Two Theory-Internal Contradictions in Generative Grammar[*]
Mayumi Hosono

1.   Introduction

Since Chomsky (1959) refuted Skinner's (1957) account of language acquisition LA in terms of verbal behavior, it has long been argued that the faculty that enables a human child to acquire language is innately endowed with human beings. It is the consensus of almost all researchers that language is a genetically innate property of human beings. Generative linguists have addressed the question to what extent this genetic endowment is species-specific. Chomsky (1965) claimed that beyond correctly describing a native speaker's linguistic competence, i.e. 'descriptive adequacy', linguists should seek an 'explanatorily adequate' theory that accounts for the language faculty which LA of a child is based on. The following three observations on LA were focused. (i) In the same way as an adult, a child understands and produces new sentences that she has not heard before. (ii) A child acquires language in a remarkably short period. (iii) LA is uniform in that any human child can acquire the language(s) in her community.[1]

In Chomsky (1981, 1986), the focus of the LA studies was transferred to the question how a human child comes to acquire so rich and complex knowledge of language only with limited data ('Plato's problem', or the problem of the 'poverty of stimulus', Chomsky 1986:7). It was argued that the linguistic theory must be rich enough to achieve descriptive adequacy so that it can apply to any natural languages. It must not be so rich to achieve explanatory adequacy, however: the LA device endowed with human beings must be restricted to enable a human child to precisely determine a few languages given limited data. Thus, it was claimed that the linguistic theory should consist of a finite set of the principles of *Universal Grammar* UG, with the difference between languages derived from different parameter settings and the parametric difference in the lexicon (the *Principles and Parameters* P&P theory).[2]

Since the *Minimalist Program* MP (Chomsky 1995), it has been claimed that the language faculty must meet the requirement imposed by the conceptual system and the sensory-motor system. In the subsequent works (Chomsky 2000, 2001, 2004, 2008), it has been assumed that not only the semantic component is uniform, but also the syntactic computation proceeds uniformly for all human languages. As we will see in detail, this assumption leads to the claim that the parametric difference does not exist in the syntactic component.

---

[*] Many thanks to Anders Holmberg and Halldór Á. Sigurðsson for their interest in and comments on this paper. Any errors are my own.

[1] The LA theories were formulated as the learnability theory in this period (e.g. Culicover and Wexler 1977, Wexler and Culicover 1980). The LA process was once assumed to be instantaneous. This is now widely rejected, as represented by e.g. Hyams'(1983/1986) work that I introduce later.

[2] The literature has addressed the question whether the initial state of the language faculty changes as a child grows up or it does not change from child grammar to adult grammar. See, e.g. Pinker (1984) for the continuity hypothesis and Borer and Wexler (1987) for the maturational hypothesis.

In this paper, I point out two theory-internal contradictions that arise between the LA theory and the currently assumed theoretical framework in generative grammar. The organization is as follows. Section 2 introduces the theoretical transition, from the P&P theory (Chomsky 1981, 1986), through the MP (Chomsky 1995), to the currently assumed framework (Chomsky 2000, 2001, 2004, 2008). Section 3 points out that the assumption that the parametric difference lies in the lexicon causes the LA paradox based on Sigurðsson (2004). Section 4 points out that the LA theory is incompatible with the current theoretical framework: the former assumes that the syntactic computation can differ between languages, whereas the latter does not accept that claim under the assumption of the uniformity in syntax. Section 5 briefly concludes this paper.

2.   The theoretical transition in generative grammar

The P&P theory (Chomsky 1981, 1986) assumed that the linguistic theory consists of a finite set of the UG principles, with the difference between languages derived from the parametric difference in the lexicon as well as different parameter settings. It was argued that explanatory adequacy is achieved by the restricted character of the UG principles, whereas descriptive adequacy is achieved by fixing different parametric values that can produce far-reaching different properties among languages.

