure.

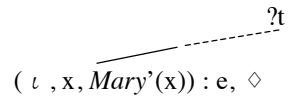
The growth of semantic structure is driven by combination of (a) general action, (b) lexical action, and (c) pragmatic action. (Pragmatic action is not directly relevant in this paper, and thus disregarded.) For illustration, let's consider how the string (5) is progressively mapped onto its semantic structure word-by-word.

(5) *Mary-ga* *nai-ta*.
M.-NOM cry-PAST
'Mary cried.'

The first item to be parsed is *Mary*. Japanese allows the permutation of arguments, and thus a parser cannot see at this point whether *Mary* turns out to be a subject or an object, etc. So, the node for *Mary* is unfixed at this stage. Formally, the general action of LOCAL *ADJUNCTION induces structural underspecification, or an unfixed node, which must be fixed in a local proposition. The lexical action of *Mary* then decorates

this unfixed node. These two actions give rise to the semantic structure (6). In (6), the structural underspecification is notated with the dashed line.

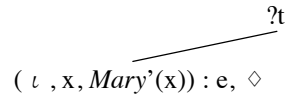
(6) Parsing *Mary*



In (6), the top node is decorated with ?t, a requirement that this node be decorated with a type-t content (i.e. proposition). The unfixed daughter node has three pieces of information. First, (ι , x, *Mary*'(x)) is a content of *Mary*. (ι is an iota operator that binds the variable x, which is restricted by *Mary*'.) Second, e states that the content is of semantic type-e. Third, the pointer ◊ indicates a node under current development.

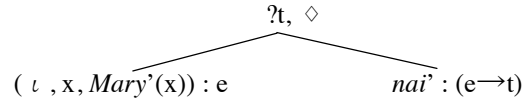
The underspecification in (6) is fixed by the lexical action of the nominative case particle *ga*, which fixes the node for *Mary* as a subject node, as in (7), where the dashed line becomes solid.

(7) Parsing *Mary-ga*



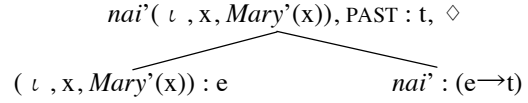
The next item is the predicate *nai*. Japanese being a pro-drop language, a predicate constructs an open proposition with argument slots. Thus, the lexical actions of *nai* updates (7) into (8), where a subject slot has been already identified.

(8) Parsing *Mary-ga nai*



Finally, the general action of functional application composes the content at each daughter node, and the parse of *ta* puts tense information PAST at the root node.[1] The final state is given in (9); this structure is well-formed in the sense that no outstanding requirements ? remain.

(9) Parsing *Mary-ga nai-ta*



This section has introduced the framework of Dynamic Syntax: a parser constructs semantic structure incrementally on the basis of left-to-right word-by-word parsing.

[1] In a fully articulated structure, tense information is represented as an event term in Epsilon Calculus [2].

3 Uniform Analysis of Clefts[2]

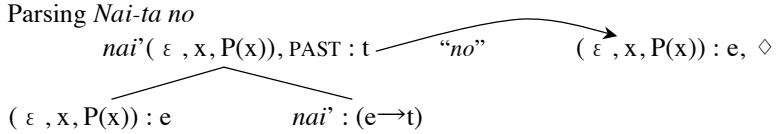
3.1 Clefts$_{+P}$

This sub-section provides a tree-transition for simple cases of clefts$_{+P}$, such as (10).[3]

(10) [*Nai-ta* *no*]-*wa* *gakusee-ga* *san-nin* *da*.
[cry-PAST NO]-TOP student-NOM 3-CL COP
'It is 3 students that cried.'

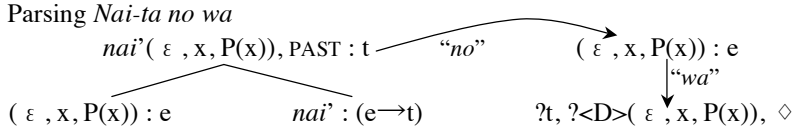
First, the parse of *Nai-ta* yields a proposition; the gap is notated as (ε , x, P(x)), where ε is an existential operator, and P is an abstract restrictor [8, p.65]. Next, following [1, p.285], Seraku [15] regards *no* as a nominalizer that copies a type-e term within a proposition, and pastes it at a newly created node that is LINKed to the proposition. In (11), the LINK relation is notated by the curved arrow, and what is copied/pasted is the type-e term (ε , x, P(x)).

