IN THIS ISSUE

LETTERS & REPLIES

A RESPONSE TO ALAN PRINCE'S LETTER IN ISSUE 2-6

Morris Halle & William Idsardi

We write to respond to Alan Prince's letter in *Glot International* 2-6, and to continue the discussion of Optimality Theory (OT) following on Luigi Burzio's article in *Glot International* 1-6.

In section 1 of his letter, Prince contends that metrical theory now has developed answers to questions regarding English trochaic shortening, and more generally to questions such as:

Why shorten rather than lengthen? Indeed why do anything at all?

Prince's contention is wrong: there can be no answers to these questions because languages freely exercise all available options in this domain. For example, English shortens the head vowel in some trochaic feet, but lengthens it in others (CiV lengthening); Polish and many other languages differ from English in that they never change vowel quantity and thus effectively "do nothing at all." Both rule-based theories and OT are readily able to account for these facts. As noted by Halle, rule-based theories do this with the help of two disjunctively ordered rules, while OT deals with these facts by a judicious ranking of Stress-to-Weight and Faithfulness constraints so as to insure that stressed vowels are lengthened in some environments, shortened in others and left as they are in still other environments. Both theories are also equipped to deal with the facts of Polish and such other languages that have no vowel quantity alternations. In rule-based theories this is reflected formally by not including in the grammar rules that change vowel quantity; in OT Faithfulness (MaxIO, DepIO, IdentIO) constraints are allowed to outrank the constraints mandating change in vowel quantity. In all these cases "the principles of prosodic form" invoked by Prince do not and cannot play a determinative role. And analogouos answers hold for the Prince's other questions.

In section 2 Prince argues that in a derivational account the Elsewhere Condition is an extraneous add-on and that the required outputs can also be generated by ordering the more general rule of Shortening before the more special rule(s) of Lengthening. In other words, as Prince points out, if Shortening is ordered before Lengthening, "overwriting" will generate the correct outputs as illustrated in (1) (where "O" stands for a phonologically long vowel and "o" for its short cognate).

(1)
micrOb-ial --Shortening --> microb-ial --Lengthening--> micrOb-ial

While any account certainly must generate the correct outputs, this is not the only consideration that determines its acceptability. Among other factors to be considered is the nature of the derivations. As Chomsky (1995, p.220) has noted "a linguistic expression L cannot be defined just as a pair (π, λ) formed by a convergent derivation. Rather, its derivation must be optimal, satisfying certain natural economy conditions: locality of movement, **no 'superfluous steps' in derivations,** and so on" (emphasis added.)

Derivations such as (1), which have been dubbed by Pullum (1976) the "Duke of York gambit", have justly been singled out as unacceptable because of the superfluous steps in them. The fact that (1) is ruled out by the Elsewhere Condition is part of the evidence in favor of the Elsewhere Condition as a general principle of UG.

In Halle and Idsardi (1997) we suggested an additional motivation for the appearance in a language of pairs of rules of which the Elsewhere Condition holds. We proposed that such rule pairs are the result of an attempt by speakers of a dialect to reinstitute a phonological contrast which had been eliminated in certain contexts. In view of the Elsewhere Condition the rule reinstituting the contrast must be more specific than the rule eliminating the constrast. In our example we considered Eastern Massachusetts English, which like many English dialects has long had a rule eliminating coda r's. Because of the Elsewhere Condition, the innovative rule that re-establishes coda r's must be more specific than the rule already in the dialect. The solution chosen by the

Continued on page 22

GLOT INTERNATIONAL

Volume 3
ISSUE 1

Monthly Magazine for Linguists

January 1998

ISSN 1381–3439

Price single issue: Dfl 14,50

Glot International appears monthly, except June & July.

Colophon

Editors
Lisa Lai-Shen Cheng
Rint Sybesma
Holland Institute of Generative Linguistics
address: Department of General Linguistics | Leiden University | P.O. Box 9515 | 2300 RA Leiden | The Netherlands | phone: +31 71 5272104 | fax: +31 71 5272615 | e-mail: glot@rullet.leidenuniv.nl/ llcheng@uci.edu

Conference reports and announcements editor
Jan-Wouter Zwart
Groningen University | Department of linguistics | Oude Kijk in 't Jatstraat 26 | 9712 JK Groningen | The Netherlands | e-mail: zwart@let.rug.nl

Goodies editor
Andrew Carnie
Department of Linguistics | Harvard University | 77 Dunster Street | Cambridge MA 02138 | USA | e-mail: carnie@linguistlist.org

Correspondents/ Regular contributors
Eulàlia Bonet (Barcelona) | Marcel den Dikken (Vrije Universiteit Amsterdam) | Paula Fikkert (Konstanz) | Helen de Hoop (Utrecht) | Guido Vanden Wyngaerd (Brussels) | Shohei Yoshida (Yokohama National University)

Columnists
Elan Dresher (Toronto), Neil Smith (London)

Information for contributors
Linguists who want to contribute a State-of-the-Article or a Review are invited to contact the Editors.
Linguists who have recently finished their dissertation are invited to forward a copy to the Editors.
Everybody is welcome to contribute to *Glot International*. But before you send in anything, please, contact the Editors.
Glot International prints conference announcements, calls for papers, conference programs, etc.

Subscription information
Glot International appears every month, except in June and July. Subscriptions are for the printed or the electronic (pdf) version.

Prices for Volume 3 (1998):

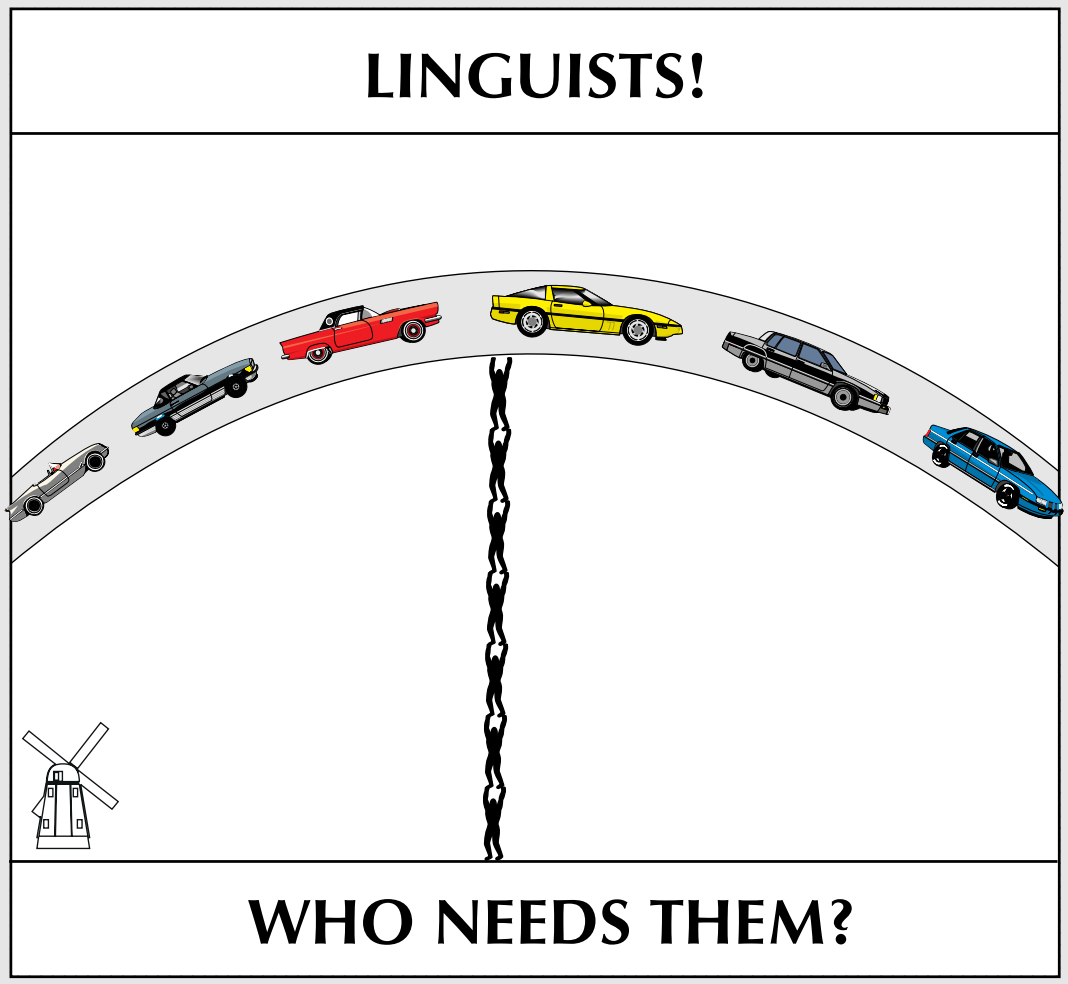
INSTITUTIONAL RATE		NLG	120
INDIVIDUAL RATE			
pdf only		NLG	35
printed version		NLG	50
POSTAGE			
surface mail	Netherlands	NLG	32.5
	other countries	NLG	37.5
air mail	Europe	NLG	42.5
	other countries	NLG	47.5
surface shipping in three batches (March—August—December)			
	Netherlands	NLG	15
	other countries	NLG	20

Subscriptions at the individual rate must be ordered directly from the publisher.
Subscriptions at the institutional rate may be ordered from your bookseller or subscription agent, or directly from the publisher.
Prices exclude 6% VAT (EC only).
Orders must be prepaid.
Payments preferably by giro, credit card (American Express, Euro/Access/Master, Visa) or international money order. Credit card orders must include the complete cardholder's address. There is an additional charge of Dfl 17,50 for payments by international bank cheque or bank transfer.

Publisher
Holland Academic Graphics
Julius Röntgenstraat 15 | 2551 KS The Hague | The Netherlands | phone +31 70 4480203 | fax +31 70 4480177 | e-mail: mail@hag.nl | http://www.hag.nl | bank acc. ABN AMRO 49 88 22 796 | postal giro 6261107 | VAT nr. NL 0588.61.798.B01

Advertisements
Contact the publisher

Printing
Printed in The Netherlands



Contributors to this issue

Sjef Barbiers
HIL/Leiden University | Dutch Studies | P.O. Box 9515 | 2300 RA Leiden | The Netherlands
barbiers@rullet.leidenuniv.nl

Kleanthes Grohmann
University of Maryland | Linguistics Department | 1401 Marie Mount Hall | College park | MD 20742 | U.S.A.
grohmann@wam.umd.edu

Morris Halle
M.I.T. | Department of Linguistics and Philosophy | Bldg. E39-245, 77 Massachusetts Ave. | Cambridge, MA 02139 | U.S.A
halle@mit.edu

Heidi Harley
University of Pennsylvania | Department of Linguistics | 619 Williams Hall | Philadelphia PA 19104-6305 | U.S.A.
hharley@ling.upenn.edu

Hiroto Hoshi
S.O.A.S./University of London | Russell Square | London WC1H OXG | U.K.
hh5@soas.ac.uk

William Idsardi
University of Delaware | Department of Linguistics | 46 E. Delaware Ave | Newark DE 19716 | U.S.A.
idsardi@udel.edu

David Lightfoot
University of Maryland | Linguistics Department | 1401 Marie Mount Hall | College Park | MD 20742 | U.S.A.
dlight@deans.umd.edu

Neil Smith
University College London | Department of Linguistics | Gowerstreet | London WC1E 6BT | U.K.
neil@linguistics.ucl.ac.uk

Yuji Takano
Keio University | 4-1-1 Hiyoshi, Kohoku-ku | Yokohama 223 | Japan
ytakano@hc.keio.ac.jp

Keren Rice
University of Toronto | University of Toronto | 130 St. George Street | Toronto, Ontario M5S 3H1 | Canada
rice@chass.utoronto.ca

Jan-Wouter Zwart
University of Groningen | Department of Linguistics | Oude Kijk in 't Jatstraat 26 | 9712 EK Groningen | The Netherlands
zwart@let.rug.nl

Dissertations reviewed

Sjef Barbiers, *The syntax of interpretation*. November 1995. HIL/Leiden University. Supervisors: J. de Vries, H. Bennis. 225 pp. ISBN: 90-5569-011-2. [HIL Dissertations 14.] Available from Holland Academic Graphics, The Hague. mail@hag.nl. Price: hfl. 40,-.