The assumption that the parametric difference lies in the lexicon dates from Borer (1984), who argues that the parametric difference is reduced to functional categories only. Many syntactic differences between languages are observed, though the logical meaning of sentences is not affected. Lebanese Arabic and Hebrew are taken as an illustration. In both languages, the object of the verb meaning *talk* is assigned Case through a preposition (1-2a). When an object clitic appears, the preposition absorbs Case (1-2b). When Case is absorbed, Lebanese Arabic allows an object to appear by inserting another preposition, *la* (*clitic doubling*) (1c). Such an option is not allowed in Hebrew (2c).

(1) a. hkit   ma9 Karim.   [Leb. Ara.]
         talk-PAST.1sg with Karim
         'I talked with Karim.'

     b. hkit   ma9-o.
         talk-PAST.1sg with-him
         'I talked with him.'

     c. hkit   ma9-o   la Karim.
         talk-PAST.1sg with-him to Karim
         'I talked with Karim.'

(2) a. dibarti   'im Anna.   [Heb.]
         talk-PAST.1sg with Anna
         'I talked with Anna.'

    b.   dibarti   'im-a.
        talk-PAST.1sg with-her
        'I talked with her.'

    c.   *dibarti   'im-a   (le/šel) Anna.
        talk-PAST.1sg with-her (to/of) Anna
        (Borer 1984:27,(34-35))

The fact above is accounted for in the way that the inflectional rule 'Ø → *la* [PP … NP]' (Borer 1984:28,(37)) is included in the grammatical system of Lebanese Arabic but not included in that of Hebrew. Thus, the presence or absence of inflectional rules and various applications of them to functional categories produce syntactic differences among languages, without affecting the change in the meaning of sentences. It has then been argued that LA is reduced to the acquisition of inflectional rules as well as that of idiosyncratic properties in each language, without the need to learn the UG principles.

        The parameter setting is accounted for as follows. Take *wh*-questions as an example. The *wh*-phrase moves to sentence-initial position in English (3a) but remains in the original position in Japanese (3b).

(3)   a.   What did you eat?

        b.   Kimi-wa nani-o   tabe-ta-no?   [Jap.]
            you-TOP what-ACC eat-past-Q
            'What did you eat?'

On the assumption that the *wh*-phrase is an operator that binds a variable (represented by a trace *t* below), the difference between these languages was accounted for as follows. The *wh*-phrase moves to the operator position, [Spec,CP], 'overtly' at S-structure, in English (4a). It remains there at L(ogical)F(orm) too (4b). In Japanese, on the other hand, the *wh*-phrase does not move at S-structure (5a). It moves to [Spec,CP] 'covertly' at LF (5b). According to this account, the syntactic difference in *wh*-questions is derived from the different setting of the *wh*-movement parameter, i.e. either *wh*-movement at S-structure (English) or *wh*-movement at LF (Japanese). The semantic universality is ensured by the same representations at LF as illustrated in (4-5b).[3]

(4)   a.   [CP what … [IP … [VP … t_{what}]]]   (S-structure)
        b.   [CP what … [IP … [VP … t_{what}]]]   (LF)

(5)   a.   [CP … [IP … [VP … nani-o]]]   (S-structure)
        b.   [CP nani-o … [IP … [VP … $t_{nani\text{-}o}$]]]   (LF)

The acquisition of *wh*-movement has then been reduced to the process in which a child comes

---

[3] See, e.g. Huang (1982) for a representative work on *wh*-movement in the P&P period.

to fix the parametric value of the target language, either '*wh*-movement at S-structure' (for English-speaking children) or '*wh*-movement at LF' (for Japanese-speaking children), by being exposed to primary linguistic data.

Since Chomsky (1995), it has been argued that the language faculty must satisfy the *interface conditions*, which are imposed by the conceptual-intentional system and the articulatory-perceptual system. Only theoretical tools and representational levels that are required to meet the demands of the interface should be assumed. Thus, different syntactic levels such as S-structure and LF have been eliminated from the syntactic model.[4] In the subsequent works (Chomsky 2000, 2001, 2004, 2008), it has been assumed that not only the semantic component but also the syntactic component are uniform for all languages, with the surface difference between languages restricted to the phonological component, as well as the parametric difference in the lexicon.

