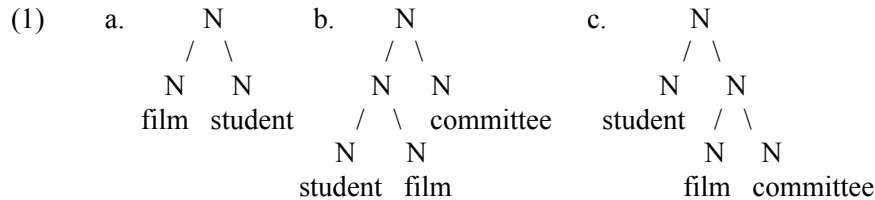# Learnability in a Minimalist Framework:
Root Compounds, Merger, and the Syntax-Morphology Interface

Thomas Roeper, University of Massachusetts
William Snyder, University of Connecticut and Haskins Laboratories
Kazuko Hiramatsu, University of Connecticut

1. Cross-linguistic Variation and the Minimalist Program

The Minimalist Program in syntactic theory seeks to factor cross-linguistic variation out of the syntax, and entirely into other grammatical components, chiefly the lexicon and the interfaces with semantics and phonetics. At the same time, current Minimalist research accords a central role to the operation of "Merge," a generalized transformation combining two autonomous subtrees as daughters of a single node (Chomsky 1995, 1998). We observe that as a consequence, the once peripheral phenomenon of root compounding, as illustrated for English in (1), becomes a critical testing ground for current syntactic theory.

```
(1)     a.     N        b.     N            c.        N
             /  \            /  \                   /  \
            N    N          N    N                 N    N
         film  student    /  \  committee     student  /  \
                         N    N                        N    N
                      student film                   film  committee
```

Following Chomsky (1970), Baker (1988) and Lieber (1992) among others, we take fully productive processes of complex word-formation to occur in the syntax. Productive root compounding then receives a natural analysis as the syntactic merger of two open-class words. Indeed, from this perspective, root compounding is an exceptionally "pure" case of syntactic merger.

As a consequence, however, any cross-linguistic variation in the availability or form of root compounds poses a direct challenge to the Minimalist goal of a universal syntax. This is because the open-class words (e.g. *student*, *film*, *committee*) that are merged in examples such as (1a-c) lack any relevant selectional requirements in their lexical entries. Moreover, root compounding is resistant to the standard Minimalist re-statement of syntactic variation as variation in the features of functional heads. First, there is no independent motivation for the introduction of (phonetically null) functional heads within structures such as (1a-c). Second, overt functional heads such as determiners are typically disallowed within endocentric root compounds, as illustrated in (2).

(2)   a. student (*the) film committee   [= 'student committee on the films']
        b. student film (*the) committee   [= 'the committee on student films']

Hence, from a Minimalist perspective it would be convenient if productive root compounding were a universal of human language.

Yet, the productivity of root compounding is in fact an unequivocal point of cross-linguistic variation. For example, Bauer (1978) provides extensive arguments for a sharp contrast between English and French: In the latter, endocentric root compounding is entirely unproductive, limited to frozen forms and self-conscious coinages. The point is illustrated in (3), where the English compound (3a) has both a lexical (drifted) sense ('undersea diver') and the potential for alternative readings, according to context.

(3)   a. English: *frog man* [= 'undersea diver', or 'man who collects frogs', or 'man resembling a frog', or 'man who sells statues of frogs', etc., ad infinitum]

         b. French: *homme grenouille* (lit. 'man frog') [= 'undersea diver']

Creation of novel compounds to express novel meanings occurs freely and unconsciously in English, as we should expect if root compounding is an operation of the syntactic component. In contrast, even though French has the compound in (3b) as a lexical item, its meaning is frozen, and any attempt to use it in a novel sense would require a deliberate, self-conscious act of coinage.

2. Learnability: The Role of Recursive Structures

From the perspective of learnability, cross-linguistic variation in productivity poses a considerable challenge: When attempting to judge the productivity of compounding, how can a child distinguish adults' innovations from their use of frozen, lexical forms? Frequency information (for compound tokens or types) appears unreliable in a language like French, because of its many lexical compounds (cf. Bauer 1978). Crucially, the child acquiring French becomes an adult for whom root compounding is not simply a quantitatively low-frequency process, but rather a qualitatively unproductive process.

A possible solution is suggested by Namiki's (1994) observation that *recursive* compounds are (nearly or entirely) absent, even as frozen forms, in languages of the French type. Thus, the child could in principle use the presence or absence of recursive compounds in the input to decide whether compounding is productive. As reported in (Snyder 1999), a search of the longitudinal corpora of child-adult interactions for ten children learning English (CHILDES; MacWhinney & Snow 1990) revealed that recursive compounds were well-attested in the early adult input to *every* child. Some representative examples are provided in (26).