(11) Parsing *Nai-ta no*



What comes next is the topic marker *wa*. The standard analysis of *wa* [1, p.268] is to put at a new type-t-requiring node the requirement ?<D>(ε , x, P(x)), which requires that a node below this node should be decorated with (ε , x, P(x)). (This requirement will be met in (13) below, since the term (ε , x, *gakusee*'(x)) is stronger than the term (ε , x, P(x)).)

(12) Parsing *Nai-ta no wa*



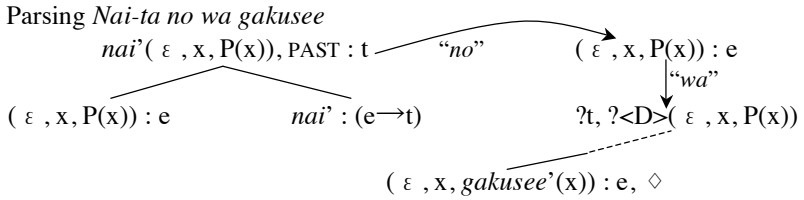
In order to parse the next item *gakusee*, structural underspecification is reuiqred. As mentioned in (6) in Section 2, the general action of LOCAL *ADJUNCTION induces an unfixed node (marked by the dashed line), which allows *gakusee* to be parsed.

---

[2] The theoretical insights of this section are sketched in [15].

[3] To many speakers, if a focus item has a nominative case particle *ga*, the string is not perfectly fine. As noted in [11], the string becomes acceptable if a numeral quantifier, such as *san-nin* in (10), is used. Also, a *ga*-marked focus is acceptable in multiple foci, as in (2) in Section 1. This tendency is stronger in Korean, where clefts with a nominative case-marked focus are completely unacceptable [6]. This paper sets aside the speaker variation concerning the acceptability judgments of case-marked foci in Japanese clefts.

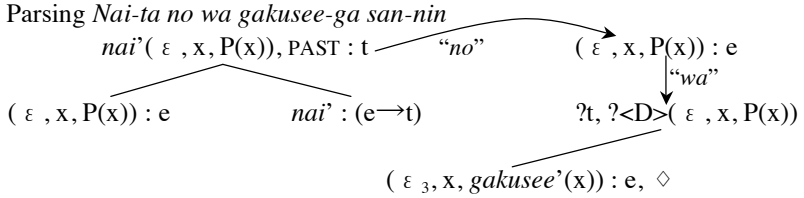
(13) Parsing *Nai-ta no wa gakusee*



So far, the tree-transition is neutral between clefts$_{-P}$ and clefts$_{+P}$.

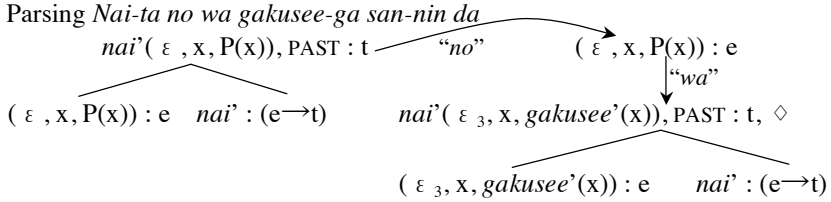
At this stage, the transition starts to differ between clefts$_{-P}$ and clefts$_{+P}$, since only the latter has a case particle. In the cleft$_{+P}$ (10), the case particle *ga* is parsed, which fixes the unfixed node as a subject node, as indicated by the solid line in (14). (In (14), the quantifier *san-nin* has been parsed; note the strengthened operator ε_3.)

(14) Parsing *Nai-ta no wa gakusee-ga san-nin*



Finally, the copula *da* is parsed. Seraku [15] argues that *da* provides a type-t meta-variable, which licenses the re-use of previous representation or actions. In the case of (15), the representation (i.e. the structure built up by the parse of *Nai-ta*) is re-used, where a subject node has been already present due to the parse of *gakusee-ga*.

