Conceptual Basis of One-Cycled Syntax*

Hiroyuki TANAKA

1. Introduction

In this article I will argue that a theory encompassing the idea of "multiple Spell-Out" provides a conceptual basis for what may be called "one-cycled syntax." Section 2 describes the system and the schematic effects derived from it. Section 3 discusses the postulated asymmetry between attractor and attractee. Empirical consequences will be examined in sections 4 and 5.

2. Multiple Spell-Out and the Minimalist Operation

From the "minimalist" perspective (Chomsky (1995)), the computational system of human language syntax (henceforth CHL), is conceived of as a process of derivation (or mapping) from an array of lexical items (called numeration) to PF and LF representations which inform the articulatory-perceptual system and the conceptual-intentional system, respectively. The assumption here is that during the derivation from a numeration N to an LF representation λ, there is a single arbitrary point at which an operation called Spell-Out strips morphology-/phonology-related information away from the syntactic object already formed and sends it to the PF component, which ultimately produces a phonetic representation π of the linguistic expression.

This assumption, however, is not well-grounded and has already been challenged by some researchers (e.g. Uriagereka (1997)). It is conceptually undesirable for the following two reasons at least: (i) Spell-Out is assumed to apply only once in the derivation while others (Merge, Attract etc.) are allowed to apply arbitrarily many times, and (ii) the object it applies to is the whole syntactic object, which can in principle be of any degree of complexity, while other operations affect only two elements and construct a maximally local structural relation (i.e. immediate dominance) from them. Behind
the recognition of the above two problems lies the desideratum that all operations in \( C_{HL} \) should be uniform with respect to the number of application available in a derivation and the range of their input. To restrict the application of an operation to a constant number of times would require that \( C_{HL} \) has a counter, which is generally believed to be false. As for the second point, the other operations in \( C_{HL} \) are typically local, and if Spell-Out is to be fully implemented, it is best understood not as a simplex operation but as a sequence of local and simple operations each of which searches through the tree and sends the relevant information to the PF component. Claiming that Spell-Out applies at a single point in the derivation is tantamount to saying that these simplex operations involved in Spell-Out happen to apply in a cluster, without intervention of any other operation. But this is not a conceptual necessity: We could as well maintain the effect of Spell-Out as it is while having the simplex Spell-Out operations scattered among other types of operations.

A model of \( C_{HL} \) that emerges from such a consideration is one in which all kinds of operations are grouped together as suboperations of a single operation. The operations necessary for \( C_{HL} \) are of three kinds: (i) structure building operation (Merge and Attract), creating immediate dominance relations among nodes, (ii) feature checking operation, licensing the elements occurring in syntax, and (iii) Spell-Out, making features of syntactic elements available for interpretive components. We therefore take these three operations to constitute the sole super-operation which, if applied recursively, ultimately produces the entire derivation. The Operation (capitalized hereafter) can be informally written as follows:

(1) Operation (suboperations applied in the order indicated)
   a. Merge/Attract; create \( \{H(\alpha), \{\alpha, \beta\}\} \) from \( \alpha \) and \( \beta \), where \( \alpha \) is necessarily a member of the set \( \Sigma \) of syntactic objects. (Merge and Attract differ in the choice of \( \beta \), which is determined by the internal algorithm of each operation.)
   b. Check; eliminate a –Interpretable feature \( F \) of \( H(\alpha) \) or \( \beta \), provided that \( F \) is in a checking relation with a matching feature \( F' \).
c. Spell-Out; send the features of $\beta$ to the morphophonological and semantic components.

An important idea encoded in (1a) is that Merge and Attract apply only to the root category at each stage of their application. If we can expect that operations in $C_{HL}$ is considerably simple, then it would be the ideal case that they can only target the members of $\Sigma$, the set of syntactic objects already formed by former applications of them, but not the elements embedded by the members of $\Sigma$. The checking operation (1b) is to be understood in a usual way, though it is an expression of an idea to reduce the notion “checking domain” to mere immediate dominance. This idea has to be worked out more thoroughly in future research, and is not the focus of discussion here. The core proposal of this paper bears on (1c). Among many possible implementations of the “multiple Spell-Out” thesis, I assume the version formulated in (1c): If an element $\beta$ gets embedded by a different projection $\alpha$, $\beta$ is spelled-out to the both interpretive components after it undergoes checking if necessary. It means that information from syntax is not read off from a fully constructed syntactic representation, but derivationally accumulated to the interpretive components every time the Operation is applied. Thus, the whole process of $C_{HL}$ can be seen simply as the sequence of the recursive applications of the Operation. The model can be schematically represented as (2b). Compare it with the “single Spell-Out” model in (2a).