Yuji Takano, *Movement and Parametric Variation in Syntax*. March 1996. University of California, Irvine. Supervisor: Naoki Fukui. 272 pp. Available from Irvine Linguistics Students Association, School of Social Sciences, University of California, Irvine, CA 92697, U.S.A. Price: $14.

Goodies

Moraic for Macintosh. Cascadilla Press, 1997. <www.cascadilla.com>, <sales@cascadilla.com>, or +1-617-776-2370. Price: $20 (US); $24 (outside of US; incl. Shipping); 20% discount for purchase of ten or more; site licenses: $600 (US), 620 (outside of US) for: Arboreal (Mac and Windows), Moraic (Mac and Windows), Logic Times.

THE DEVELOPMENT OF GRAMMARS

David Lightfoot

Some aspects of how syntactic systems change over time are a function of the way in which they are acquired by children. In the following article, David Lightfoot introduces us to a number of aspects of the art of explaining change through acquisition.

1. The development of grammars in children

Grammars are mental entities which develop in the mind/brain of individual children. This development is data-driven only in part. Researchers have postulated genotypical principles which are available independently of experience and which do not have to be learned. These principles determine similarities among grammars, recurrent properties which hold of all grammars. Alongside the invariant principles, it is customary to postulate grammatical *parameters*, which children set on the basis of their linguistic experience and which account for grammar variation. So language acquisition proceeds as children set the parameters defined by Universal Grammar (UG), i.e. those genotypical principles and parameters which are relevant for the emergence of language in an individual. The parameters of UG are structural and abstract and that accounts for the "bumpiness" of language variation; even closely related languages generally differ from each other in terms of *clusters* of superficial phenomena. Here I shall discuss the nature of the experience which triggers the development of grammars, arguing that children scan their environment for certain designated structures or "cues" and that they are not influenced by the set of sentences generated by their grammars. Indeed, there are no independent "parameters"; rather, some cues are found in all grammars and some are found only in certain grammars — the latter constitute points of variation.

There is a second kind of development, which we shall turn to in the next section. Grammars may also develop from generation to generation. This is diachronic change. The central mystery for historical linguists taking our grammatical perspective is why they have anything to study: why do changes take place and why are languages not generally stable? In particular, why do changes sometimes take place abruptly and catastrophically? If people produce utterances corresponding fairly closely to the capacity of their grammars, then children exposed to that production would be expected to converge on the same grammar. This is what one would expect if grammars have structural stability, as they must to some degree; children are not "trigger happy," developing different grammars whenever their trigger experiences differ just a little. In that case, change would be expected only if there is some major disruption due to population movement (see below).

Not only is stability what one would expect naively and pretheoretically, but it is also what current learnability models would lead one to expect. Chomsky (1965) viewed children as endowed with a metric evaluating grammars which could generate the primary data to which they are exposed, along with appropriate structural descriptions for those data. The evaluation metric picked the grammar which conformed to the invariant principles of UG and was most successful in generating those data and those structural descriptions. The child selected a grammar which matched her input as closely as possible. Again, if the data and the associated structural descriptions to which the child is exposed correspond fairly closely to the grammatical capacity of some older individual, one would expect the child's evaluation metric to select the same grammar as that older individual's.

The same point holds for more recent models. Gibson & Wexler (1994) posit a Triggering Learning Algorithm (TLA), under which the child-learner uses grammars to analyze incoming sentences and eventually converges on the correct grammar. Gibson & Wexler distinguish global and local triggers, but both are sentence-types (op cit: p.409). If the child-learner cannot analyze a given sentence with the current grammar, then she follows a certain procedure to change one of the current parameter settings and tries to reprocess the sentence using the new set of parameter values. If analysis is now possible, then the new parameter value is adopted, at least for a while. So the TLA is error-driven and permits the child to pinpoint which parameter setting is incorrect when the learner's grammar does not give the right results. There is much to be said about the way that this model works and Dresher (1997) has illuminating discussion, but what is crucial here is that the model has the child seeking grammars which permit analysis of incoming data, where the data consist of more or less unanalyzed sentences. Gibson & Wexler's Table 3 correlates sets of three parameter settings (Specifier-final/initial, Comp-final/initial, +/−verb-second) and sets of data (listed here in terms of primitives like Subject, Verb, First Object, Second Object). When exposed to some data set (righthand column), the child selects the appropriate grammar (lefthand column) … although it would not be easy for the child to know which data set she is exposed to.

Parameter settings	*Data in defined grammar*
Spec-final Comp-final -V2 (VOS)	V S, V O S, V O1 O2 S Aux V S, Aux V O S, Aux V O1 O2 S, Adv V S Adv V O S, Adv V O1 O2 S, Adv Aux V S Adv Aux V O S, Adv Aux V O1 O2 S
Spec-final Comp-final +V2 (VOS + V2)	S V, S V O, O V S, S V O1 O2, O1 V O2 S, O2 V O1 S S Aux V, S Aux V O, O Aux V S S Aux V O1 O2, O1 Aux V O2 S, O2 Aux V O1 S Adv V S, Adv V O S, Adv V O1 O2 S, Adv Aux V S, Adv Aux V O S, Adv Aux V O1 O2 S
Spec-final Comp-first -V2 (OVS)	V S, O V S, O2 O1 V S V Aux S, O V Aux S, O2 O1 V Aux S, Adv V S Adv O V S, Adv O2 O1 V S, Adv V Aux S Adv O V Aux S, Adv O2 O1 V, Aux S
Spec-final Comp-first +V2 (OVS + V2)	S V, O V S, S V O, S V O2 O1, O1 V O2 S, O2 V O1 S S Aux V, S Aux O V, O Aux V S S Aux O2 O1 V, O1 Aux O2 V S, O2 Aux O1 V S Adv V S, Adv V O S, Adv V O2 O1 S, Adv Aux V S, Adv Aux O V S, Adv Aux O2 O1 V S
Spec-first Comp-final -V2 (SVO)	S V, S V O, S V O1 O2 S Aux V, S Aux V O, S Aux V O1 O2, Adv S V Adv S V O, Adv S V O1 O2, Adv S Aux V Adv S Aux V O, Adv S Aux V O1 O2
Spec-first Comp-final +V2 (SVO + V2)	S V, S V O, O V S, S V O1 O2, O1 V S O2, O2 V S O1 S Aux V, S Aux V O, O Aux S V S Aux V O1 O2, O1 Aux S V O2, O2 Aux S V O1, Adv V S Adv V S O, Adv V S O1 O2, Adv Aux S V, Adv Aux S V O, Adv Aux S V O1 O2
Spec-first Comp-first -V2 (SOV)	S V, S O V, S O2 O1 V S V Aux, S O V Aux, S O2 O1 V Aux, Adv S V Adv S O V, Adv S O2 O1 V, Adv S V Aux Adv S O V Aux, Adv S O2 O1 V Aux
Spec-first Comp-first +V2 (SOV + V2)	S V, S V O, O V S, S V O2 O1, O1 V S O2, O2 V S O1 S Aux V, S Aux O V, O Aux S V S Aux O2 O1 V, O1 Aux S O2 V, O2 Aux S O1 V Adv V S, Adv V S O, Adv V S O2 O1, Adv Aux S V, Adv Aux S O V, Adv Aux S O2 O1 V

Clark (1992) offers a similar kind of model but one which differs from that of Gibson & Wexler in that the child cannot pinpoint the source of a grammar's failure, revising particular parameter settings. Clark posits a Darwinian competition between grammars needed to parse sets of sentences. All grammars allowed by UG are available to each child and some grammars are used more than others in parsing what the child hears. A "genetic algorithm" picks those grammars whose elements are activated most often. A Fitness Measure compares how well each grammar fares, and the fittest grammars go on to reproduce in the next generation, while the least fit die out. Eventually the candidate grammars are narrowed to the most fit and the child converges on the correct grammar. Clark & Roberts (1993) used this model to give an account of changes affecting the verb-second properties of early French, by allowing an arbitrary degree of misconvergence by children.

There is a serious, technical problem with Clark's Fitness Measure. There is no reason to suppose that a grammar with more parameters set correctly will be more successful in parsing/generating incoming data. Dresher (1997) illustrates this by considering the settings needed to generate the phonological stress system of Selkup, computing the relative score the Fitness Measure would give them when applied to eight representative words. It isn't obvious what criterion the Fitness Measure should use, so he tried three different criteria: words correct, syllables correct, and main stress correct. Some results were (1).

(1)

	Parameters	*correct*	*Words*	*correct*
a.	4/10	40%	2/8	25%
b.	6/10	60%	1/8	12.5%
c.	7/10	70%	4/8	50%
d.	8/10	80%	5/8	62.5%
e.	9/10	90%	5/8	62.5%
f.	9/10	90%	3/8	37.5%

	Syllables	*correct*	*Main stress*	*correct*
a.	7/20	35%	4/8	37.5%
b.	7/20	35%	6/8	62.5%
c.	12/20	60%	7/8	50%
d.	14/20	70%	8/8	62.5%
e.	14/20	70%	9/8	62.5%
f.	10/20	50%	9/8	37.5%

Candidates (e) and (f) are each correct in all but one (different) parameter, but they are very different in their apparent fitness. (e) scores high, but no higher than (d), which has fewer correct settings. Candidate (f), with only one parameter wrong, scores worse in every category than (c), which has three parameters wrong. And (a) does better than (b), despite having only four correct parameter settings. Dresher also points out that these results can be influenced in unpredictable ways by the chance occurrence of various types of words. As a result, there is no simple relationship between success and the number of parameters set correctly, which is a problem for Clark's Fitness Measure.

What these models have in common is that learners eventually match their input, in the sense that they select grammars which generate the sentences of the input. Models of this type can characterize instances of language stability straightforwardly. The child converges on a grammar which analyzes the input successfully, where the input consists of sets of sentences, elements of E-language in the terminology of Chomsky (1986). In that case, the grammar will resemble closely the grammar/grammars which generate that input. Such models can also handle cases of mixed input under conditions of population movement. There again the child is presented with a set of data, in this case data yielded by diverse grammars; she converges on a grammar which is most successful in generating that dataset, sometimes a grammar quite different from any of those in the previous generation. This would be a case of grammar change and the new grammar might yield structural descriptions and some sentences which differ from those of the

input; but the new grammar would result from the child's effort to match the input sentences as closely as possible.

Of course, these are not *pure* input-matching models of the type advocated by MacWhinney & Bates (1989), in which it is mysterious why children should ever produce non-adult forms in any systematic way. Clark, Gibson, and Wexler's children are not dependent only on the input; they operate in a space defined by UG. Consequently, each intermediate stage for the developing child is represented by some set of UG-defined parameter settings, and that set may generate non-adult forms. However, it is a fact that sometimes children do not match their input at any stage, including the final stage. One instance would be abrupt, catastrophic change. It is difficult to see how failure to match input would be handled by models which are inherently input-matching.

Here I want to argue the following: existing models of learnability commit us to insisting that languages are basically stable. This conforms to the views of many historians that change is inherently piecemeal and gradual. But a better model of learnability enables us to better understand historical change. Under this model we would expect language change sometimes to be abrupt, sudden, and "catastrophic." Our learnability model, in turn, will allow us to capture the contingent nature of historical change and to avoid the excessively principled accounts of change offered by many historians.

Ironically, the best worked out model of parameter setting comes from phonology and the work of Dresher & Kaye (1990). Parameters have not played an extensive role in the phonological literature, but Dresher and Kaye identified parameters for stress systems, a rather well-studied area of phonology. Furthermore, they developed a "cue-based" theory of acquisition, now clarified, elaborated, and generalized by Dresher (1997). Under this view, UG specifies not only a set of parameters, but also for each parameter a cue. I amend this view slightly and say that cues which are realized only in certain grammars constitute the parameters, the points of variation between grammars. A cue is some kind of structure, an element of I-language, which is derived from the input. The cues are to be found in the mental representations which result from hearing, understanding, and "parsing" utterances. As a child understands an utterance, even partially, she has some kind of mental representation of the utterance. These are partial parses, which may differ from the full parses that an adult has. The learner scans those representations, derived from the input, and seeks the designated cues. If a cue is found, it is incorporated into the emerging grammar. Furthermore, the child scans the linguistic environment for cues only in simple domains; this is the "degree-0 learnability" of Lightfoot (1991, 1994). Learners do not try to match the input; rather, they seek certain abstract structures derived from the input (elements of I-language), looking only at structurally simple domains, and they act on this without regard to the final result. That is, a child seeks cues and may or may not find them, regardless of what the emerging grammar can generate; the output of the grammar is entirely a by-product of the cues that the child finds, and the success of the grammar is in no way based on the set of sentences that it generates, unlike in input-matching models. The child's triggering experience, then, is best viewed as a set of abstract structures manifested in the mental representations which result from parsing utterances; some of those representations constitute partial parses, which lack some of the information found in mature, adult parses.