(15) Parsing *Nai-ta no wa gakusee-ga san-nin da*



This semantic representation is the final state of the tree transition for the cleft$_{+P}$ (10).

3.2 Clefts$_{-P}$

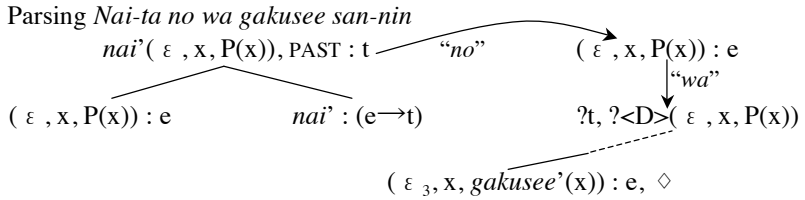
The analysis of clefts$_{+P}$ is straightforwardly applicable to clefts$_{-P}$, like (16).

(16)	[*Nai-ta*	*no*]-*wa*	*gakusee*	*san-nin*	*da*.
	[cry-PAST	NO]-TOP	student	3-CL	COP

'It is 3 students that cried.'

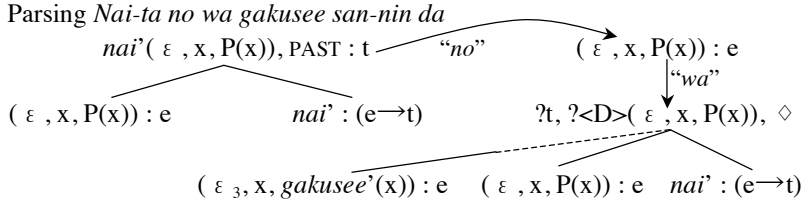
Given the transition above, the parse of *Nai-ta no wa gakusee san-nin* yields (17).

(17) Parsing *Nai-ta no wa gakusee san-nin*



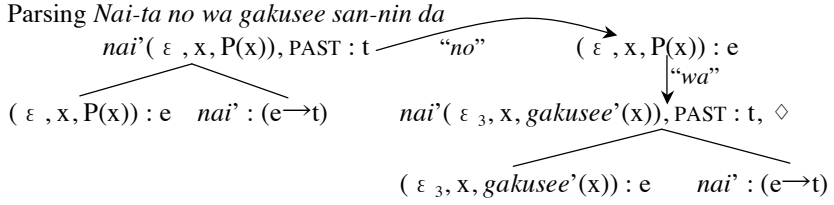
This time, the cleft$_{-P}$ (16) has no particle, and the transition proceeds without fixing the unfixed node. The parse of *da* then licenses the re-use of the previous structure.

(18) Parsing *Nai-ta no wa gakusee san-nin da*



The subject node then unifies with the unfixed node by the general action of MERGE [1], engendering the final state (19).[4]

(19) Parsing *Nai-ta no wa gakusee san-nin da*



This structure of the cleft$_{-P}$ (16) is identical to that of the cleft$_{+P}$ (10). This ensures that the two types of clefts are truth-conditionally equivalent.

In sum, this section has sketched the unitary account of the two types of clefts in Japanese. The next section examines the issues in multiple foci.

4 Multiple Foci

In Section 3, when a focus item was parsed, an unfixed node was introduced as a result of the general action of LOCAL *ADJUNCTION. This general action may be conducted more than once from the same node only if an unfixed node has been fixed before another unfixed node is introduced [1, p.235]: this is because multiple nodes that are unfixed from the same node are indistinguishable, leading to inconsistency. This single-unfixed-node constraint accounts for the difference in the availability of

[4] MERGE outputs (ε 3, x, *gakusee*'(x)&P(x)), but it is simply notated as (ε 3, x, *gakusee*'(x)). The omission of the abstract restrictor P makes no difference if there is a concrete restrictor.

multiple foci between the two types of clefts. First, in clefts$_{+P}$, multiple foci are licensed, since an unfixed node for each focus is fixed by each case particle. Second, in clefts$_{-P}$, the only means of fixing an unfixed node is a unification process. In order for unification to occur, however, a proposition must have been built up, and, as shown in Section 3, such a proposition is created only after the parse of *da*. Thus, more than one item cannot be clefted in clefts$_{-P}$.