The assumption of the uniformity in syntax and semantics is ensured by the cartographic system (Rizzi 1997, Cinque 1999). In this system, the position an argument occupies in the syntactic component must correspond to the interpretation it receives in the semantic component. For instance, the argument located in [Spec,FocP] in syntax must be interpreted as the focus in the semantic component. Conversely, the argument that receives the focus interpretation in the semantic component must be located in [Spec,FocP] in syntax.

In the current framework introduced above, *wh*-movement is accounted for as follows. The syntactic derivation proceeds in the same manner and the *wh*-phrase cross-linguistically moves to [Spec,CP] in syntax in all languages, as long as the interpretation as a constituent *wh*-question does not differ. The difference only lies in which copy in the *wh*-chain, either the highest one (<*what,~~what~~*> (6a)) or the lowest one (<~~nani-o~~,nani-o> (6b)), is assigned phonological features in the phonological component (cf. Groat and O'Neil 1996, Pesetsky 2000).[5]

(6) a. [$_{CP}$ what … [$_{TP}$ … [$_{v*P}$ … ~~what~~]]]   (=1a)
        wh-chain: <what,~~what~~>

     b. [$_{CP}$ ~~nani-o~~ … [$_{TP}$ … [$_{v*P}$ … nani-o]]]   (=1b)
        wh-chain: <~~nani-o~~,nani-o>

In this way, the computational system of human language now works in the same manner for all speakers of all languages. Therefore, the parametric difference does not lie in syntax in the current framework, contrary to the P&P period, in which the difference between languages was derived from different parameter settings at the lexical and syntactic levels.

---

[4] It was also claimed that traces are problematic, since they are created in the course of syntactic derivations. It has been assumed that an argument moves, leaving a *copy*, the property of which does not differ from that of the moved argument.

[5] See Watanabe (1992) for a seminal work of *wh*-movement in the transitional period from the P&P theory to the MP.

3.   The language acquisition paradox

Chomsky (2001) accounts for the assumption that the parametric difference lies in the lexicon as follows. The faculty of language specifies a set of universal features {F}. An individual language selects a subset of features [F] from {F} and makes a lexicon LEX. Features in {F} that were not selected are discarded. Lexical items are then selected from LEX and enter a syntactic operation (Chomsky 2001:10).

        Sigurðsson (2004) argues against the assumption of feature selection. Suppose that UG specifies a set of universal features $\{F_1, F_2, F_3, F_4\}$. Suppose that language A selects a subset consisting of $\{F_1, F_2, F_3\}$ and makes a lexicon, whereas language B selects a subset consisting of $\{F_1, F_2, F_4\}$ and makes a lexicon. Since features that were not selected are abandoned, language A could not access F_4. Language B could not access F_3, either. However, the concept of definiteness exists in languages that do not have articles (e.g. Japanese). The concept of future tense exists in languages that do not have the inflection of future tense (e.g. English). Hence, the universal cognitive faculty of human beings requires that all universal features should be available to any languages.

        Sigurðsson's argument points out the LA paradox. Assume above that F_3 is the property of future tense and F_4 that of definiteness. Assume that a Japanese-speaking child learns and acquires the lexicon made by a subset consisting of $\{F_1, F_2, F_3\}$, whereas an English-speaking child learns and acquires the lexicon made by a subset consisting of $\{F_1, F_2, F_4\}$. F_4 that is not contained in $\{F_1, F_2, F_3\}$ could not be accessed by the former. F_3 that is not contained in $\{F_1, F_2, F_4\}$ could not be accessed by the latter. The Japanese-speaking child (and grown-up adult) then could not come to know the concept of definiteness attributed to F_4. The English-speaking child (and grown-up adult) could not come to know the concept of future tense attributed to F_3, either. Both are contrary to fact.