Dresher (1997) illustrates the cue-based model of acquisition with some phonological parameters. The essential feature is that a cue-based learner does not try to match target input forms, but uses them as sources of cues. The trigger consists not of sets of sentences but rather of partially analyzed syntactic structures, elements of I-language; these are the mental representations resulting from parsing utterances. So cues are intensional elements, grammar fragments. A cue-based learner determines whether a Specifier precedes or follows its head on the basis of exposure to data which must be analyzed with a Spec preceding its head, e.g. [$_{\text{Spec}}$[John's] $_{\text{N}}$[hat]]. This cue may be identified only when the child has a partial analysis which treats *John's* and *hat* as separate words, the latter a head noun, etc. In this way, the order in which parameters appear to be set, the "learning path" (Lightfoot 1989), reflects dependencies among cues and follows from their internal architecture.

Less trivially, a cue-based learner acquires a verb-second grammar not by evaluating grammars against sets of sentences but on exposure to structures commencing with a XP followed immediately by a finite V, where there is no fixed grammatical or thematic relation between the initial phrasal category and the finite verb, effectively where the initial XP is a non-subject (Lightfoot 1997b). This requires analyzing the XP as in SpecCP and so $_{\text{SpecCP}}$[XP] is the cue for a verb-second system; the cue must be represented robustly in the mental representations resulting from parsing the primary linguistic data (PLD).

Some version of this cue-based approach to acquisition is **implicitly** assumed in some earlier work, notably in the work of Nina Hyams (1986, 1996) and in my own work (Lightfoot 1989, 1991). It has been productive for phonologists concerned with the parameters for stress systems (Dresher & Kaye 1990; Dresher 1997; Fikkert 1994, 1995), it has been invoked for some syntactic problems by Fodor (1998), and it represents something quite different from the input-matching learning algorithms of Gibson & Wexler, Clark, and others. In fact, I see no reason to believe that there is any learning algorithm beyond the information provided specifically by UG.

2. Diachronic development of grammars

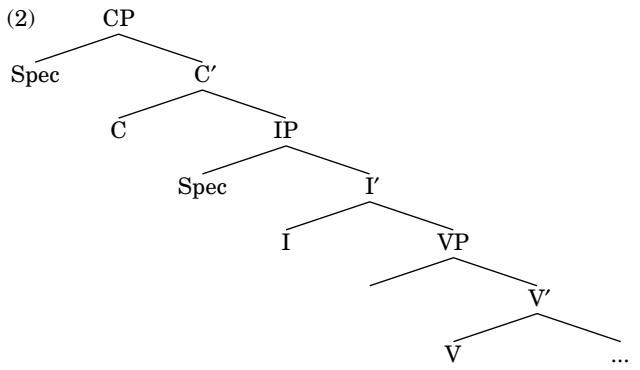
Turning now to language change, we note that the speech of no two people is identical, so it follows naturally that if one takes manuscripts from two eras, one will be able to identify differences and so point to language "change". In this sense languages are constantly changing in piecemeal, gradual, chaotic, and relatively minor fashion. However, historians also know that languages sometimes change in a bumpy fashion, several things changing at the same time, and then settle into relative stasis, in a kind of "punctuated equilibrium," to borrow a term from evolutionary biology. From the perspective adopted here, it is natural to try to interpret cascades of changes in terms of unitary changes in **grammars,** sometimes having a wide variety of surface effects and perhaps setting off a chain reaction. So grammatical approaches to language change have focussed on these large-scale changes, assuming that the clusters of properties tell us about the harmonies which follow from particular parameters, from identifying particular cues. By examining the clusters of simultaneous changes and by taking them to be related by properties of UG, we discover something about the nature of cues and about how they are identified.

Let us consider one case of a grammatical change, which is partially understood, using it as a case-study to show what further work is needed. It will show how the study of a change is intimately connected, under this approach, with work on grammatical theory and on cue-based acquisition.

Operations which associate inflectional features with the appropriate verb appear to be parameterized, and this has been the subject of a vast amount of work covering many languages (see, for example, the collection of papers in Lightfoot & Hornstein 1994). We can learn about the cues by considering how the relevant grammars could be attained, and that in turn is illuminated by how some grammars have changed.

For ease of exposition, I follow work by Emonds (1978) and Pollock (1989) and I adopt the familiar basic clause-structure of (2).

(2)



Subjects occur in SpecIP and *wh* elements typically occur in SpecCP. Heads raise from one head position to another, so verbs may raise to I and then further to C. In fact, many grammars raise their verbs overtly to the position containing the inflectional elements (3), but English grammars, unusually, do not. We know this because English finite verbs cannot be separated from their complements by intervening material (4a) and do not occur in some initial C-like position (4b).

(3)

a. Jeanne $_{\text{I}}$[lit$_{\text{i}}$] $_{\text{VP}}$[toujours e$_{\text{i}}$ les journaux]
b. lit$_{\text{i}}$ $_{\text{IP}}$[elle e$_{\text{i}}$ $_{\text{VP}}$[toujours e$_{\text{i}}$ les journaux]

(4)

a. *the women visited not/all/frequently Utrecht last week.
b. *visited you Utrecht last week?

What is it that forces French children to have the overt V-to-I operation and what forces English children to lack the operation?

It is reasonable to construe the English analysis as the default, as argued in Lightfoot (1993), Lasnik (1995), Roberts (1998). There is no evidence available to the English-speaking child which would force her to select a covert movement over an overt, syntactic V-to-I movement. Children would need to know that (4a,b) do not occur, but these are negative data, therefore unavailable as input to children. In that case, the covert movement is the default setting.

Now one can ask what triggers the availability of an overt, syntactic V-to-I raising operation in grammars where it may apply. Some generalizations have emerged over the last several years. One is that languages with rich inflection may have overt V-to-I operations in their grammars, and rich inflection could be part of the trigger (Rohrbacher 1994). However, the presence of V-to-I raising cannot be linked with rich inflection in a simple, one-to-one fashion. It may be the case that if a language has rich inflection, then overt V-to-I raising is available (Lightfoot 1991; Roberts 1998). If there is no rich inflection, a grammar may have the raising operation (Swedish — see Lightfoot 1997a, note 5) or may lack it (English). Indeed, English verb morphology was simplified radically and that simplification was complete by 1400; however, overt V-to-I movement disappeared only in the eighteenth century, so there was a long period when English grammars had very little verbal inflection but did have V-to-I movement. In that case, there needs to be a *syntactic* trigger for V-to-I movement. So, for example, a finite verb occurring in C, i.e. to the left of the subject NP (as in a verb-second language or in interrogatives), could only get there by raising first to I, and therefore inversion forms like (3b) in French could be syntactic triggers for V-to-I (see also Faarlund 1990 and Vance 1995 for illuminating discussion bearing on these matters).

Under a cue-based acquisition approach, one would say that the cue for grammars raising V to I is a finite verb in I, i.e. $_{\text{I}}$[V], an element of I-language. Children seek this cue in the representations resulting from their (partial) parses. One unambiguous instance of $_{\text{I}}$[V] is an I containing the trace of a verb which has moved on to C, as in the structure of (3b). Indeed, I would guess that this would be a very important expression of the cue, and I doubt that structures like (4a) would be robust enough to trigger V-to-I in isolation; this can be tested (see below). Adopting terminology from Clark (1992), one can ask how robustly the cue is "expressed;" it is **expressed** robustly if there are many simple utterances which can be

analyzed by the child only as $_{I}$[V]. So, for example, the sentences of (3a,b) can only be analyzed by the French child (given what the child has already established about the emerging grammar) if the V *lit* raises to I; a simple sentence like *Jeanne lit les journaux* 'Jeanne reads the newspapers,' on the other hand, could be analyzed with *lit* raised to I overtly or covertly in the English style, and therefore it does not express the cue for the V-to-I operation.

Early English grammars manifested the V-to-I operation, but later grammars do not; the operation was lost at some point. From the perspective adopted here, the operation ceased to be cued. The cue for V-to-I raising, $_{I}$[V], came to be expressed less in the light of three developments in early Modern English.

First, the modal auxiliaries (*can, could, may, might, shall, should, will, would, must*), while once instances of verbs that could raise to I, were re-categorized such that they came to be base-generated as instances of I; they were no longer verbs and so sentences with a modal auxiliary ceased to include $_{I}$[V] and ceased to express the cue for V-to-I movement. Sentences with a modal auxiliary, *Kim must leave*, are very common in ordinary speech addressed to young children, and the recategorization meant that they no longer expressed the cue. Sentences of this form existed at all stages, of course, but they came to be analyzed very differently after the change in category membership. The evidence for the recategorization is the obsolescence of (5), which follows if the modal auxiliaries are generated in I and therefore can occur only one per clause (5a), without an aspectual affix (5b,c), and mutually exclusively with the infinitival marker *to*, which also occurs in I (5d).

(5)
a. John shall can do it
b. John has could do it
c. canning do it
d. I want to can do it

This change has been discussed extensively in Lightfoot (1979, 1991), Kroch (1989), Roberts (1985, 1993a), Warner (1983, 1993), and there is consensus that it was complete by the early sixteenth century.

Second, as periphrastic *do* came to be used in negatives like *John did not leave* and interrogatives like *did John leave?*, so there were still fewer instances of $_{I}$[V]. Before periphrastic *do* became available, sentences like *the women visited all Utrecht last week* (4a) and *visited you Utrecht?* (4b) expressed the $_{I}$[V] cue. Periphrastic *do* began to occur in significant numbers at the beginning of the fifteenth century and steadily increased in frequency until it stabilized into its modern usage by the mid-seventeenth century. Ellegård (1953) shows that the sharpest increase came in the period 1475–1550.

Third, in early grammars with the much-discussed verb-second system all matrix clauses had a finite verb in C. Therefore many matrix clauses expressed the cue for V-to-I, $_{I}$[V], (on the assumption that V could move to C only by moving first to I). As these grammars were lost and as finite verbs ceased to occur regularly in C, so the expression of the cue for overt V-to-I raising was reduced correspondingly.

By quantifying the degree to which a cue is expressed, we can understand why English grammars lost the V-to-I operation and why they lost it after the modal auxiliaries were reanalyzed as non-verbs, as the periphrastic *do* became increasingly common, and as the verb-second system was lost. We can reconstruct a plausible history for the loss of V-to-I in English. What we are doing here is identifying when grammars changed and how the available triggering experiences, specifically those expressing the cue, seem to have shifted in critical ways prior to the grammatical change (Warner 1995 adopts the same logic). We know from acquisition studies that children are sensitive to statistical shifts in input data. For example, Newport, Gleitman & Gleitman (1977) showed that the ability of English-speaking children to use auxiliaries appropriately results from exposure to non-contracted, stressed forms in initial positions in yes-no questions: the greater the exposure to these subject-auxiliary inversion forms, the earlier the use of auxiliaries in medial position. Also Richards (1990) demonstrated a good deal of individual variation in the acquisition of English auxiliaries as a result of exposure to slightly different trigger experiences. The issue is when trigger experiences differ critically, i.e. in such a way as to set some parameter differently.

Our conclusion in earlier work was that V-to-I movement was lost in the seventeenth century, much later than suggested by Kroch (1989), Roberts (1993a) and others (in fact, Kroch's own figures from Ellegård show several sentence-types — positive intransitive questions, negative declaratives, and positive *wh*-object questions — with *do* less than 40% of the time at the very end of the sixteenth century, showing that V-to-I grammars were still very much in use). Warner (1997) now argues that the operation may have been lost as late as in the eighteenth century. He offers some statistics from Ellegård (1953) and Tieken-Boon van Ostade (1987). Ellegård shows that interrogative inversion with nonauxiliary in positive clauses (i.e. *came he to London?* as opposed to *did he come to London?*) occurred 27% of the time for 1625–50; 26% for 1650–1700. Tieken-Boon van Ostade shows a drop to 13% in the eighteenth century. Negative declaratives with a nonauxiliary (*he came not to London* as opposed to *he did not come to London*) occur 68% in 1625–1650, 54% in 1650–1700, dropping sharply to 20% in the eighteenth century. The drop is actually sharper than these figures suggest; Tieken-Boon van Ostade's figures for the later period include a high proportion of recurrent items (*know, doubt*, etc) which Ellegård omitted. A particularly interesting feature of these figures is the discrepancy between the interrogatives and the negatives, which lends some support to the hunch (above) that structures like those underlying (3b) are a more effective expression of the cue $_{I}$[V] than structures like those of (4a). In any case, we see that structures like (4a) were robust and widely attested in the texts of the late seventeenth century and then they disappeared rapidly — the kind of bumpiness that the abstractness of the cues leads us to expect.