(2)  a. Single Spell-Out Model

\[
N \rightarrow \Sigma_1 \rightarrow \Sigma_2 \rightarrow \ldots \rightarrow \Sigma_k \rightarrow \Sigma_{k+1} \rightarrow \ldots \rightarrow \Sigma_n
\]

\[
\text{Morphophonology} \quad \text{Semantics}
\]

b. Multiple Spell-Out Model

\[
N \rightarrow \Sigma_1 \rightarrow \Sigma_2 \rightarrow \ldots \rightarrow \Sigma_n
\]

\[
\text{Morphophonology} \quad \text{Semantics}
\]
This mechanism causes an interesting effect when combined with the principle of Full Interpretation (FI), requiring that there be no illegitimate element in interface representations. In a “single Spell-Out” model, FI has to apply to the (interim) result of syntactic computation. As Spell-Out is conceived of as a repetitive and cumulative process, it is quite natural to have FI work on every occasion of Spell-Out. For the sake of exposition let us take the following formulation:

(3) Every output of Spell-Out must not contain any unchecked attractor.

Here I use the technical term *attractor* to refer to the feature on the target in a checking relation, as opposed to the corresponding *attractee* attracted into its checking domain.\(^1\) Suppose a–Interpretable formal feature \(F\) of a head \(X\) enters into a checking relation with a corresponding feature \(F'\) of \(W(P)\), in such typical configurations as those in (4). Then we call \(F\) the attractor and \(F'\) the attractee in the checking relation in question.

(4) a. \([x_P WP [x_X X YP]]\) (specifier)
b. \([x_X W X]\) (head-adjointed position)

Returning to (3), it requires that categories spelled out from syntax must have all of their attractors (but not necessarily the attractees) already checked off. If an unchecked –Interpretable attractor remains, we take it that the derivation is canceled at that point. Recall that the Operation (1) ensures that a category is necessarily spelled out when it is embedded by a different category. Then, (5) falls out as a theorem.

(5) An attractor must be checked off before it is embedded.

Following Chomsky (1995: ch.4) we assume that attractors are always –Interpretable, in order to reduce the unwanted complexity in computation. Furthermore, when I say “a feature \(F\) is embedded,” as in (5), I mean precisely that the category \(C\) containing \(F\) is embed-
ded by another category which is not a projection of C.

Now note that (5) does not prevent attractees from being embedded. In fact, attractees of argument DPs, for example, must be embedded in a structure (under a projection of predicates, perhaps) to be thematically interpretable at all.

As a result of (5), it now follows that there is only one "cycle" in the derivation, contra Bures (1993) and Watanabe (1995). In order for a derivation to converge, even weak features have to attract some feature before they are embedded. Thus, what have been considered as overt and covert movement are not distinguishable in terms of timing of application in the derivation. Rather, we reinterpret the distinction as follows: Attraction involving only weak features does not pied-pipe the whole category, while that involving a strong feature does (see various implementations of this kind, especially Bobaljik (1995), Groat and O’Neil (1996) and Collins (1997) among others). What our system requires is that once a category is introduced into a structure, all attractors of it, strong or weak, trigger the application of Attract until all of them get checked off. Once all -Interpretable attractors are eliminated from a head H, Attract does not need to (and cannot, due to Last Resort) apply counter-cyclically to a feature of H after it is embedded.

The “one-cycled syntax” emerging from this proposed system has a broad range of empirical effects. After discussing the controversial status of the notions attractor and attractee, I will sketch out the empirical consequences of the proposal.

3. Asymmetry between Attractors and Attractees

Before we go on to examine the consequences, we have to validate the use of the notions attractor and attractee. Conceptually, the distinction is not necessary because there is no a priori reason to claim that asymmetry exists between the features involved in a checking relation. But empirically, it seems that which role an element plays in the asymmetric relationship is determined by UG. Just as it is universally determined that DP has a Case feature, it is also determined that the Case feature of DP is an attractee (or assignee, in traditional terms), not an attractor (assigner). Thus, we want to pre-
vent the obviously unwanted result that DP checks Case with another DP and v or T checks Case with another v or T.\textsuperscript{2} As Chomsky notes, the asymmetry between attractors and attractees "play[es] a certain role in computation (Chomsky 1995: 278)." Having no principled account at hand of how we can determine whether a given feature is an attractor or an attractee, suffice it here to say that UG prescribes that DP have an attractee Case feature and v and T an attractor Case feature, and likewise for other types of features.