        This paradox cannot be saved even if the parametric variation is assumed to lie in syntax, as in the P&P period: lexicons of individual languages themselves differ, being made by different features. It might be argued that interpretation such as present, past and future is a matter of value assignment in *Agree*, the syntactic operation in which uninterpretable features are assigned values by interpretable features (Chomsky 2000). This does not save the problem here. It is assumed that uninterpretable features enter a syntactic operation without values, unlike interpretable features. As the tense feature T is assumed to be interpretable, T cannot enter a syntactic operation without values and be assigned, e.g. the future value, in Agree. It might be also argued that interpretation, e.g. of future tense and definiteness, is attributed to the interface conditions. It is entirely unclear, however, whether and how a speaker who would have grown up without accessing the concept, e.g. of definiteness, and would not have that concept can interpret it in the conceptual-intentional interface.

        In general, the argument that the parametric difference lies in the lexicon holds only for languages that give some phonetic form to a target linguistic property, regardless of what phonetic form the linguistic property is given in relevant languages. In section 2, we saw the argument based on Borer (1984) that LA is devoted to learning the rules on functional categories, as well as idiosyncratic properties, in the language that a child is exposed to.

Crucially, however, the relevant languages there, Lebanese Arabic and Hebrew, both lexicalize the same grammatical item, i.e. a preposition. The difference between these languages is accounted for in terms of the different behaviors of that grammatical item.

To maintain a theoretical coherence, the parameter setting would have to entirely belong to the phonological component, i.e., be reduced to a matter of whether a relevant linguistic feature is assigned phonological features in individual languages. Syntactic operations would then proceed only with features for all languages, as correctly suggested by a series of Sigurðsson's (e.g. 2006a,b) works. This argument would further lead to the claim that a syntactic operation cannot and must not start with lexical items selected from lexicon, which issue I leave for future discussions.

In sum, the assumption that the parametric difference between languages lies in the lexicon causes the LA paradox based on Sigurðsson (2004): a child who acquires a lexicon that consists of a subset of universal features cannot access discarded universal features. The child and grown-up adult could not come to know the concepts attributed to those abandoned universal features, contrary to fact.

4.   Incompatibility of the language acquisition theory with the current framework

The account of LA in terms of the parameter setting has been widely accepted since Hyams' (1983/1986) work. Adult Italian speakers can omit an overt subject (7a), whereas adult English speakers do not allow a subject position to be null (7b). However, English-speaking children often drop subjects (8a) in the same way as Italian-speaking children (8b).

(7)   a.   Telefonerà.   [Ita.]
           telephone-FUT.3sg
           '(he/she) will telephone.'

      b.   *(he, she,…) will telephone.

(8)   a.   Play it
           Shake hands
           (Hyams 1986:63,(1))

      b.   Taglio   E   mia palla   [Ita.]
           cut-PRES.1sg   be-PRES.3sg my ball
           'I cut'   'It is my ball'
           (Hyams 1986:111,(1c,e))

This is accounted for in terms of the pro-drop parameter (Chomsky 1981, Rizzi 1982). The pro-drop parameter is initially set as 'pro-drop', the unmarked value, for both Italian-speaking and English-speaking children. Not being exposed to the data against the first setting, Italian-speaking children do not change the first value. Triggered by the data against the first setting (e.g. the exposure to expletives), English-speaking children reset the first value as

'non-pro-drop'.

Roeper and Weissenborn (1990) claim that the *pro*-drop parameter setting is dependent on the *wh*-movement parameter setting. After subordinate clauses are learned, children of non-*pro*-drop languages such as German (9a) and French (9b) begin to have an overt subject in embedded clauses without exception.

(9)   a.   wenn der is dunkel   [Ger.]
           when it   is dark
           'when this is dark'

      b.   c'est ça   qu'on   fait   [Fre.]
           it's   it   what one does
           'that's it what one does'
           (Roeper and Weissenborn 1990:157,(20c,g))

Whether subordinate clauses are available to a child depends on whether the CP system works. It is activated by movement of a *wh*-phrase to [Spec,CP]. *Wh*-movement in embedded clauses is obligatory in languages such as French (10).