The historical facts, then, suggest that lack of rich subject–verb agreement cannot be a sufficient condition for absence of overt V-to-I, but it may be a necessary condition. Under this view the **possibility** of V-to-I not being triggered first arose in the history of English with the loss of rich verbal inflection; similarly in Danish and Swedish. That possibility never arose in Dutch, French, German, where verbal inflections remained relatively rich. Despite this possibility, V-to-I continued to be triggered and it occurred in grammars well after verbal inflection had been reduced to its present-day level. However, with the re-analysis of the modal auxiliaries, the increasing frequency of periphrastic *do* and the loss of the verb-second system, the expression of $_{I}$[V] in English became less and less robust. That is, there was no longer anything very robust in children's experience which had to be analyzed as $_{I}$[V], i.e. which **required** overt V-to-I, given that the covert operation was always available as the default. In particular, sentences like (4a) with post-verbal adverbs and quantifiers had to be analyzed with the V in I but these cues were not robust enough to set the parameter and they disappeared quickly, a by-product of the loss of V-to-I.

This suggests that the expression of the cue dropped below some threshold, leading to the elimination of V-to-I movement. The next task is to quantify this generally; we should recognize that the gradual reduction in the expression of $_{I}$[V] is not crucial, but rather the point at which the phase-transition took place, when the last straw was piled on to the camel's back. This can be demonstrated by building a population model, tracking the distribution of the $_{I}$[V] cues in the PLD, and identifying the point at which the parameter was reset and V-to-I ceased to be triggered (differing, of course, from one individual or one dialect area to another). This work remains to be done (see below) but one hopes to find correlations between the changing distribution of the cue and the change in grammars.

3. Some context and comparisons

This grammatical, cue-based approach to diachrony **explains** changes at two levels. First, the cues postulated as part of UG which embody the points of parametric variation explain the unity of the changes, why superficially unrelated properties cluster in the way that they do. Second, the cues permit an appropriately contingent account of why the change took place, why children at a certain point converged on a different grammar: the expression of the cues changed in such a way that a threshold was crossed and a new grammar was acquired. That is as far as this model goes and it has nothing to say about why the distribution of the cues should change. That may be explained by claims about language contact or socially defined speech fashions but it is not a function of theories of grammar, acquisition or change, … except under one set of circumstances, where the new distribution of cues results from an earlier grammatical shift; in that circumstance, one has a "chain" of grammatical changes. One example would be the recategorization of the modal auxiliaries (above) which contributed to the loss of V-to-I. One can, of course, embed these grammatical accounts in an appropriate model of population change; so, if Kroch & Taylor (1997) are correct, the loss of verb-second from English grammars (another contributing factor for the loss of V-to-I) was a function of contact between speakers with distinct grammars.

Notice that this approach to change is independent of any particular grammatical model. Warner (1995) offers a persuasive analysis of parametric shift using a lexicalist HPSG model, quite different from the one assumed here. Interesting diachronic analyses have been offered for a wide range of phenomena, invoking different grammatical claims: Fontana (1993), van Kemenade (1987), Pearce (1990), Roberts (1993a,b, 1994, etc.), Sprouse & Vance (1997), Vance (1995) and many others.

Our general approach to abrupt change, where children acquire very different systems from those of their parents, comports with work on creolization under the view of Bickerton (1984, 1998), and the acquisition of signing systems by children exposed largely to unnatural input (Goldin-Meadow & Mylander 1990; Newport 1998; Supalla 1990). For several years Bickerton has worked on plantation creoles, where new languages appear to be formed in the space of a single generation. He argues, surely correctly, that situations in which "the normal transmission of well-formed language data from one generation to the next is most drastically disrupted" will tell us something about the innate component and how it determines acquisition (Bickerton 1998); it certainly shows that children do not always proceed by converging on grammars which match the input.

It is hard to see how input-matching models can succeed when children are exposed to unusual amounts of artificial and degenerate data, which in fact are not matched. In particular, it is hard to see how they could account for the early development of creole languages, as described by Bickerton and others. In these descriptions, early creole speakers are not matching their input, which typically consists to a large degree of pidgin data. Pidgins are primitive communication systems, cobbled together from fragments of two or more languages. They are not themselves natural languages and they tend not to last long, before giving way to a creole with all the hallmarks of a natural grammar. The first speakers of creoles go way beyond their input in some ways and in other ways fail to reproduce what they heard from their models, arriving at gram-

mars which generate sentences and structural descriptions quite different from those of the input.

Let us call this the "abrupt" view of creolization (following Thomason & Kaufman 1988). There is a dramatic discrepancy between what early creole speakers hear in childhood and what their mature grammars eventually characterize as well-formed, much greater than in non-creole contexts. The abrupt view of creolization is more controversial than it should be. It offends a commitment to the proposition that languages generally change only gradually. This commitment is linked to a highly data-driven view of language acquisition, and it is widely and deeply held, including by creolists. Creolists committed to gradualism (e.g. Carden & Stewart 1988) insist that creoles emerge gradually as a result of changes introduced primarily by adults, as they relexify their own languages. Clearly gradual change exists and this is part of the story. However, if this is generally true, if this is most of the story, and if creolization for the most part mirrors adult second language learning and is not abrupt and instantiated by children, then there is little reason for theoreticians to be interested in the phenomenon. Our data about the early stages of creole languages generally are not very rich, and if one is interested in adult second language learning, one is probably better off refining theories in the light of better data-sources.

Bickerton's enterprise is limited by the sketchiness of the available data for the earliest stages of creole languages, but the view that new languages emerge rapidly and fully formed despite very impoverished input receives striking support from work on signed languages. The critical fact here is that only about 10% of deaf children in the US are born to deaf parents who can provide early exposure to a conventional sign language. This means that the vast majority of deaf children are exposed initially to fragmentary signed systems which have not been internalized well by their primary models. This is often some form of Manually Coded English (MCE), which maps English into a visual/gestural modality. Goldin-Meadow & Mylander (1990) take these to be artificial systems and they show how deaf children go beyond their models in such circumstances and "naturalize" the system, altering the code and inventing new forms which are more consistent with what one finds in natural languages. Supalla (1990) casts more light on this, showing that MCE morphology fails to be attained well by children, who fail to use many of the markers that they are exposed to and use other markers quite differently from their models. He focusses on deaf children who are exposed only to MCE with no access to American Sign Language (ASL), and he found that they restructure MCE morphology into a new system. Clearly this cannot be modeled by input-matching learning devices, because the input is not matched. Not even close. Furthermore, it is not enough to say that MCE morphology simply violates UG constraints, because that would not account for the way in which children devise new forms. More is needed from UG. The unlearnability of the MCE morphology suggests that children are cue-based learners, programmed to scan for clitic-like, unstressed, highly assimilable inflectional markers. That is what they find standardly in spoken languages and in natural signed languages like ASL. If the input fails to provide such markers, then appropriate markers are invented; children seize appropriate kinds of elements which can be interpreted as inflectional markers. The acquisition of signed languages under these circumstances offers an opportunity to understand more about abrupt language change, creolization and about cue-based acquisition (Lightfoot 1998). One particular case of great interest in this regard is the emergence of Nicaraguan Sign Language, as described by Kegl, Senghas & Coppola (1998).

The characterization of abrupt grammatical change sketched here makes sense only if one views grammars as individual mental entities, and not as some kind of social entity codifying the data attested in the texts of some period. Failure to make this simple distinction has entailed confusion in the literature, discussed in Lightfoot (1995). Of course, one can talk about the social distribution of these grammars. There has been interesting work on the replacement of one grammar by another, i.e. the spread of change through a speech community. So, Kroch and his associates (Kroch 1989; Kroch & Taylor 1997; Pintzuk 1990; Santorini 1992, 1993, Taylor 1990) have argued for coexisting grammars. That work postulates that speakers may operate with more than one grammar in a kind of "internalized diglossia" and it enriches grammatical analyses by seeking to describe the variability of individual texts and the spread of a grammatical change through a population.

Niyogi & Berwick (1995) have recently offered a population genetics computer model for describing the spread of new grammars. It is generally agreed that certain changes progress in a S-curve but now Niyogi & Berwick provide a model of the emergent, global population behavior, which **derives** the S-curve. They postulate a learning theory and a population of child learners, a small number of whom fail to converge on preexisting grammars, and they produce a plausible model of population changes for the loss of null subjects in French. The fact that changes can be shown to progress through populations in a S-curve is not surprising to those who have written about chaotic systems and catastrophic changes (Lightfoot 1991, ch.7), but the success of Niyogi & Berwick is to show that it is not impossibly difficult to compute (or simulate) grammatical dynamical systems; they show explicitly how to transform parameterized theories and memoryless learning algorithms to dynamical systems, producing results along the way.

However, the approach sketched here runs counter to three other pervasive lines of thought. One is the idea that all change is gradual and that abrupt, catastrophic change does not happen (Harris & Campbell 1995; Hopper & Traugott 1993; Carden & Stewart 1988). This is sometimes modeled in "lexicalist" theories of grammar, in which particular grammars differ from each other in terms of features on individual lexical items (see Lightfoot 1991; ch.6 for disussion). This approach to change implies that language acquisition is data-driven, that children match their input, which may vary without limit. Where children appear not to match their input, it is claimed that access to more complete data would reveal that abrupt transitions do not happen. Of course, in dealing with historical texts, one is dealing with performance data which do not match grammars perfectly, least of all single grammars. This means that grammarians must interpret the data and each interpretation must find the most appropriate level of abstraction. For example, Fries (1940) offered statistical data showing that Old English alternated between object–verb and verb–object order freely and that "the order of … words … has no bearing whatever upon the grammatical relationships involved" (p.199). He found that object–verb order occurred 53% of the time around the year 1000 and that it was "gradually" replaced by verb–object order, reducing to 2% by the year 1500. However, his counts ignored the distinction between matrix and embedded clauses and he had no analysis of verb-second effects. If one makes such distinctions, one can show that Old English grammars most typically had object–verb order underlyingly and an operation of verb movement raising finite verbs to C in matrix clauses to yield verb-second order (van Kemenade 1987). Kroch & Taylor (1997) show that there was a dialect difference involving movement of finite verbs to C, and consequently the **grammatical** change consisted in a change in the order of the verb and its complement and in the loss of verb-second grammars, each of which were catastrophic (Lightfoot 1997a).

A second incompatible line of thought is that there exists a theory of change with some content. If one has a theory of grammar and a theory of acquisition, it is quite unclear what a theory of change is supposed to be a theory of. Presumably a "theory of grammaticalization" (Hopper & Traugott 1993) is a sub-part of such a theory of change, insofar as it involves a claim that there is more grammaticalization over time.

A third approach with which I take issue is the tendency to incorporate historicist elements into UG. Keyser & O'Neil (1985) propose a condition that "whenever possible the language acquisition device reduces the level of optionality, either by change of status or rule loss"; their evidence comes from changes which they analyze as the loss of optional rules. Similarly, Bauer (1995) construes Latin as a thorough-going left branching (LB) language which changes into a thorough-going right branching language (French). She explains this on the grounds that LB languages (with non-agglutinating morphology) were hard to acquire: "Latin must have been a difficult language to master, and one understands why this type of language represents a temporary stage in linguistic development" (p.188). So she explains her change not in a mysterious theory of history, but rather in terms of human biology: our brains work in such a way that complex structures in LB languages without agglutinative morphology are hard to acquire. This, of course, immediately raises the question of why early Latin would have been LB: "If left-branching structures are … acquired with greater difficulty, it is indeed legitimate to wonder why languages, in an early period, exhibit this kind of structure" (p.216). She concludes that this "still remains to be explained" (p.217). In the same vein, Kiparsky (1997) appeals to "endogenous optimization" and Roberts (1993b) builds a weighting into UG so that UG effectively encourages learners to "grammaticalize" independently of what they experience through their PLD; this is said to promote Diachronic Reanalyses. Historical linguists often see general directions to change and they explain this either by invoking laws of history (i.e. a "theory of change;" see Lightfoot 1979) or by attributing historical effects to genetic predispositions. So Keyser & O'Neil (1985) build a clause into UG predisposing us against optional rules. But for optional rules to be lost, they must first be introduced; if we are predisposed not to attain optional rules, one wonders how they would be triggered in the first place. The identical point holds of the inbuilt tendencies to branch to the right, to "optimize", and to grammaticalize.