A severer problem, for which I have no convincing solution, is why the difference between attractors and attractees bears on the legitimacy with respect to FI (3), and hence on the tolerability of embedding. It is a mere stipulation that attractees can remain unchecked at the point of interim Spell-Out, though obviously they must be checked off by the end of the derivation. I will obviously leave this question open for future investigation.

Our system departs in some crucial respects from Chomsky’s (1995: ch.4), in which only strong features must avoid embedding. (5) refers only to attractors, leaving attractees exempt from the “avoid embedding" requirement. This “weakening” aspect of the revision allows for the possibility that even a strong feature can be embedded if it is an attractee. The revised system allows a DP argument with a strong feature to be embedded until a functional head with an appropriate feature gets merged with the structure. This is a desirable consequence because, as we will show in section 4, the widely observed phenomenon of “weak pronoun shift” strongly suggests that weak pronouns are possessed with a strong feature.\textsuperscript{3} The “strengthening” aspect of (5) with respect to Chomsky’s system is that it requires that all –Interpretable features, as long as they are attractors, be eliminated before embedded. This leads us to the interesting prediction that no head can check its attractor features in its raised position, which is examined in section 5.

4. Embedding Strong Attractees Allowed

As an instance of the situation in which a strong attractor is embedded, we will consider what may be called “weak pronoun shift" and its theoretical implications. Generally across languages, weak
(i.e. unstressed) pronouns tend to be placed higher in the syntactic structure than do other noun phrases. One of the clearest illustrations is pronominal clitics.\textsuperscript{4)} Take for example the cliticized pronoun in French functioning as the direct object of the verb.

\begin{enumerate}
   Jean loves Marie and Charlotte
   'Jean loves Marie and Charlotte.'
\item[6] b. Jean les aime.
   Jean them loves 'Jean loves them.'
\end{enumerate}

Assuming with Kayne (1975) that this kind of displacement is caused by syntactic movement,\textsuperscript{5)} the problem relevant here is what triggers the movement. If in our framework strong features are the cause of overt movement, we want to say that the clitic has a strong affixal feature that requires it to be attached to an appropriate head.

Faced with this situation, Chomsky's system is too strong in prohibiting the embedding of strong features under a projection of different categories: If a clitic bearing a strong feature is introduced in the structure, it must not be embedded before the feature is checked off, but at the same time it must be embedded by a projection of a predicate before the checking takes place in order to be interpreted as an argument. One might propose that this conflict be resolved by assuming that the host (target) of cliticization has the strong feature to attract the clitic, thereby allowing the clitic to have a weak feature. But this move would obscure the insight that the cause of movement is on the part of the pronoun, and would also fail to capture the locality of clitic movement, as shown in the impossibility of moving an object clitic to a higher, non-restructuring verb in Italian (see Rizzi 1982: ch.1).\textsuperscript{6)}

Our (5), on the other hand, readily allows the clitic to have a strong feature and to be embedded in the structure. (5) states that every attractor must not be embedded. Since the pronominal clitic is an argument DP, its affixal feature is an attractee and therefore it can be embedded. Since it is strong, the movement is overt.\textsuperscript{7)}
5. Embedding Weak Attractors Disallowed

The most remarkable contribution of our one-cycled view of syntactic derivation is that it serves to capture a systematic gap in natural language, which has been left unnoticed so far, as an automatic consequence of (5). (5) entails that every attractor has to accomplish checking before it is embedded, and hence within its own projection. In other words, it is impossible for an attractor to wait until it is embedded by another category and to undergo checking after the head containing it has raised to a higher head. Thus we predict the generalization in (7), which as far as I know is exceptionless crosslinguistically.

(7) Absence of Rendezvous Checking (ARC)
No head raises to have its attractor features checked off.

In general, elements with attractee features (DPs, wh-phrases, etc.) usually move up to the checking domain of the target, and this is allowed in our system, as seen in section 4. What is banned is the movement of the head to check its attractor features against the attractees in the checking domain of an upper head it has adjoined to. In other words, the checking domains for each head must be disjoint from those of the others. No attractor can enter into a checking relation with an attractee "away from home," hence the name of the generalization.