(26) a. Christmas tree cookie
      b. peanut butter sandwich
      c. baby doll napkin
      d. nursery school book

Hence, as a potential trigger, recursive compounds meet the basic criterion of reliable availability in the input. Accordingly we propose that recursive compounds are a possible trigger for the

marked setting of the Root Compounding Parameter. We speculate that recursive structures may play a privileged role in language acquisition more generally, but we leave this as a direction for future research.[1]

3. Merger Theory and the Morphological Parameterization Hypothesis

To account for the English/French contrast in availability of productive, endocentric root compounding, we first propose the Morphological Parameterization Hypothesis: Syntactic merger is constrained, at the point of its operation, by conditions of *morphological* well-formedness. These well-formedness conditions, in turn, are a point of parametric variation. The child's acquisition of a Minimal syntax can then be understood as reducing, in large measure, to the setting of morphological parameters.

In particular, we propose that English takes a marked parameter-setting that allows Merge to create objects with the status of complex words (root compounds). Our approach requires two principal assumptions. First, the syntactic component's representation of phrase structure must make a formal distinction between maximal and non-maximal projections. Second, as proposed by Chomsky (1998), the available syntactic operations must include both "pair merger" and "set merger." Following Chomsky, we take set merger as the basis for the syntactic composition of head and complement, and pair merger as the basis for adjunction of moved elements.

On our approach, root compounding is the result of set merger applied to a pair of non-maximal projections. We assume that languages universally permit set merger of a head and a *maximal* projection, which serves as the basis for the head-complement relation. Similarly, we assume that languages universally permit *pair* merger of non-maximal projections, which is the basis for head-to-head adjunction in typical cases of head movement. Where languages vary, however, is in the availability of set merger between two non-maximal projections, as stated in (4).

(4)   The Root Compounding Parameter (RCP):
          Set-merger can(not) combine non-maximal projections.

The RCP, in our view, is a morphological parameter, in the sense that the objects resulting from set-merger of non-maximal projections are interpreted as complex words (endocentric root compounds) at the semantic and phonological interfaces. The positive ("can") setting of the parameter makes root compounding a fully productive mode of complex word-formation.

4. Root Compounds, Antisymmetry, and the Abstract Clitic Hypothesis

In this section we relate the RCP to the Abstract Clitic Hypothesis of (Keyser & Roeper 1992), and to considerations of antisymmetry in morphology (Roeper 1999). First, notice that the concept of parameter carries a deliberate ambiguity. The term covers both *formal* and

---

[1] Leftward recursion may engage the grammar in another abstract way which is relevant but goes beyond the scope of this essay. We find that both English and German allow possessives to the left, but only in English is it both productive (all nouns allow it while in German only proper nouns allow it) and recursive: *my sister's cousin's hat.*

substantive connections, within a decision tree for the specification of a grammar, that enable acquisition to be efficient (Chomsky 1965).   The RCP originally emerged from empirical work (Snyder 1995, 1996, 1999; Snyder & Chen 1997; Slabakova 1999; Miyoshi 1999; Sugisaki & Isobe, to appear) as a substantive parameter: The original empirical observation was (roughly) that if verb-particle constructions are attested in a given language, then recursive compounds are also fully productive in that language.

While irreducible substantive connections are biologically possible, one hopes to find a simple formal principle that underlies them.[2] An analogy may be helpful: At the substantive level of description we note that facial hair and sexual interest may emerge together in the male human. In the absence of an obvious logical connection, one expects to find an explanation at a deeper, biological level of analysis.

The challenge for syntactic theory, in our view, is to discover a formal connection between the complexity of rightward-building verb-particle constructions, on the one hand, and the complexity of (typically) leftward-building compounds, on the other, that expresses this parameter with maximal simplicity. While all languages allow some form of object complements and some form of incorporation, only a subset of languages allow particles. What is the connection, then, between righthand particles and leftward compounds?