These ideas reflect nineteenth-century efforts to find deterministic laws of history and the view that historical change is principled and law-governed. Historians are attracted to this view when they focus attention on changes which they believe recur in the history of many languages (see Lightfoot, in press, for discussion).

Rather, one needs a more contingent approach: two people attain different grammars only if exposed to triggers which differ in some relevant way, and therefore grammatical shifts are to be explained **only** by a prior change in the trigger experience. Language acquisition takes place by an interaction of UG, the trigger experience, and nothing else. Our model embodies this kind of contingency and characterizes change as chaotic and flukey. Change is chaotic and flukey, but nonetheless explainable to a degree.

4. Conclusion

I submit that work on abrupt creolization, the acquisition of signed languages, and on catastrophic historical change shows that children sometimes do not match their input. This work invites us to think of acquisition as cue-based; children scan the environment for certain elements of I-language in unembedded domains. These elements are not in the input directly, but they are derived from the input, in the mental representations yielded as children understand

and "parse" their input. So a cue-based learner acquires a verb-second grammar not by evaluating grammars against sets of sentences but on exposure to structures commencing $_{\text{SpecCP}}$[XP].

The name is new but the cue-based approach to acquisition is assumed in earlier work, as we noted, and it comports well with work on the visual system, which develops as organisms are exposed to very specific visual stimuli, horizontal lines for example (Hubel 1978; Hubel & Wiesel 1962; Sperry 1968). Current theories of the immune system are similar; specific antigens amplify pre-existing antibodies. In fact, this is the kind of thing which is typical of selective learning quite generally (Piattelli-Palmarini 1986).

Cue-based acquisition is a radical departure from much current work on learnability, which postulates various forms of input matching. It is striking that so much of this work has children dealing with elements of E-language, often requiring that the system perform elaborate calculations, in effect. For example, one of the best known results of work on learnability, the Subset Principle of Berwick (1985), is usually construed as calculating subset relations among sets of E-language and choosing among grammars accordingly. Dresher & Kaye (1990) show that the Subset Principle can be defined intensionally with respect to cues. The model advocated here plays down the centrality of E-language, and postulates children seeking elements of I-language in the input and selecting grammars accordingly; the model makes no reference to elements of E-language or to the output of the grammar.

The cue-based approach assumes with Lightfoot (1989) that there is a learning path, an order in which parameters are set. We have seen that a child cannot determine whether Specifiers precede heads until some analytical vocabulary has been developed. Similarly, the child cannot determine whether SpecCP is filled (in a V2 language) until she has identified phrasal categories, learned that initial categories do not have any fixed grammatical or thematic role and (therefore) are followed directly by a finite verb. All of this represents prior stages of acquisition. Representations are elaborated step-by-step in the course of acquisition, and the cues needed become increasingly abstract and grammar-internal. In this model the learning path is part of linguistic theory, a function of the way in which the parameters and their cues are stated.

Consequently, there may be no general learning algorithm distinct from the content of the grammar, along the lines of Gibson & Wexler's TLA or Clark's genetic algorithms, which are learning algorithms quite distinct from the grammars assumed. Dresher (1997) argues that the cue-based acquisition strategy is "deterministic" in the sense of Berwick (1985), in that the learner may not backtrack or undo parameter settings that have already been set.

Under this view, one would expect there to be grammatical changes which are abrupt, and one would expect languages to differ from each other in bumpy ways. We may seek to quantify the degree to which cues are represented in the PLD, showing that abrupt, catastrophic change takes place when those cues are expressed below some threshold of robustness. This enables us to avoid the circularity of historical explanations which, for example, attribute the loss of verb-second grammars in Middle French to the prior introduction of *XP subject V …* forms.

If we produce productive models for historical change along these lines, relating changes in simple cues to large-scale grammatical shifts, our results will have consequences for the way in which we study language acquisition. In particular, we shall not be surprised that changes sometimes occur abruptly. With the development of computer corpora, Niyogi and Berwick's results, and an explicit cue-based theory of acquisition, we have all the ingredients for success in the historical domain, synthesizing work on language change, acquisition, and variation.

For the bibliography, please see page 8.

CONFERENCE CALLS

Lexical Functional Grammar Conference

The University of Queensland, Brisbane, Australia.
30 June–2 July 1998.

The conference aims to promote interaction and collaboration among researchers interested in nonderivational approaches to grammar, where grammar is seen as the interaction of constraints from multiple levels, including category information, grammatical relations, and semantic information.

Deadline: 30 January 1998.
Info: thking@parc.xerox.com (Tracy King);
http://www.sultry.arts.usyd.edu.au/LFG98/

Chicago Linguistic Society 34

The University of Chicago, USA
17–19 April, 1989

The organizers invite original, unpublished work on any topic of general linguistic interest. Invited speakers: Diana Archangeli, David Dowty. Panel topics: The Status of Constraints in Linguistic Theory, The Acquisition of Spoken Language, Language Acquisition and the Lexicon.

Deadline: 31 January, 1998.
Info: cls@tuna.chicago.edu;
http://humanities.uchicago.edu/humanities/cls/

1998 Henry Sweet Society Colloquium

Amsterdam, The Netherlands
16–19 September, 1998

Papers (20 mins) are invited on any aspect of the history of linguistics.

Deadline: 31 January, 1998.
Info: rdsmith@rullet.leidenuniv.nl (Robin Smith).

European Summer School in Logic, Language, and Information Student Session

University of Saarbrücken, Germany
17–28 August, 1998

Two talks a day for two weeks in a row for student participants of ESSLLI 98. For undergraduate and graduate students only. Subject areas include: Linguistics, Logic & Linguistics, Linguistics & Computation.

Deadline: 15 February, 1998.
Info: korbay@ufal.mff.cuni.cz (Ivana Kruijff-Korbayova);
http://www.coli.uni-sb.de/esslli/

Joint Conference on Formal Grammar, Head Driven Phrase Structure Grammar, and Categorial Grammar

University of Saarbrücken, Germany
14–16 August, 1998

This conference combines the 4th conference on Formal Grammar and the 5th conference on HPSG. Themes of interest: formal and computational syntax, semantics, and pragmatics; HPSG and Categorial Grammar; model-theoretic and proof-theoretic methods in linguistics; constraint-based and resource-sensitive approaches to grammar; foundational, methodological, and architectural issues in grammar.

Deadline: 15 March, 1998.
Info: http://www.dfki.de/events/hpsg98/

Effects of Morphological Case

Utrecht University, The Netherlands
28–29 August, 1998

This workshop aims to bring together theoretical and empirical considerations on the effects of morphological case. Issues for consideration include the implications of morphological case for abstract case theory, the difference between structural and inherent case, agreement, word order phenomena, grammaticalization processes, discourse theory, and semantics.

Deadline: 1 April, 1998.
Info: koeneman@let.ruu.nl (Olaf Koeneman).

Second Language Research Forum '98

University of Hawaii at Manoa, USA
15–18 October, 1998

Title: Complementary Perspectives on Second Language Research.

Deadline: 15 April, 1998
Info: slrf98@hawaii.edu;
http://www.lll.hawaii.edu/slrf98

European Second Language Association (EUROSLA) 8

The British Institute in Paris/University of Paris VIII, France
10–12 September, 1998

Main theme: the bilingual individual.

Deadline: 23 April, 1998.
Info: Buxton@ext.jussieu.fr (Candi Buxton);
http://www.bip.lon.ac.uk/eurosla8

Western Conference on Linguistics '98

Arizona State University, Tempe, Arizona, USA
9–11 October, 1998

Abstracts (20 mins) are invited in all areas of linguistic theory. Invited speaker: Jim McCloskey.

Deadline: 1 June, 1998.
Info: Teresa.Wells@asu.edu

A Grammatical Development Bibliography

Baker, M. (1996). *The polysynthesis parameter*. Oxford: Oxford University Press.

Battye, A. & I.G. Roberts, eds. (1995). *Clause structure and language change*. Oxford: Oxford University Press.

Bauer, B. (1995). *The emergence and development of SVO patterning in Latin and French*. Oxford: Oxford University Press.

Berwick, R.C. (1985). *The acquisition of syntactic knowledge*. Cambridge, Mass.: MIT Press.

Bickerton, D. (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences* **7,** 173–222.

Bickerton, D. (1998). How to acquire language without positive evidence: What acquisitionists can learn from creoles. In DeGraff, ed.

Carden, G. & W.A. Stewart (1988). Binding theory, bioprogram and creolization: Evidence from Haitian Creole. *Journal of Pidgin and Creole Languages* **3,** 1–67.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.

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

Clark, R. (1992). The selection of syntactic knowledge. *Language Acquisition* **2.**

Clark, R. & I.G. Roberts (1993). A computational approach to language learnability and language change. *Linguistic Inquiry* **24,** 299–345.

DeGraff, M. ed. (1998). *Language creation and change: Creolization, diachrony and development*. Cambridge, Mass.: MIT Press.

Dresher, B.E. (1997). Charting the learning path: Cues to parameter setting. To appear in *Linguistic Inquiry*.

Dresher, B.E. & J. Kaye (1990). A computational learning model for metrical phonology. *Cognition* 137–195.

Ellegård, A. (1953). *The auxiliary DO: The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.

Emonds, J. (1978). The verbal complex V′- V in French. *Linguistic Inquiry* **9,** 49–77.

Faarlund, J.T. (1990). *Syntactic change: Toward a theory of historical syntax*. Berlin: Mouton de Gruyter.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. Diss., Leiden University. The Hague: Holland Academic Graphics.

Fikkert, P. (1995). Models of acquisition: How to acquire stress. In J. Beckman, ed., *Proceedings of NELS 25*. GLSA, University of Massachusetts.

Fodor, J.D. (1998). Unambiguous triggers. *Linguistic Inquiry* **29**.

Fontana, J.M. (1993). *Phrase structure and the syntax of clitics in the history of Spanish.* Diss., University of Pennsylvania.

Fries, C. (1940). On the development of the structural use of word-order in Modern English. *Language* **16,** 199–208.

Gibson, E. & K. Wexler (1994). Triggers. *Linguistic Inquiry* **25.**

Goldin-Meadow, S. & C. Mylander (1990). Beyond the input given: The child's role in the acquisition of language. *Language* **66,** 323–355.

Harris, A. & L. Campbell (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Holmberg, A. (1986). *Word order and syntactic features*. Diss., University of Stockholm.

Hopper, P. & E. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Hubel D. (1978). Vision and the brain. *Bulletin of the American Academy of Arts & Sciences* **31,** no.7:28.

Hubel, D. & T. Wiesel (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology* **160,** 106–154.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Foris.

Hyams, N. (1996). The underspecification of functional categories in early grammar. In H. Clahsen, ed., *Generative perspectives on language acquisition*. Amsterdam: John Benjamins.

Kegl, J., A. Senghas & M. Coppola (1998). Creation through contact: Sign language emergence and sign language change in Nicaragua. In DeGraff, ed.

Kemenade, A. van (1987). *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.

Kemenade, A. van & N. Vincent, eds. (1997). *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press.

Keyser, S.J. & W. O'Neil (1985). *Rule generalization and optionality in language change*. Dordrecht: Foris.

Kiparsky, P. (1997). The rise of positional licensing. In Van Kemenade & Vincent, eds.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Journal of Language Variation and Change* **1,** 199–244.

Kroch, A. & A. Taylor (1997). Verb movement in Old and Middle English: Dialect variation and language contact. In Van Kemenade & Vincent, eds.

Lasnik, H. (1995). *Verbal morphology: Syntactic structures meets the Minimalist Program.* Ms., University of Connecticut.

Lightfoot, D.W. (1979). *Principles of diachronic syntax.* Cambridge: Cambridge University Press.

Lightfoot, D.W. (1989). The child's trigger experience: Degree-0 learnability. *Behavioral and Brain Sciences* **12,** 321–334.

Lightfoot, D.W. (1991). *How to set parameters: Arguments from language change*. Cambridge, Mass.: MIT Press.