In terms of Case, the ARC (7) states that a DP may raise to get Case from higher head, but no head may raise to assign Case to a higher DP. It is safe to say that this is a correct generalization. The assignment of Case to DP may involve movement of DP into the domain of the Case-assigner (either overtly (8) or covertly (9)), but no cases of Case-assignment have been attested which crucially involve movement of the Case-assigner to the vicinity of the assignee, as evidenced by the lack of sentences like (10).

(8) Hei is believed [ti to be a genius].
(9) There is likely to be a riot around here.
(10) *It is likely [TP himi to [vP ti mention that the earth is flat]].
Suppose that in (10) *him, the logical subject of mention, has moved up to the specifier of the embedded TP to satisfy the EPP requirement. The verb mention can have an accusative Case feature even when it takes a that-clause complement (note its ability to Case-mark the expletive it as in They never mentioned it to the candidate that the job was poorly paid; see the discussion in Authier (1991)). Without the ARC, which is derived from (5), it is possible that the verb moves up in covert syntax to the infinitival T to check its Case feature (an attractor) against that of him (an attractee), leading the derivation to convergence. This can be avoided if the Case feature of mention does not tolerate embedding, as required by the ARC.

One might point out that it would have to be inserted into the Spec,TP instead of moving him (thus generating (11), which would win over (10)) in Chomsky’s system of economy where Merge is preferred to Attract/Move.

(11) *Iti is likely [TP t; to [vP him mention that the earth is flat]].

Then we want this derivation to crash in order for (10) to be relevant to our discussion. A possible solution to this problem may be that the “more economical” derivation does not converge because him and mention cannot enter into a checking relation inside vP due to the ban against checking in θ-positions, nor outside vP, due to the ARC.

This system disallows Case-assigning heads to be stacked up together before their Case features are eliminated. It therefore rejects Chomsky’s mechanism of LF accusative Case feature checking, in which the formal features of the verb and the object enter into a checking relation in the checking domain of T. Without the ARC, it is also unclear how to block the unwanted derivation in which the Case features of nominative-assigning T and accusative-assigning v, for instance, are wrongly associated with the logical object and subject, respectively, when v has overtly adjoined to T (e.g. in French) before finishing accusative Case checking. Consider (12), an active transitive sentence in French.

Considerational Basis of One-Cycled Syntax

(In the sense ‘He is likely to mention that the earth is flat.’)
(12) \[ TP \text{Jean} [T \text{embrasse} T] \text{souvent} [vP \text{t embrasse} \text{Marie}] \]

Jean kisses often Marie

'Jean often kisses Mary.'

In a "standard" model in which Case feature checking of object takes place after overt operations are applied (and hence in the LF cycle, counter-cyclically), nothing prevents Jean from checking accusative Case with the verb, while T attracts the formal features of Marie covertly and check nominative Case with it. In our system v has to attract FF(Marie) before T is merged with vP, so there can be no confusion as to which argument should enter into a Case checking relation with which functional head.

The ARC also serves to disambiguate the possible \(\phi\)-feature checking relations among arguments and functional heads in just the same way as it does for Case feature checking. If T is responsible for subject agreement and v for object agreement, then the ARC excludes the possibility that in such a configuration as in (12), the logical object but not the subject controls subject agreement.\(^9\) Agreement as well as Case is one of the important grammatical functions that determine the status of arguments in the sentence. To have an explicit device for reducing the number of possible agreement relations in an empirically correct way is a virtue that any theory of natural language morphosyntax should achieve.

It is also important to notice that the ARC has a direct implication for the phenomena of Case absorption. If all Case checkers have to undergo checking before embedded, and if Case absorption is determined by the external context of the projection of the Case assigning head (Watanabe (1996)), one needs a sophisticated theory of the interaction between selection and the status of Case features.

In other words, we must at least leave open the possibility that the property of attractors (i.e., whether they are absorbed or not) is affected by what kind of projection they are immediately embedded under. If, as suggested in section 2, the notion \textit{checking domain} reduces to immediate dominance (thus including complements in the checking domain of the head) and Case absorption reduces to feature checking between the embedding head (the follow-up check-
er, in Watanabe’s terminology) and the maximal projection of the Case assigner, then Case absorption finds a natural place within our framework. Furthermore, we predict that the follow-up checker must be the head immediately above the Case assigner; otherwise absorption would never take place. Whether this should be the case is unclear in Watanabe’s framework. (And the prediction should be verified empirically, of course.)