Within Dynamic Syntax, there are other two types of structural underspecification: *ADJUNCTION and GENERALIZED ADJUNCTION [1]. A question, then, is: can these be utilized more than once from the same node? The answer is no, since an unfixed node induced by these actions is not fixed by the parse of a particle [1]. Thus, if these types of underspecification are induced, multiple foci are prohibited in clefts$_{+P}$ and clefts$_{-P}$. One may object that this wrongly predicts that both clefts$_{+P}$ and clefts$_{-P}$ do not allow multiple foci. Nevertheless, the availability of multiple foci in clefts$_{+P}$ is adequately modeled, as there is a successful tree transition based on LOCAL *ADJUNCTION.

This analysis also accounts for the partially case-specified foci in (3) in Section 1, repeated here as (20). In this case, the unfixed node for the second focus may be fixed by unification, and thus needs not to be fixed by the parse of a case particle.

(20)	[e_i	e_j	*purezento-o*	*age-ta*	*no*]-*wa*
	[		present-ACC	give-PAST	NO]-TOP
	Tom$_i$-*ga*		*Mary*$_j$	*da*.	
	T.-NOM		M.	COP	
	Lit. 'It is Tom$_i$ to Mary$_j$ that e_i gave a present to e_j.'				

To some speakers, (20) may not be completely fine, but it is much better than (21).

(21)	*[e_i	e_j	*purezento-o*	*age-ta*	*no*]-*wa*
	[		present-ACC	give-PAST	NO]-TOP
	Tom$_i$		*Mary*$_j$-*ni*	*da*.	
	T.		M.-DAT	COP	

In (21), *Tom* has no particle, and the underspecification for *Tom* is not resolved. Thus, another structural underspecification cannot be induced for *Mary*.

As mentioned above, unification requires that a proposition be present. Given that a proposition is constructed by the parse of *da*, unification is applicable only to a final focus. Thus, it is further predicted that a cleft is acceptable where there are three foci and only the final focus lacks a case particle. This prediction is borne out:

(22)	[e_i	e_j	e_k	*age-ta*	*no*]-*wa*
	[			give-PAST	NO]-TOP
	Tom$_i$-*ga*		*Mary*$_j$-*ni*	*purezento*$_k$	*da*.
	T.-NOM		M.-DAT	present	COP
	Lit. 'It is Tom$_i$ to Mary$_j$ a present$_k$ that e_i gave e_k to e_j.'				

Again, (22) may not be perfectly fine, but it is much better than cases in which a case particle is dropped off foci other than the final focus (i.e. either the first or the second

focus, or both). Such cases are all unacceptable, since any unfixed nodes for a non-final focus without a case particle cannot be fixed.

In a nutshell, multiple foci are constrained by the incremental updating of semantic structure that involves structural underspecification and subsequent resolution.

5 Long-distance Clefts

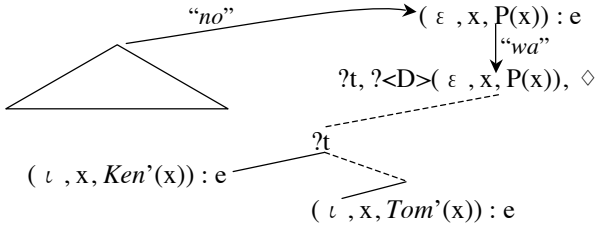
The analysis developed so far can be extended to long-distance clefts, like (23). Based on the treatment of long-distance scrambling proposed in [7], the transition for (23) is as in (24); after *wa* is parsed, a type-t-requiring node is introduced by *ADJUNCTION, and the foci are parsed by LOCAL *ADJUNCTION below this node.[5]

(23)

[*Mary-ga*	[e_i e_j	*purezento-o*	*age-ta*	*to*]
[M.-NOM	[	present-ACC	give-PAST	COMP]
it-ta	*no*]-*wa*	*Ken*$_i$-*ga*	*Tom*$_j$-*ni*	*da*.
say-PAST	NO]-TOP	K.-NOM	T.-DAT	COP

Lit. 'It is Ken$_i$ to Tom$_j$ that Mary said that e_i gave a present to e_j.'