(10)   a.   il me demande   où   je vais   [Fre.]
            he me ask-PRES.3sg where I go-PRES.1sg
            'he asks me where I go'

       b.   *il me demande   je vais   où
             he me ask-PRES.3sg I go-PRES.1sg where
            'he asks me I go where'
            (Roeper and Weissenborn 1990:159,(23))

The non-pro-drop property is ultimately derived from the setting of the *wh*-movement parameter as '*wh*-movement at S-structure'. Thus, the *pro*-drop parameter setting is dependent on that of the *wh*-movement parameter setting.[6]

Note that the account is above provided in terms of the setting of parameters that affect syntactic operations. The *pro*-drop parameter setting can be attributed to morphological properties of individual languages, i.e., whether a language has a rich or poor inflectional system (Chomsky 1981, Rizzi 1982). Based on Roeper and Weissenborn (1990), however, the *pro*-drop parameter setting is preceded by the *wh*-movement parameter setting, which distinguishes languages that move *wh*-phrases overtly (e.g. English) from *wh*-in-situ languages (e.g. Japanese).

Recall the transition of the theoretical framework, specifically, the different accounts of *wh*-movement between the P&P theory and the currently assumed framework since Chomsky (2000). In the P&P period, the difference between English and Japanese was accounted for in terms of whether a *wh*-phrase moves at S-structure in English or at LF in

---

[6] There is a huge amount of the LA studies. See Guasti (2002) for a recent summarizing work of LA in generative grammar. See also Friedemann and Rizzi (2000, eds.) for recent comparative studies of LA.

Japanese. The syntactic difference was directly associated with the setting of the *wh*-movement parameter, which determined the syntactic level at which movement occurs. The current framework, however, posits that not only the semantic component but also the syntactic component are uniform. The parametric difference does not exist in the syntactic component. The syntactic operation that derives *wh*-questions proceeds in the same manner both in English and in Japanese. The only difference lies in which copy in a *wh*-chain is assigned phonological features, i.e. either the highest copy in English or the lowest copy in Japanese.

Thus, the LA theory is not compatible with the currently assumed theoretical framework: the argument for the setting of parameters that affect syntactic operations in the LA theory leads to the claim that the computational system differs between languages, contrary to the claim for the syntactic uniformity in the current framework. It might be argued under the cartographic system that final representations of syntactic operations have only to be the same among languages. This argument does not hold since there are neither different syntactic levels such as S-structure and LF nor movement operations that could occur at those different levels in the current system.[7]

In sum, the LA theory is incompatible with the currently assumed framework of the syntactic theory: the former assumes that the syntactic computation can differ between languages, whereas the latter does not accept that claim.

5.   Conclusion

In this paper, I have pointed out two theory-internal contradictions in generative grammar. First, the assumption that the parametric difference lies in the lexicon (Chomsky 2001) causes the LA paradox based on Sigurðsson (2004). That is, a child who acquires a lexicon that consists of a subset of universal features cannot access discarded universal features. The child and grown-up adult could not come to know the concepts that are attributed to those abandoned universal features, contrary to fact. I have claimed that the argument of the parametric difference in the lexicon holds only for languages that give some phonetic form to a target linguistic property, regardless of what phonetic form that linguistic property is given in relevant languages. Second, the LA theory provides accounts in terms of the setting of the parameters that affect syntactic operations. This argument leads to the claim that the computational system of human language can differ between languages, which is incompatible with the currently assumed framework: the latter does not accept that claim under the assumption of the uniformity of syntax.

---

[7] The same claim applies to the other literature. Rizzi (1994) proposes the 'truncation' theory, claiming that the CP structure is absent in English child language. But the CP layer occurs after *wh*-movement is acquired. Wexler (1994, 1996) proposes the 'very early parameter setting' of the verb movement parameter, which either raises finite verbs up to C (e.g. the Scandinavian languages), to T (e.g. the Romance languages) or do not raise them (e.g. English).