Also important in this connection is the Government Transparency Corollary (GTC) effect of Baker (1988). Baker’s intuition is that head movement affects the assignment of grammatical functions to arguments. This seems at first sight to argue against my claim that Case/agreement is checked within the projection of the head responsible for determining these properties. But what the GTC says is that head movement enables the upper head (Y in (13)) to govern what the lower head (X) used to govern prior to movement.

\[(YP \ Y \ [xp \ldots \ X \ldots \ \alpha \ldots])\]

Thus, the only thing head movement makes available is the Case/φ-feature checking relation between Y and an argument (α) which was in the domain of the lower head. And this checking can be done within YP, in conformity with the ARC. As long as the features of X can be absorbed (as suggested above), there is nothing wrong with this situation.

Let us note finally that the ARC has another important theoretical implication, namely that it imposes an Agr-less system of Case checking. Since a Case assigner has to check its Case within its own projection, the Agr-mediated system of Case checking (Chomsky (1993)), where v and T have to raise to Agr to check Case features, can no longer be maintained. Whereas Chomsky (1995: ch.4) rejects the existence of Agr as an independent syntactic head on a purely conceptual ground, our argument here suggests another way to do so, though in a rather technical fashion.

6. Conclusion

In this article I investigated the consequences that a particular implementation of the multiple Spell-Out thesis brings, when
combined with a version of Full Interpretation sensitive to the attractor/attractee distinction. The resultant system is “one-cycled” in that the Operation targets only the root categories at each stage of the derivation and can never go back to affect the proper subpart of the structure already formed. Each category is spelled out as soon as it is embedded and has its FFs checked, and the principle of FI requires that the output of Spell-Out contain no unchecked attractors. Thus it follows that every attractor has to be checked before embedded. This abstract generalization accounts for the existence of strong attractees in argument DPs that may be embedded, and for the absence of feature checking outside the projection of the head containing the attractor (the ARC).

Although the technical implementation is tentative and the empirical analyses still more so, I would like to place this work as a preliminary sketch of a possibility that recursive applications of a highly simple and restrictive computational process (here temporarily dubbed “the Operation”) ultimately derives an important behavior of $C_{HL}$.

Notes

*This is an extended version of the paper presented at a workshop in the 22nd annual meeting of Kansai Linguistic Society, held at Kyoto University on November 8th, 1997. I would like to thank the participants for their useful comments.

1) I assume with Chomsky (1995: ch.4) and Collins (1997) that all attractors are -Interpretable.

2) I assume that it is $v$ in the split VP hypothesis adopted in Chomsky (1995: ch.4) that assigns the accusative Case in active transitive sentences.

3) Watanabe (1997) independently reaches the conclusion that strong features can move in overt syntax.

4) Other examples include object shift affecting only weak pronouns in languages such as Swedish, and the English Verb-Particle constructions taking pronominal objects, and the impossibility of weak pronouns to induce singular concord in Belfast English (Henry (1995)).

5) See Borer (1986) for an overview of movement vs. non-movement.
Conceptual Basis of One-Cycled Syntax

approaches to clitic pronouns.

6) One might argue that the motivation of the clitic movement in (6b) may be attributed to Kayne's (1994) Linear Correspondence Axiom (LCA), as suggested in Chomsky (1995: 337). It is, however, highly dubious that this is the sole motivation of clitic movement. The relevance of the LCA to clitics is that a clitic, being a non-complex DP under the bare phrase structure theory, violates the LCA if it is generated in a complement position and remains there. If it is generated elsewhere, however, the LCA becomes irrelevant and cannot force the movement. (i) represents a case in which the clitic is generated as the subject of a small clause, which cannot be considered as a complement; but the clitic still has to move.

(i) a. Jean les, croit [t, contents de leur chambre].
   'Jean believes them satisfied with their room.'
   b.*Jean croit [les contents de leur chambre].

7) This analysis lends itself to the claim that the operation Move exists, in the sense that the operation is triggered by the moving element but not by the target. It should be seen, however, whether Move can be formulated so that it might not be an exception to the claim that the suboperations of the Operation can only target the root category.

8) My proposal also runs counter to Miyagawa's (1997) analysis of Japanese clause-internal scrambling, in which objects enter into Case checking with v within the checking domain of T.

9) This is not to deny that the grammatical functions of the subject and object can ever be reversed. Inverse voice studied in Ura (1996) is a clear example of this kind, and there is a way to derive this construction without violating (7).

References


Orlando.


(Assistant)