A possible answer comes from the Abstract Clitic Hypothesis (Keyser and Roeper 1992), which posits a special position for particles. According to the ACH, this position is a simple verb projection, but with one crucial property: Only non-maximal projections can appear in it. In current terms, it is an instance of set merger applied to a head and a complement head. Now we can understand the effects of the RCP in the manner illustrated in (5,6):

(5)   a. [- Root Compound Parameter] (e.g. French)
        b. G1: V - Maximal Projection
        c. partir de   l' école / * partir   d' école
            leave from the school / * leave from school

(6)   a. [+ Root Compound Parameter] (e.g. English)
        b. G2: V - Maximal Projection, or V - Head
        c. leave the school / leave school

Languages with the negative setting of the RCP, such as French, disallow the abstract clitic position, and in general permit only maximal projections as complements to V (cf. 5c). Languages such as English, which are [+RCP], permit the abstract clitic position, and therefore permit both maximal and non-maximal complements to V (cf. 6c).[3]   Hale and Keyser (1993)

---

[2] See (Greenberg 1966) and subsequent work for extensive cataloguing of grammatically-linked phenomena for which the facts are clear and the links obscure.

[3] Note that (apparent) bare-N complements occur in certain frozen expressions in French, especially with the verbs *faire* 'make' and *avoir* 'have', as in *faire (*du) mal* 'cause (*some) pain' and *avoir (*beaucoup de) faim* 'have (*much) hunger'. In the

(see also Ferguson 199_**) add the insight that the clitic position is a form of "inner complement," which can saturate a thematic role in the same manner as a maximal projection.

Turning to considerations of antisymmetry, recall that Kayne (1994) and Chomsky (1995) have advanced the proposals in (7).

(7)   a. All movement rules operate to the left.
        b. Leftward movement is universal.
        c. All structures are represented in the form Specifier-Head-Complement.

Under these proposals, the child has nothing to learn about movement. Let us now consider the Abstract Clicit Hypothesis from this perspective.

The projection of an abstract clitic position, created by set-merger of heads, correctly predicts that the syntactic category of clitics can vary, and that clitics normally stand in complementary distribution, as illustrated in (8,9).[4]

(8)   VP
             /   \
            V   Prt
            | |
          pick   up
          play   dumb
          play   chess
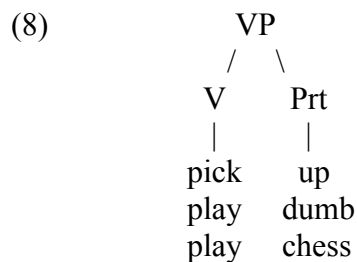
 (9)   a. V-P   pick up
       b. V-P   lose out   *pick up out
       c. V-N   lose faith   *lose faith out
       d. V-A   play dumb   *play dumb up

We also expect clitics to be eligible for leftward movement, and indeed each type of clitic in (9) permits leftward incorporation, as illustrated in (10).

(10)   a. lose faith => faith-losing
        b. stand out => outstanding
        c. play dumb => dumb-acting

---

absence of any evidence for recursion, however, these expressions more plausibly result from lexical selection of a bare NP, rather than from syntactic compound formation.

[4]The indirect object pronoun also occupies the ACP, which accounts for the ungrammaticality of (ia-c). Example (ic) contrasts with the grammatical (ii), as expected.

(i)   a. *He played me dumb.
        b. *He gave me the money up.
        c. *He walked me out. [In the sense of (ii)]

(ii)   He walked out on me.

(11) The disease broke out. => the outbreak of disease

(12)   a. outbreak
         b. outcast
         c. input
         d. overview

Certain derived nominals show the same characteristic of leftward movement, as illustrated in (11,12). (See Roeper 1999 for extensive discussion of the distinction between *outbreak* and *breakout*.)

   Our system is readily extended to prefixes like *re-*, if such prefixes belong to the particle class, originate in the clitic position, and undergo leftward movement, as illustrated in (13).

(13)   *play re-* => *replay*

Moreover, as in noun compounds, there is considerable potential for leftward recursion (14) (although there do exist restrictions). This leftward recursion contrasts sharply with the complementarity effects observed for rightward clitics in (9). The recursive forms can appear with root nominalizations, as in (15), and we also observe interactions of *re-* with leftward-moved prepositional particles, as in (16).
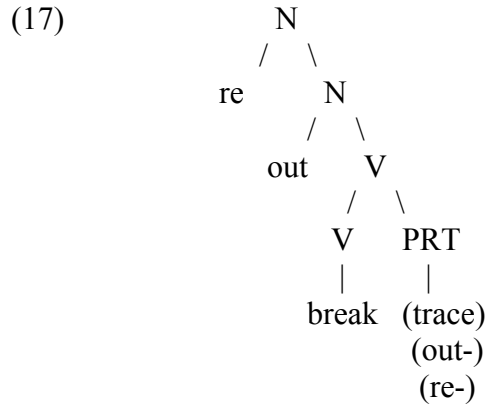
(14)   re-over-reimburse