References

Borer, Hagit. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris Publications.

Borer, Hagit, and Kenneth Wexler. 1987. The Maturation of Syntax. In *Parameter Setting*, ed. by Thomas Roeper and Edwin Williams, 123-172. Dordrecht: D. Reidel Publishing Company.

Chomsky, Noam. 1959. A Review of B. F. Skinner's *Verbal Behavior*. *Language* 35:26-58.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. New York: Mouton de Gruyter.

Chomsky, Noam. 1986. *Knowledge of Language*. Westport, CT: Praeger.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step By Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89-156. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and Beyond*: *The Cartography of Syntactic Structures, Volume 3*, ed. by Adrian Belletti, 104-131. Oxford: Oxford University Press.

Chomsky, Noam. 2008. On Phases. In *Foundational Issues in Linguistic Inquiry: Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero and Maria Luisa Zubizarreta, 133-166. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads*. Oxford: Oxford University Press.

Culicover, Peter W., and Kenneth Wexler. 1977. Some syntactic implications of a theory of language learnability. In *Formal Syntax*, ed. by Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 7-60. New York: Academic Press, Inc.

Friedemann, Marc-Ariel, and Luigi Rizzi. (eds.) 2000. *The Acquisition of Syntax*. London: Longman.

Groat, Erich, and John O'Neil. 1996. Spell-Out at the LF Interface: Achieving a Unified Syntactic Computational System in the Minimalist Framework. In *Minimalist Ideas: Syntactic Studies in the Minimalist Framework*, ed. by Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and C. Jan-Wouter Zwart, 113-139. Amsterdam: John Benjamins Publishing Company.

Guasti, Maria Teresa. 2002. *Language Acquisition: the Growth of Grammar*. Cambridge, MA: MIT Press.

Huang, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. PhD dissertation, MIT.

Hyams, Nina. 1983/1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: D. Reidel Publishing Company.

Pesetsky, David. 2000. *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.

Pinker, Steven. 1984. *Language Learnability and Language Development*. Cambridge, MA:

Harvard University Press.

Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris.

Rizzi, Luigi. 1994. Early Null Subjects and Root Null Subjects. In *Language Acquisition Studies in Generative Grammar*, ed. by Teun Hoekstra and Bonnie D. Schwarz., 151-176. Amsterdam: John Benjamins Publishing Company.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar*, ed. by Liliane Haegeman, 281-338. Dordrecht: Kluwer.

Roeper, Thomas, and Jürgen Weissenborn. 1990. How to make parameters work: Comments on Valian. In *Language Processing and Language Acquisition*, ed. by Lyn Frazier and Jill de Villiers, 147-162. Dordrecht: Kluwer Academic Publishers.

Sigurðsson, Halldór Á. 2004. Meaningful Silence, Meaningless Sounds. In *Linguistic Variation Yearbook 2004*, ed. by Pierre Pica, 235–259. Amsterdam: John Benjamins Publishing Company.

Sigurðsson, Halldór Á. 2006a. Remarks on features. ms, Lund University.

Sigurðsson, Halldór Á. 2006b. The case of PRO. ms, Lund University.

Skinner, Burrhus F. 1957. *Verbal Behavior*. New York: Appleton-Century-Crofts.

Watanabe, Akira. 1992. Subjacency and S-Structure movement of *wh*-in-situ. *Journal of East Asian Linguistics* 1:255-291.

Wexler, Kenneth. 1994. Optional infinitives, head movement and the economy of derivations. In *Verb Movement*, ed. by David Lightfoot and Norbert Hornstein, 305-350. Cambridge: Cambridge University Press.

Wexler, Kenneth. 1996. The development of inflection in a biologically based theory of language acquisition. In *Toward a Genetics of Language*, ed. by Mabel L. Rice, 113-144. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Wexler, Kenneth, and Peter W. Culicover. 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.