Lightfoot, D.W. (1993). Why UG needs a learning theory: Triggering verb movement. In C. Jones, ed., *Historical linguistics: Problems and perspectives,* 190–214. London: Longman. [Reprinted in Battye & Roberts, eds.].

Lightfoot, D.W. (1994). Degree-0 learnability. In B. Lust, G. Hermon & J. Kornfilt, eds., *Syntactic theory and first language acquisition.* Vol. 2: *Crosslinguistic perspectives*. Hillsdale, N.J.: Lawrence Erlbaum.

Lightfoot, D.W. (1995). Grammars for people. *Journal of Linguistics* **31,** 393–399.

Lightfoot, D.W. (1996). Review of Bauer 1995. *Language* **72,** 156–159.

Lightfoot, D.W. (1997a). Catastrophic change and learning theory. *Lingua* **100,** 171–192.

Lightfoot, D.W. (1997b). *Shifting triggers and diachronic reanalyses.* In Van Kemenade & Vincent, eds.

Lightfoot, D.W. (1998). *Creoles and cues.* In DeGraff, ed.

Lightfoot, D.W. (in press). *The development of language: Acquisition, change, and evolution*. Oxford: Blackwell.

Lightfoot, D.W. & N. Hornstein, eds. (1994). *Verb movement*. Cambridge: Cambridge University Press.

MacWhinney, B. & E. Bates, eds. (1989). *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

Newport, E.L. (1998). *Reduced input in the acquisition of signed languages: Contributions to the study of creolization.* In DeGraff, ed.

Newport, E., H. Gleitman & L.R. Gleitman (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow & C. Ferguson, eds. *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

Niyogi, P & R.C. Berwick (1995). The logical problem of language change. *MIT, A.I. Memo* No. **1516.**

Pearce, E. (1990). *Parameters in Old French syntax*. Dordrecht: Kluwer.

Piattelli-Palmarini, M. (1986). The rise of selective theories: A case study and some lessons from immunology. In W. Demopoulos & A. Marras, eds. *Language learning and concept acquisition: Foundational issues*. Norwood, N.J.: Ablex.

Pintzuk, S. (1990). *Phrase structures in competition: Variation and change in Old English word order*. Diss., University of Pennsylvania.

Platzack, C. (1986). Comp, Infl, and Germanic word order. In L. Hellan & K. Koch Christensen, eds. *Topics in Scandinavian syntax*. Dordrecht: Reidel.

Pollock, J.-Y. (1989). Verb movement, UG and the structure of IP. *Linguistic Inquiry* **20,** 365–424.

Richards, B.J. (1990). *Language development and individual differences: A study of auxiliary verb learning*. Cambridge: Cambridge University Press.

Roberts, I.G. (1985). Agreement patterns and the development of the English modal auxiliaries. *Natural Language and Linguistic Theory* **3,** 21–58.

Roberts, I.G. (1993a). *Verbs and diachronic syntax*. Dordrecht: Kluwer.

Roberts, I.G. (1993b). A formal account of grammaticalization in the history of Romance futures. *Folia Linguistica Historica* **13,** 219–258.

Roberts, I.G. (1994). Two types of head movement in Romance. In Lightfoot & Hornstein, eds.

Roberts, I.G. (1998). *Verb movement and markedness.* In DeGraff, ed.

Rohrbacher, B. (1994). *The Germanic VO languages and the full paradigm: A theory of V to I raising*. Diss., University of Massachusetts, Amherst.

Santorini, B. (1992). Variation and change in Yiddish subordinate clause word order. *Natural Language and Linguistic Theory* **10,** 595–640.

Santorini, B. (1993). The rate of phrase structure change in the history of Yiddish. *Journal of Language Variation and Change* **5,** 257–283.

Sperry, R. (1968). Plasticity of neural maturation. *Developmental Biology Supplement* **2,** 306–327.

Sprouse, R. & B. Vance (1998). An explanation for the loss of null subjects in certain Romance and Germanic languages. In DeGraff, ed.

Supalla, S. (1990). *Segmentation of manually coded English: Problems in the mapping of English in the visual/gestural mode.* Diss., University of Illinois.

Taylor, A. (1990). *Clitics and configurationality in Ancient Greek*. Diss., University of Pennsylvania.

Thomason, S. & T. Kaufman (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of Calirofnia Press.

Tieken-Boon van Ostade, I. (1987). *The auxiliary DO in Eighteenth-century English: A socio-historical-linguistic approach*. Dordrecht: Foris.

Vance, B. (1995). On the decline of verb movement to Comp in Old and Middle French. In Battye & Roberts, eds.

Vikner, S. (1994). Finite verb movement in Scandinavian embedded clauses. In Lightfoot & Hornstein, eds.

Warner, A.R. (1983). Review article on Lightfoot 1979. *Journal of Linguistics* **19,** 187–209.

Warner, A.R. (1993). *English auxiliaries: Structure and history*. Cambridge: Cambridge University Press.

Warner, A.R. (1995). Predicting the progressive passive: Parametric change within a lexicalist framework. *Language* **71,** 533–557.

Warner, A.R. (1997). The structure of parametric change, and V movement in the history of English. In Van Kemenade & Vincent, eds.

DOES CHOMSKY EXIST?

Neil Smith

Let me reassure those of a nervous disposition that the answer we shall arrive at is, probably "yes". As most people reading this will never have doubted the fact, it may seem curious that I pose the question at all, especially as there is a common myth (1994b, p.170) that there are two Noam Chomskys: the political activist and the linguist. I have even interviewed prospective students for a linguistics degree who expressed mild disbelief when I assured them that the author of *Deterring Democracy* and *Year 501* had made a not insignificant contribution to the subject they were blithely preparing to devote the next three years of their lives to.

The question is prompted by puzzles about the relation between words and things, between thoughts and states of affairs. The accepted wisdom is that our utterances reflect our beliefs, our beliefs are made true or false by facts that exist independently of us, and hence we may perhaps not be the ultimate authority on what our own words mean. The conclusion is discomfiting.

In his discussion of reference, Chomsky is wont to make remarks about "there being no external reality" (1993, p.43), to say that "The question, 'to what does the word X refer?', has no clear sense" (forthcoming, p.14), and that relating linguistic mental representations to things in the world is not simple and "perhaps even a misconceived project" (1994, p.159). If there is no external reality, then there is no Chomsky (and you're not reading this, either). But this sceptical conclusion, while immune to refutation, is neither fruitful nor interesting, and is clearly not what the putative Chomsky had in mind.

Rather, he is calling into question the traditional philosophical treatment of semantics in terms of a systematic word/world relation such that the word 'red' refers to red things, 'elm' refers to elm trees, 'water' refers to H_2O, and so on. Chomsky (1995, etc.) argues at length that, although we can refer by using words, words themselves do not refer. "**People** use words to refer to things in complex ways, reflecting interests and circumstances, but the **words** do not refer" (1996, p.22).

The idea that words refer to things, and sentences refer to states of affairs, in the world is deeply entrenched, but misguided. The source of the misapprehension is the assumption that meaning is necessarily to be described in terms of truth, that is, the meaning of a sentence is given by its truth conditions: the set of conditions whose fulfilment would guarantee the truth of the sentence. At first blush this idea is not entirely silly: if someone asks whether the assertion that 'Scandinavians are dolichocephalic' is true, then a prerequisite to being able to answer the question is knowing what the word means ('long-headed'). That is, to say that *'Scandinavians are dolichocephalic' is true* is equivalent to saying that *'Scandinavians are long-headed' is true*, or 'Scandinavians are dolichocephalic' means that Scandinavians are long-headed. In such cases truth and meaning seem to be inter-convertible in a reasonably simple way, but there are serious problems. These problems fall into different classes and have prompted a panoply of sophisticated accretions to truth-conditional theories of meaning. Chomsky's solution is more radical: to give up on truth as relevant for I-language. Let us look first at some of the reasons for denying the word-world relation, and then at what might constitute an alternative.

One of Chomsky's favourite examples (e.g. 1993, p.22; 1995, p.21; 1996, p.22) is 'London'. It seems straightforward to say that the word 'London' refers to the city which is the capital of the UK. But Chomsky draws a distinction between our use of the word to refer to London: something which is unproblematic, and the unjustified further claim that, because we can use the word to refer, the word itself refers to some real entity in the world. London might be destroyed and rebuilt elsewhere, but it would still be London; I can refer to the same entity by using terms like 'the city where I work', 'the great wen', or indefinitely many others; or I can use the word 'London' with no referential intention at all as when I say in exasperation "Look, either you work in London or you don't work in London, there are no other logical possibilities". Thus his position is ultimately a reflection of a disagreement about the nature of semantics. The philosophical tradition is wedded to the assumption, usually associated with the nineteenth century philosopher Frege, that it is possible to assign truth conditions to every possible (assertoric) utterance. The alternative internalist view is that semantic relations are nothing to do with things in the world but are relations between mental representations: they are entirely inside the head. It is no more appropriate to assume that 'London' picks out a specific entity in the world, than it is to postulate a real world entity of 'nearness' which is possessed to a greater or less extent by various objects. Whether London is 'near' Birmingham is not something which is determined by facts in the real world, but by our perspective on those facts. In the context of European cartography, London and Birmingham are 'near'; in the context of the Ramblers' Association they are not. No-one has suggested that 'nearness' should be assigned the status of an externally valid absolute; Chomsky's point is that standard philosophical assumptions about language and reference are equivalent to assigning that status — unjustifiably — to language: in particular, the assumption that there are common meanings expressed in a public language (E-language) is at the root of a deep confusion.

There is an intuitive appeal to the notion that there is an external language that different people speak. Indeed, it is so self-evidently true that it would be pointless to deny it. However, the idea turns out to be problematic, as the notion of 'language' involved is different from the notion that linguists theorise about. Why? First, we all differ from each other in vocabulary: you may not have known the word 'dolichocephalic' I used earlier, and you certainly know words I don't. Within an externalist framework, this problem is not difficult to accommodate: we each know a subset of the words of the language we speak, where the totality of that vocabulary is encapsulated in major dictionaries. Matters are (supposedly) not so easy for internalists, where there is no external standard provided by a public language. Problematic examples are provided by disagreements over meaning of the kind seen in Burge's 'arthritis' (1992, p.217) or Putnam's twin earth 'water'. It is supposed to be necessary to a correct understanding of 'arthritis' that there be social agreement as to what the word "really" means; hence there has to be a public language code in which this meaning is represented. But first this external code doesn't obviously solve the problem of partial communication; and second there is no problem in making the simple internalist assumption that your interlocutors are like you until you get evidence that this is false, and then you modify your interpretation of their utterances accordingly. If you tell me that you have arthritis in your cheek, I shall think you ignorant of standard medical usage and interpret your hypochondria accordingly. No more needs to be said.

Similarly with twin-earth, or the problems of Pierre who is confused about whether London (which he considers ugly) is the same as Londres (which he considers beautiful). If Oscar is transported to Putnam's twin earth and has a drink from the tap, believing it to be H_2O when in fact it is XYZ, he is mistaken, but nothing of import follows about the nature of language.

Chomsky's scepticism about "external reality" is then best interpreted as an affirmation of the discontinuity between two perspectives on language. Specifically, his preoccupation with developing a scientific theory of the language faculty — and his ignoring for the most part of how that faculty is put to use in comprehension, and of how it is used to refer, for instance, leads him to believe that there is no role for a theory of reference **within language:** "It is possible that natural language has only syntax and pragmatics" (1995, p.26). This does not deny that humans use language to refer, but referring crucially involves going outside the language faculty. Answering the question we started with ultimately requires an understanding of Chomsky's contribution to epistemology: knowledge of language is not "grounded" in the way conscious knowledge is traditionally supposed to be, it is there as part of our make-up as humans.

Chomsky exists all right, but 'Chomsky' doesn't refer to him.

References

(All references are to Chomsky …)
(1992). Explaining language use. *Philosophical topics.* **20,** 205–231.
(1993). Mental constructions and social reality. In E. Reuland & W. Abraham (eds.), *Knowledge and language.* Vol. 1, *From Orwell's problem to Plato's problem,* 29–58. Dordrecht: Kluwer.
(1994a). Chomsky, Noam. In S. Guttenplan (ed.), *A companion to the philosophy of mind*, 153–167. Oxford: Blackwell.
(1994b). *Keeping the rabble in line: Interviews with David Barsamian.* Edinburgh: AK Press.
(1995). Language and nature. *Mind* **104,** 1–61.
(1996). *Powers and prospects: Reflections on human nature and the social order.* London: Pluto Press.
(forthcoming). Internalist explorations. In M. Hahn & B. Ramberg (eds.), *Festschrift for Tyler Burge*. Cambridge, Mass.: MIT Press.