(15)   the re-outbreak of measles

(16)   a. over-repay
         b. re-overpay

   We propose that a form such as *re-outbreak*, in (15), results from *repeated* introduction of heads into the abstract clitic position, with leftward movement and adjunction. In other words, we propose that *break* is first merged with *out*, and then *out* moves leftward, vacating the abstract clitic position. We assume that word-internal traces are permitted to delete, and that the abstract clitic position can thereby become available for insertion of another element. The prefix *re-* is then inserted into the abstract clitic position, moved leftward, and adjoined to a position higher than *out*. This process is illustrated in (17).[5]

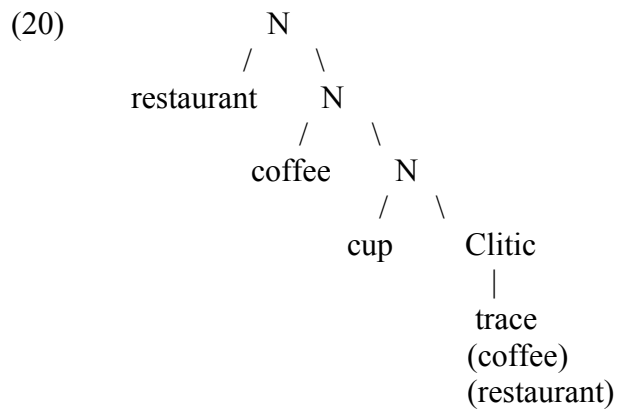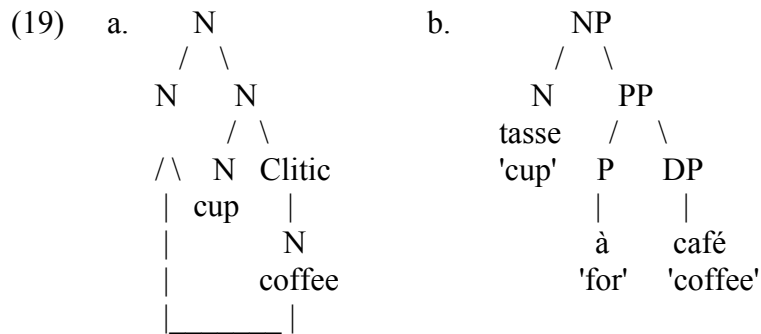[5] We require the further assumption that material introduced through the abstract clitic position and moved leftward would not have been eligible for simple base-adjunction (pair-merger) in its surface position. This effect might derive from considerations of antisymmetry, if some record of the derivational history permits the moved element, but not its base-generated counterpart, to be identified correctly as an adjunct.

(17)
```
                    N
                  /   \
              re        N
                      /   \
                  out       V
                          /   \
                      V       PRT
                      |         |
                    break    (trace)
                             (out-)
                             (re-)
```

The goal now is to provide a structure which gives the same derivation for (18a) and (18b). This requires one further step: the generation of a clitic position for nouns within the same model.

(18)　　a. re　　　out　break
　　　　　b. restaurant coffee cup

Adapting proposals of (Roeper 1999), we arrive at the analysis illustrated in (19a), in which the English compound *coffee cup* is initially a close counterpart to the French phrasal expression, *tasse à café* (lit.) 'cup for coffee' (19b). Re-using the clitic position to introduce a second compound-internal modifier, we obtain *restaurant coffee cup* as in (20).

(19)　　a.
```
          N                      b.            NP
        /   \                               /   \
      N       N                           N       PP
            /   \                       tasse   /   \
         /\   N   Clitic               'cup'   P     DP
         |  cup    |                           |      |
         |         N                           à     café
         |       coffee                      'for'  'coffee'
         |_____|
```

(20)
```
                  N
                /     \
        restaurant      N
                      /    \
                  coffee      N
                            /   \
                        cup       Clitic
                                    |
                                  trace
                                  (coffee)
                                  (restaurant)
```

In (19a) the expression *coffee cup* originates as [$_N$ *cup* [$_{cl}$ *coffee*]], with *coffee* in the clitic position, a structural possibility expected if N takes rightward complements. The initial structure suggests a possible interpretive similarity between a compound-internal modifier, on the one hand, and the genitival/oblique nominal modifiers permitted in a language such as Russian (cf. Partee 1999).[6]

The choice of structure then has consequences both for the input and the output of movement. Leftward movement of the maximal projection (MP) is a form of topicalization (21a), while leftward movement of the head is incorporation (21b). Note that the opposite forms (22a,b) are disallowed.

(21)   a. the house I love
        b. I am house-loving.