(24) Parsing [*Mary-ga* [*purezento-o age-ta to*] *it-ta no*]-*wa Ken-ga Tom-ni*



The rest of the transition is as usual; *da* licenses the re-use of the representation (i.e. the structure built up by the parse of *Mary-ga purezento-o age-ta to it-ta*), and after functional applications, the complete semantic structure emerges.

Interestingly, the string (23) is ambiguous; the following reading is also possible.

(25)

[e_i e_j	[*Mary-ga*	*purezento-o*	*age-ta*	*to*]
[	[M.-NOM	present-ACC	give-PAST	COMP]
it-ta	*no*]-*wa*	*Ken*$_i$-*ga*	*Tom*$_j$-*ni*	*da*.
say-PAST	NO]-TOP	K.-NOM	T.-DAT	COP

Lit. 'It is Ken$_i$ to Tom$_j$ that e_i said to e_i that Mary gave a present to someone.'

In this case, the foci are parsed without the use of *ADJUNCTION; that is, the foci are parsed by LOCAL *ADJUNCTION directly below the node decorated with ?t, ?<D>(ε , x, P(x)). The parse of *da* then licenses the re-use of the propositional structure, into which the content of each focus is reflected.

In this string, what is not possible is the reading (26).

[5] The node for *Tom* is not fixed at this point to move the pointer ◊ back to the top node [7].

(26) *[e_i [*Mary-ga* e_j *purezento-o* *age-ta* *to*]
[[M.-NOM present-ACC give-PAST COMP]
it-ta *no*]-*wa* *Ken*$_i$-*ga* *Tom*$_j$-*ni* *da*.
say-PAST NO]-TOP K.-NOM T.-DAT COP
Lit. 'It is Ken$_i$ to Tom$_j$ that e_i said that Mary gave a present to e_j.'

This indicates that foci must belong to the same local clause. This restriction has been already observed, and called the "clause-mate" condition [11]. There have been some attempts to account for this restriction, but they are not fully persuasive; Takano [16], one of the most advanced works on this matter, ends up positing the extra mechanism to create "surprising constituents".

In my analysis, however, the clause-mate condition emerges as a by-product of the gradual updating of semantic representation. First, in (23), the foci are parsed below the type-t-requiring node that is created by *ADJUNCTION. Second, in (25), the foci are parsed below the type-t-requiring node that is decorated by the parse of *wa*. Thus, the content of each focus is reflected into the same local proposition, and the reading in (26) is not possible. The upshot is that what is called the clause-mate condition is a result of the progressive development of semantic representation, which renders the condition merely an epiphenomenon.

As a further confirmation, if *kinoo* (= 'yesterday') is inserted between the foci *Ken* and *Tom* in (23)/(25), as in *Ken-ga kinoo Tom-ni*, the expected pattern occurs. In (23), *kinoo* can modify only the most embedded clause; in (25), *kinoo* can modify only the clause whose main predicate is *it* (= 'say'). This is because, in each case, *kinoo* is parsed below the same type-t-requiring node as the foci are parsed.

In closing, let me discuss a potential problem. To prevent a *ga*-marked item from being long-distance scrambled, Cann et al. [1, p.254] suggest that *ga* encodes a constraint that the node for a *ga*-marked item be fixed immediately. If this constraint is maintained, long-distance clefts with a *ga*-marked focus should not be possible; as shown in (23), however, such clefts are possible. Indeed, the long-distance scrambling of *ga*-marked item is acceptable if an embedded predicate is compatible only with the scrambled item [14, pp.176-7]. Thus, the immediate-closure constraint should be relaxed from the lexical entry of *ga*, and the acceptability of (23) is not unexpected.[6]

6 Conclusion

This paper has argued that Japanese clefts are uniformly characterized, and the issues in multiple foci are nicely handled, by the incremental growth of semantic structure that involves "structural underspecification and subsequent resolution". The analysis