THE SYNTAX OF INTERPRETATION

by *Sjef Barbiers*
Reviewed by *Heidi Harley*

Summary by *the author*

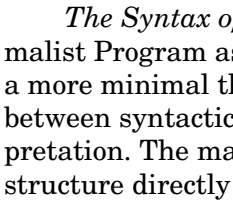
1. The hypothesis

If economy principles are fundamental to language design (Chomsky 1993) and if the sole goal of building a syntactic structure is to construct a complex meaning by combining the simplex meanings of individual morphemes, then each and every step of a syntactic derivation should make a direct contribution to the semantic interpretation of the resulting syntactic structure. The Minimalist view on language is not maximally economical in this sense. First, syntactic movement is driven by the need to check morphological features that are irrelevant for semantic interpretation. Secondly, the syntactic relations established by Merge, as expressed in an X-bar structure, do not directly correspond to semantic relations. Rather, certain syntactic relations such as sisterhood function as a condition on semantic relations, that are established by lexical properties of morphemes in the tree, e.g. by θ-role assignment.

The Syntax of Interpretation takes the Minimalist Program as its starting point and develops a more minimal theory of the correspondence between syntactic structure and semantic interpretation. The main hypothesis is that X-bar structure directly establishes the semantic relations between nodes in a tree, independently from the lexical content of the terminals that are present in the tree. At LF, the semantic relations established by X-bar structure must be fully compatible with the lexical relational information associated with the morphemes in the tree. If not, the structure is not fully interpretable. The proposed theory is minimal in that it assumes just one, conceptually necessary, primitive syntactic relation, corresponding to just one semantic relation. The semantic relations between a verb and its arguments or between an adjunct and the constituent that it modifies can all be derived from this one primitive relation. As such, the theory is a step towards eliminating θ-theory. An outline of the proposed theory is given in chapter 1.

The primitive syntactic and semantic relation assumed in the proposed theory is given in (1a). The structures in (2) are examples of this syntactic and semantic atom.

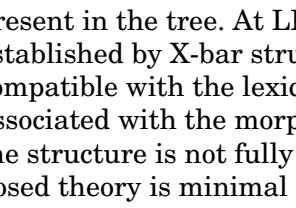
(1)



'there is a z-relation between x and y'

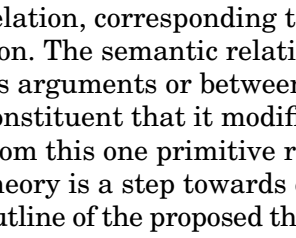
(2)

a.



'there is an *after*-relation between John and Mary'

b.



#'there is a *green*-relation between John and Mary'

The main assumption is that a primitive relation consists of three elements: the relator and the two objects of which the relation holds. What is new in this proposal is that in the primitive relation R(X,Y) not only X and Y are nodes, but R is a node as well. Whereas more traditional theories have relations such as *modifier_of*(PP,VP) or *theme_of*(DP,V), the present proposal dispenses with such additional relations, using only relations established directly by X-bar structure.

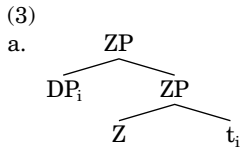
The examples in (2) demonstrate what is meant by "the semantic relations established by X-bar structure must be fully compatible with the lexical relational information associated with the morphemes in the structure". The P *after* in (2a) denotes a relational concept and can occur in the indicated syntactic position. The A *green* in (2b) however does not denote a relational concept, and therefore the atom in (2b) is not fully interpretable.

Although (2a) is fully interpretable, it does not give us the full interpretation of the string *John after Mary*, since the interpretation of this string is asymmetric. The string means 'John is after Mary', not 'Mary is after John'. This is an instance of a more general problem: how does a theory that has only one primitive relation, which consists of three nodes, express those semantic relations that seem to consist of two nodes? That is, how can the present theory express the semantic relation between an adverbial PP and a VP, or between an adjective and a noun, or between a small clause predicate and its subject?

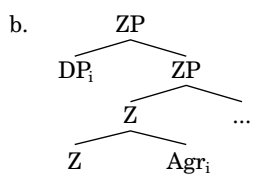
A two-element relation can be expressed in terms of a three-element relation by making the two objects of which the relation holds identical. As shown in (3), there are two ways of doing this: by movement or by agreement.

(3)

a.



b.



In (3a), the objects of the relation are identical: there is a trace or copy in the complement position that has the same index as the DP in the specifier position. In (3b), the lower Z establishes a relation between DP in spec and Agr. Since Agr has the same (ϕ-features as DP, it is reasonable to consider it to have the same index as the DP in specifier position. The configuration in (3a) is central in The Syntax of Interpretation. It plays a crucial role in the analysis of focus particles, PP Extraposition, scrambling and modal verbs. The configuration in (3b) opens an interesting perspective for an explanation of the role of Agreement in natural language. The claim is here that Agreement is a [+interpretable] feature. For empirical domains in which (3b) is relevant, see Barbiers (1997). The definition of the correspondence between syntactic structure and semantic interpretation is given in (4).

(4) Principle of Semantic Interpretation (PSI)

I. A node Z establishes a Semantic Relation between a node X and a node Y iff X immediately c-commands Z and Z immediately c-commands Y.

II. A node Z establishes a Qualification Relation between a node X and a node Y iff Z establishes a Semantic Relation between X and Y, and X and Y are coindexed.

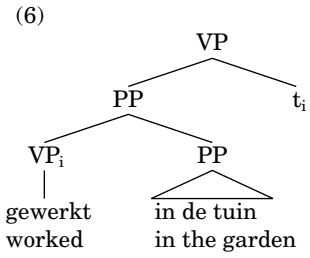
The notion of Qualification Relation subsumes semantic relations such as the relation between a small clause predicate and its subject and between a modifier and its modifiee; roughly, those semantic relations that fit into the frame "X is Z", with Z the predicate.

2. Implementation

In chapter 2, the proposed theory is made technically explicit. The theory adopts many of the properties of the Antisymmetry framework (Kayne 1994). In particular, the theory does not allow rightward movement and right-adjunction and there is a direct mapping from c-command relations to linear ordering. Kayne's definition of c-command is adapted to capture c-command out of PP, to make multiple adjunction possible and to make segments visible for movement operations and for the PSI.

By this definition, movement as depicted in (6) becomes possible, while the common assumption that a moved constituent must c-command its trace can be retained. In (6), VPi immediately c-commands PP, and PP immediately c-commands ti. As a result, according to the PSI the PP *in de tuin* is interpreted as a Qualifier of the VP *gewerkt*.

(6)



In addition to this discussion of c-command and Qualification configurations, Chapter 2 offers a theoretical discussion of head movement. It is shown that the Head Movement Constraint (Travis 1984) and the Mirror Principle (Baker 1985) straightforwardly follow from the PSI. This is a very interesting result, since in this way mirror effects at the head movement level and mirror effects at the XP-movement level (e.g. PP Extraposition as VP-intraposition) receive a unified account.

3. Empirical domains

The theory is tested in three empirical domains of Dutch: scalar focus particles (Ch. 3), PPX (henceforth PPX) and scrambling into spec of adverbial (Ch. 4), and modal verbs (Ch. 5). Each chapter contains a detailed description and explanation of the semantic and syntactic properties of the phenomenon.

3.1. Scalar focus particles

In Chapter 3, it is argued that the multiple ambiguity of scalar focus particles can be explained as the result of the interaction between one simple lexical meaning and the syntactic environment of the focus particle. A scalar focus particle such as *pas* 'just' requires an argument that denotes a scale. It is shown that the kind of scale, or more specifically, the kind of partitioning of the scale, determines the interpretation of *pas*. Thus, the lexical ambiguity of *pas* mentioned in the dictionaries ('recently', 'no more than', 'not yet very long', 'no sooner than') is the result of the interaction between the simple lexical meaning 'short' or 'not long' and the argument of *pas*. For instance, when a verbal argument of *pas* has perfect tense, *pas* means 'recently', but when a verbal argument has present tense, *pas* means 'not yet very long'.

More interesting from a syntactic point of view is the numeral-associated interpretation of *pas*, 'no more than'. Ample empirical evidence is provided that to get a numeral-associated interpretation, *pas* must be base-generated in a position in which it immediately c-commands the DP or PP that contains a numeral. Such a DP/PP must move covertly or overtly to the spec of the projection of the focus particle (FocP). These properties of scalar focus particles are explained by the PSI. A scalar focus particle qualifies a constituent that denotes a part of a scale, e.g. *pas* 'just' qualifies a part of a scale as short. Since the particle is a qualifier, its argument must move into the spec of the projection of the particle to make a qualifcation relation syntactically available.

After movement to SpecFocP, the argument may move higher up, e.g. to SpecCP, leaving the focus particle stranded in its base-position. Focus particles therefore can be used as a diagnostic for movement. This diagnostic is extensively used in

Chapter 4 to investigate the properties of PPX and scrambling. In Chapter 3, the diagnostic is briefly applied to Dutch to see what it tells us about the syntactic structure of Dutch.

3.2. PPX and scrambling

The interpretation and distribution of scalar focus particles in PPX contexts strongly suggests an analysis of PPX in which adjunct-PPs are base-generated as left-adjuncts to an (extended) projection of the verb. This verbal projection intraposes by moving into the spec of the adjunct PP, as depicted in (6). This intraposition is obligatory: it must take place in overt or covert syntax to make the PP interpretable as a Qualifier of the verbal projection.

The distribution of scalar focus particles in PPX contexts is exceptional in two respects. First, a verb (and its arguments) can intervene in linear order between the particle and a postverbal numeral-containing PP without blocking the numeral-associated interpretation of the particle: Focus_Particle–Verb–PP. This seems to be an exception to the generalization that a scalar focus particle must be base-generated in a position in which it immediately c-commands the numeral-containing constituent. This exception ceases to be an exception if the analysis in (6) is adopted: the particle immediately c-commands the PP in this structure.

The second exception to the generalization is that in the linear sentence-final order Verb–Focus Particle–PP the particle cannot have a numeral-associated reading, despite the fact that it is linearly adjacent to a PP that contains a numeral. This follows from the PSI as well. For the particle to be a qualifier of the PP, the PP must move into the specifier of the projection of the focus particle. However, the given linear order can only be derived by movement of VP into the spec of the projection of the particle.Therefore this spec position is no longer available for PP. Since VP is a temporal and since the resulting structure determines that the particle is a qualifier of VP, the analysis predicts that the particle receives a temporal interpretation. This prediction turns out to be correct.

This analysis of PPX straightforwardly explains the well-known mirror effects resulting from multiple PPX (Koster 1974) and the impossibility of extraposing a small clause PP. Independent evidence for the existence of movement to the spec of an adjunct comes from a previously unattested type of scrambling in Dutch.

3.3. Modals

The four different types of interpretations that each modal in Dutch has is shown to be the result of the interaction of one basic lexical meaning for each modal and the syntactic environment of the modal. The traditional control–raising analysis of the root–epistemic ambiguity makes a number of wrong predictions and cannot explain the syntactic and semantic properties of modals. Yet, the intuition underlying this traditional analysis is partially correct: one distinguishing property of modal interpretations is the subject-orientedness of the modal.

The second distinguishing property is the nature of the complement of the modal. For the modal to get a non-epistemic interpretation, its complement must denote a polarity transition, i.e. there must be a (conceivable) stage in which the predicate in the complement of the modal does not hold of its subject and a stage in which it does. This point is demonstrated in a study of non-verbal complements of modals in Dutch. It is argued that these non-verbal complements cannot be analyzed as verbal complements with a phonologically silent verb. One strong argument is that a non-verbal complement to a modal must denote a bounded lattice, a simple generalization that cannot easily be expressed in a PF-deletion analysis of non-verbal complements.