(22)   a. *house I love
        b *I am the-house-loving

In order to capture these facts we allude to a general principle of homogenity, stated roughly in (23).

(23)   MP-movement rules build MP structures; head-movement rules attach to head structures.

We will not articulate this consequence further, although we note that it may have deeper significance, and conforms to the concept of head-to-head movement that has been widely discussed.

The system requires recursion, which leads to further predictions. In particular, one should be able to insert non-maximal phrases into the clitic position. Thus we find forms such as (24), where the compound *student film* is first created by inserting *student* into the clitic position of *film* (with immediate leftward movement), and where the result is then inserted into the clitic position of *committee*.

(24)   [student film] committee

Insertion of a maximal projection into the clitic position is excluded, however, with the result that the forms in (25) are ungrammatical.

(25)   a. *restaurant [good coffee] cup
        b. *elephant [the hide] wallet

---

6   The range of interpretations available to compound-internal modifiers is examined from both cross-linguistic and acquisitional perspectives in (Hiramatsu et al. 2000).

Furthermore, since our argumentation rests on universal distinctions, we correctly predict that once the RCP is set, the child language data will be free of exceptions (cf. the ungrammatical forms in 9, 25). Indeed, no such exception has ever been reported.

References

Baker, M. (1988) *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

Bauer, L. (1978) *The Grammar of Nominal Compounding, With Special Reference to Danish, English, and French.* Odense: Odense University Press.

Chomsky, N. (1970) Remarks on nominalization. In R. Jacobs and P. Rosenbaum (eds.) *Readings in English Transformational Grammar.* Waltham, MA: Blaisdell.

Chomsky, N. (1995) *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, N. (1998) Minimalist inquiries: The framework. Ms., MIT, Cambridge, MA.

Ferguson, S. (199_) ***.

Greenberg, J., ed. (1966) *Universals of Language.* Revised second edition. Cambridge, MA: MIT Press.

Hale, K. and Keyser, S. J. (1993) On argument structure and the lexical expression of syntactic relations. In K. Hale and S.J. Keyser (eds.) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger.* Cambridge, MA: MIT Press.

Hiramatsu, K., Snyder, W., and Roeper, T. (2000) Of musical hand chairs and linguistic swing. In *Proceedings of BUCLD 24.* Somerville, MA: Cascadilla Press.

Keyser, S.J. and Roeper, T. (1992) Re: The Abstract Clitic Hypothesis. *Linguistic Inquiry* 23:89-125.

Lieber, R. (1992) *Deconstructing Morphology: Word Formation in Syntactic Theory.* Chicago: University of Chicago Press.

Miyoshi, N. (1999) Compounds and complex predicates: Japanese evidence for a "global" parameter. In *Proceedings of BUCLD 23.* Somerville, MA: Cascadilla Press.

Namiki, T. (1994) Subheads of compounds. [Bibliographic data unavailable]

Partee, B. (1999) Genitives, type-shifting, and argument-modifier debates. Paper presented at the Eastern States Conference on Linguistics, University of Connecticut, Storrs, 20 November 1999.

Rizzi, L. (1989) On the format of parameters. Commentary on Lightfoot. *Behavioral and Brain Sciences* 12:355-356.

Roeper, T. (1999) Leftward movement in morphology. Ms., University of Massachusetts, Amherst.

Slabakova, R. (1999) The complex predicate / N-N compounding relation in L2 acquisition. In A. Greenhill, H. Littlefield, and C. Tano (eds.) *Proceedings of the Boston University Conference on Language Development 23.* Somerville, MA: Cascadilla Press.

Snyder, W. (1995) *Language Acquisition and Language Variation: The Role of Morphology.* Doctoral dissertation, MIT.

Snyder, W. (1996) The acquisitional role of the syntax-morphology interface: Morphological compounds and syntactic complex predicates. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, and A. Zukowski (eds.) *Proceedings of the Boston University Conference on Language Development 20.* Somerville, MA: Cascadilla Press.

Snyder, W. (1999) Compounds and complex predicates in English, Basque, and Romance. Poster presented at the meeting of the International Association for the Study of Child Language, San Sebastian, Spain, 16 July 1999.

Snyder, W. and Chen, D. (1997) The syntax-morphology interface in the acquisition of French and English. In K. Kusumoto (ed.) *Proceedings of NELS 27.* Amherst, MA: GLSA.

Sugisaki, K. and Isobe, M. (to appear) Resultatives result from the Compounding Parameter: On the acquisitional correslation between resultatives and N-N compounds in Japanese. In *Proceedings of the West Coast Conference on Linguistics 19.*