[6] In Korean, the long-distance scrambling of a nominative-case-marked item is fine. Kempson and Kiaer [7, p.166] claim that in Korean, the immediate-closure constraint is not encoded, but a default option. My point is that in Japanese, too, the constraint is a default option, though the tendency for the default option is stronger in Japanese. In the case of clefts, the structure built up by the parse of a presupposition clause highlights the content of a gap (ε , x, P(x)), and it signals a parser to override the default option.

should be evaluated in four respects. First, the single entry of *no* as a nominalizer that copies a type-e term leads to the uniform treatment of clefts$_{-P}$/clefts$_{+P}$. Second, the output of the tree growth captures the parallelism of the two types of clefts: truth-conditional equivalence. Third, the process of tree growth captures the difference: availability of multiple foci. Fourth, restrictions on multiple foci (i.e. partially case-specified foci and the clause-mate condition) are explained as a by-product of the tree growth. These four achievements show clear advantages over previous studies within Principles-and-Parameters theory, where, as stated in Section 1, multiple foci have been an obstacle to a unified treatment, and the data such as partially case-specified foci and the clause-mate condition have posed empirical puzzles.

Finally, a general implication for linguistic theory. The issues in multiple foci are normally assumed to be "syntactic" in that they have been analyzed on the basis of syntactic operation/representation. In contrast, this paper has argued that they can be fruitfully dealt with semantically (i.e. in light of the incremental semantic-structure building). This reducibility of syntactic issues to semantic tree growth illustrates the significance of incrementality in linguistic theorizing.

Acknowledgments. I'd like to thank David Cram, Ruth Kempson, and Jieun Kiaer for their valuable comments on the draft of this paper.

References

1. Cann, R., Kempson, R., Marten, L.: The Dynamics of Language. Elsevier, Oxford (2005)
2. Gregoromichelaki, E.: Conditionals in Dynamic Syntax. In: Kempson, R., et al. (eds.) The Dynamics of Lexical Interfaces, pp. 237--278, CSLI, Stanford (2011)
3. Hasegawa, N.: A Copula-based Analysis of Japanese Clefts. In: Inoue, K. (ed.) Grand-in-Aid for COE Research Report (1), pp. 15--38, KUIS, Chiba (1997)
4. Hiraiwa, K., Ishihara, S.: Syntactic Metamorphosis. Syntax (to appear)
5. Hoji, H.: Theories of Anaphora and Aspects of Japanese Syntax. Ms., USC (1990)
6. Kang, B.: Some Peculiarities of Korean *Kes* Cleft Constructions. Studia Linguistica 60, 251--281 (2006)
7. Kempson, R., Kiaer, J.: Multiple Long-distance Scrambling. Journal of Linguistics 46, 127--192 (2010)
8. Kempson, R., Kurosawa, A.: At the Syntax-Pragmatics Interface. In: Hoshi, H. (ed.) The Dynamics of Language Faculty. Kuroshio, Tokyo (2009)
9. Kempson, R., Meyer-Viol, W., Gabbay, D.: Dynamic Syntax. Blackwell, Oxford (2001)
10. Kizu, M.: Cleft Constructions in Japanese Syntax. Palgrave, New York (2005)
11. Koizumi, M.: String Vacuous Overt Verb Raising. Journal of East Asian Linguistics 9, 227--285 (2000)
12. Kuroda, S.-Y.: Tokoro-Clause (Tokoro-Setsu). In: Kuroda, S.-Y. Generative Grammar from the Viewpoint of Japanese. (Nihongo kara Mita Seesee Bunpoo) Iwanami, Tokyo (2005)
13. Kuwabara, K.: Multiple Wh-Phrases in Elliptical Clauses and Some Aspects of Clefts with Multiple Foci. MITWPL 29, 97-116 (1996)
14. Mihara, K.: Tense Interpretation and Syntactic Phenomena. (Jisee Kaisyaku to Toogo Gensyoo) Kuroshio, Tokyo (1992)
15. Seraku, T.: On the Polyfunctionality of Copula Sentences in Japanese. Proceedings of the Sixth Cambridge Postgraduate Conference in Language Research (2011)
16. Takano, Y.: Surprising Constituents. Journal of East Asian Linguistics 11, 243--301 (2002)