Since θ-role assignment is not available in the proposed theory, a syntactic analysis is developed in which the semantic relation between a subject and a modal is expressed by a number of syntactic relations between the modal, a P/D like element in the sense of Kayne (1993) and the subject. This P/D-like element projects a DP. When a modal selects this DP, it can only have an epistemic interpretation. When P/D selects the projection of a modal, the modal can only have non-epistemic interpretations. In this way, the decisive influence of the complement of the modal on the interpretation of the modal is captured. Empirical evidence is provided to show that the complement of a modal behaves as a definite DP in the epistemic interpretation but as an indefinite NumP in the non-epistemic interpretations. The PSI correctly explains that the entire complement of the modal has to move into its Spec to identify the modal as a qualifier of its complement.

References

Baker, M. (1985). The Mirror Principle and morphosyntactic explanation. *Linguistic Inquiry* **16,** 373–415.

Barbiers, S. (1997). *The syntax and semantics of agreement.* Ms., MIT & HIL.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & J. Keyser (eds.), *The view from building 20: Essays in Linguistics in Honor of Silvain Bromberger*, 1–52. Cambridge, Mass.: MIT Press.

Kayne, R. (1993). Towards a modular theory of auxiliary selection. *Studia Linguistica* **47,** 3–31.

Kayne, R. (1995). *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Koster, J. (1974). Het werkwoord als spiegelcentrum. *Spektator* **3,** 601–618.

Travis, L. (1984). *Parameters and effects of word order variation.* Diss., MIT.

Review by *Heidi Harley*

The idea that the language faculty is subject to economy conditions is the starting point for a large proportion of generative syntax in these post-Minimalist days. Everyone agrees that the brain will do the easiest, least effortful (most elegant? most beautiful?) thing to bridge the sound–meaning gap and construct representations that are intelligible both to our interpretive apparatus and our mechanical apparatus (cf. Chomsky 1995). Debate, then, centers on determining what makes one representation easier for the brain than another, a notoriously difficult question.

Barbiers proposes that the goal of syntactic structure-building is the creation of a well-formed complex meaning, given a number of meaningful simplex building blocks. Hence, in his framework, every syntactic operation must be driven by purely semantic criteria. He discards entirely the idea that the creation of syntactic structure could be driven by morphological needs. PF will manage; it is LF that is fundamentally motivational. He claims this approach is more economical than an approach which assumes both PF and LF requirements may drive movement. For Barbiers, then the question then becomes, how can we characterize the relation (presumably the simplest one consistent with the evidence) between syntactic structure and semantic interpretation such that it will account for, e.g., movement of a DP to SpecTP?

Semantic relations are necessarily ternary

Barbiers' claim is that every semantic entity is a relation represented as a triple: a relator function R and two elements X and Y which the R function relates: R(X,Y). This is mapped onto the syntax via a modified notion of c-command which in the simplest case results in the head of an X-bar structure being R, its specifier being X and its complement Y. X must immediately c-command R which must immediately c-command Y (cf. Barbiers' discussion above). Movement to a specifier position will create a syntactic structure which is interpreted as a relation between the specifier and the complement of the head.

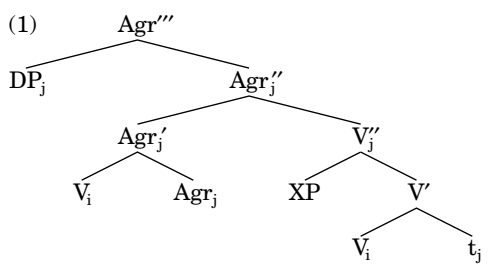
Semantic relations are often, prima facie, not triples but duples. The small clause *John tall*, for example, does not at first sight relate two entities, but merely attributes a property to a single entity. Barbiers proposes that in fact this type of relation, which he terms a Qualification relation, is a disguised triple: qualification relations are established when X and Y are coindexed — when X and Y are, in some sense, the same item. In order for *tall* to qualify *John*, a syntactic structure must be constructed which encodes the relation *tall(John, John)*. This can be done, for example, by moving *John* from the complement of *tall* to the specifier of *tall*, leaving a trace or copy of *John* behind. In Barbiers' framework, movement is always triggered by the need to establish a qualification relation — to create a binary relation from a necessarily ternary syntactic structure.

As noted above, however, the issue of "what is economical" is largely a matter of personal taste, in the absence of empirical evidence to guide the discussion. Let us examine aspects of Barbiers' case studies, which provide impressive accounts of a range of quite tangled data. In particular, Barbiers' demonstration that the varied interpretations of focus particles and modal verbs in Dutch may result solely from the nature of their complement, rather than from multiple lexical entries, is very attractive, as is his elegant account of the mirror effect in Dutch extraposed PPs. The relative economy of the approach may certainly be debatable, particularly with respect to the claim that semantic relations are necessarily ternary (why is that more economial than, e.g., a neo-Davidsonian functional application approach to semantic relations?) or its rather intricate definition of c-command, but some of the results are quite significant.

Locality of movement

Because a qualification relation is necessarily local (dependent on immediate c-command between constituents), Barbiers derives the nice result that all movement must necessarily be local, with an important exception. If mediated by an Agr projection, non-local movement may arise. Consider the illustration in (1). The head of an Agr projection is coindexed with the DP in its specifier, and if the next lowest head (e.g. V) moves up and adjoins to Agr, a qualification relation is established between the V head and the two coindexed elements, DP and Agr.

(1)



While the derivation of locality of movement is compelling, the Agr-dependent approach to locality violations seems somewhat unconstrained. So long as the projection to which movement occurs is a specifier position of an AgrP, and the next lowest head adjoins to Agr, DP-movement to SpecAgrP may be of arbitrary length; a qualification relation will be established between the moved DP and the next lowest head, whatever it is, adjoined to Agr, which is sufficient to make the triggering of movement possible. It is then unexplained why, in the languages which exhibit movement of this type, objects may not shift to high Agr projections (AgrS, for example), or why any argument DP from an embedded clause may not A-move to a matrix clause's Agr projection, skipping arbitrary numbers of specifiers en route. Solutions to this problem may certainly exist (a version of Shortest Move, e.g., or a stipulation about necessary qualification relations between the subject and the clause along the lines of the EPP) but in the absence of any proposed restriction, this is a somewhat troubling question.

The illustration in (1) also exemplifies the structure necessary to discuss Barbiers' theory of head movement, and illustrate one modification he has made to the notion of c-command. Because projections of heads bear the same index as the heads which project them, when a head moves to left-adjoin to the next head above, a qualification relation is established between the moved head and the target head — in (1), Agr relates the two nodes V_i and V''_i. V_i c-commands Agr and Agr c-commands V''_i, as there is a connected path of left branches from Agr′ to V''_i and no node intervenes between Agr′ and Agr. Heads must left-adjoin when moving, or the

left-branch condition on c-command will not be met, and obviously heads may not skip intervening heads, or no qualification relation with the complement maximal projection will be established. These requirements derive the Mirror Principle (Baker 1985) and the Head Movement Constraint (Travis 1984) respectively. (Although cf. Barbosa 1995 on the necessity of right-adjunction in the treatment of Portuguese clitics, and Carnie 1995 for right-adjunction in Irish Gaelic).

Again, this is a strong result; however, there is room for debate about whether it represents the most economical treatment of c-command. Barbiers invokes Kayne (1995) repeatedly, as his modified definition of c-command is, technically, antisymmetric. Barbiers' definition, however, is dependent upon the notion "left branch" as a primitive, and he stipulates that languages are uniformly right-branching structures. Kayne's treatment, however, is elegant precisely because it makes the claim that linear order is determined and determinable only for structures which are independently antisymmetric in their c-command relations — that is, that antisymmetry is a necessary characteristic of syntactic structure because it *enables* the determination of a left-to-right ordering of terminal nodes. Left–right relations between non-terminals are not defined and not relevant in Kayne's theory, and the right-branching nature of UG is a (nearly) derived result. Barbiers' proposal, on the other hand, achieves antisymmetry by stipulating that language is right-branching and defining c-command in terms of left branches, and it is not clear that antisymmetry is needed to determine linear order beyond the left-right notions already introduced as primitives. Much of the intuitive economy and appeal of Kayne's proposal is hence lost in Barbiers' treatment.

Scalar focus particles

The analysis Barbiers proposes of focus particles, as outlined in the summary above, seems to me to be fairly convincing. He shows that focus particles may take an argument of any category, further, he demonstrates that the interpretation of the particle depends upon the type of scale which may be associated with whatever its argument is. Of particular interest is the numeral-associated interpretation of these particles, which arises when they take a DP or PP complement with a numeral. The DP or PP which is qualified by the focus particle (by virtue of being base-generated as a complement to the particle and then moving to its specifier) may, interestingly, move away from the particle and its base position, stranding the particle. The particle thus may provide an interesting test for the base-position of arguments (much like a Japanese numeral quantifier, which exhibits interesting similarities (Miyagawa 1989), and perhaps less like an English floating quantifier, cf. Bobaljik 1995).

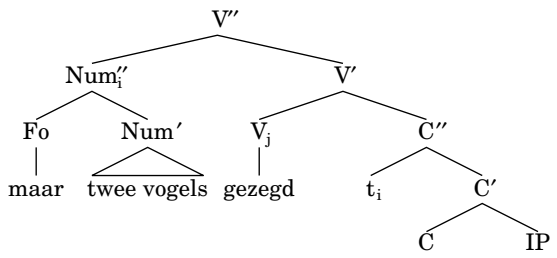
Barbiers uses this test to investigate the relative positions of subjects and objects in their scrambled and non-scrambled positions, concluding that definite objects may scramble across the base-position of the subject, as originally proposed in Chomsky (1993). In addition, he investigates the possible stranding of particles in successive-cyclic $\overline{\text{A}}$-movement, where his results indicate the viability of a Barriers-style account, forcing adjunction to the matrix VP en route to the matrix CP — a focus particle associated with the moved DP may appear in that VP-adjoined position.

Let us examine this proposal somewhat more carefully. Consider Barbiers' structure in (2b), representing the half-way point in the derivation of the sentence with the stranded focus particle in (2a), yeilding the numeral-associated reading of the stranded focus particle (I), rather than the temporal interpretation (II):

(2)

a. [$_{CP}$ TWEE vogels had Jan [$_{VP}$ maar [$_{VP}$ gezegd
two birds had Jan just said
[$_{CP}$ dat ie gezien had]
that he seen had
I: 'John had said that he had seen just two birds'
II: 'John had just said that he had seen two birds'

b.



The adjunction of the focused numeral phrase to VP creates a qualification relation between NumP and V: Num″ immediately c-commands V which immediately c-commands the trace of NumP (in SpecCP, t_i); this is essentially the interpretation which is assigned to internal arguments of V. This type of structure will arise, however, only when movement through SpecCP is independently motivated, as an argument rises to the matrix CP. Note that Num′ extracts from its position within Num″ to strand the focus particle; in Barbiers' framework, extraction from adjuncts (including extraction from specifier positions of adjuncts) is possible and necessary, as we shall see in the discussion of PP extraposition below.

The interpretation of Num″ as an internal argument is somewhat puzzling, as no verb is able to take both an internal argument and a sentential complement — the argument may not remain in situ in specVP, as Barbiers points out:

(3)

*Jan zegt TWEE vogels dat ie gezien heeft
John said two birds that he seen has.

Barbiers appeals to the analysis of Rooryck (1995), in which focus plays a crucial role in licensing the movement of arguments from embedded to matrix clauses; it is focus which will force continued movement of the *twee vogels* argument from its adjoined position to the matrix SpecCP and rule out the sentence in (3). The analysis remains somewhat murky, however, as (at least when the focus particle is stranded in the adjoined position) the trace of *twee vogels* and the Num″ projection will remain in the adjoined position. V will receive an interpretation as a qualifier, as required by the Principle of Semantic Interpretation, qualifying the Num″, but it will be remarkably different from the interpretation it receives when no argument from the embedded clause is focused. When no embedded argument is focused, in order to establish a qualification relation and hence an interpretation for V, the CP complement to V will raise to SpecVP to create the necessary triple V(CP$_i$, t$_i$). The claim is that the internal argument of the verb *say* with a CP complement is the CP when no embedded argument is focused, but in a superficially identical clause which happens to focus an embedded argument, the internal argument of *say* is the focused argument, not the CP.

Whether or not such a claim is reasonable, Barbiers does point out that this situation may only occur via this type of successive-cyclic movement via a medial SpecCP. In the unmarked case, if a V head takes both a DP internal argument and a CP internal argument — one in its specifier, one as its complement, say — no qualification relation may be established for V and the Principle of Semantic Interpretation will rule out the structure, modulo one interesting exception. If the DP and the CP are coindexed, the Principle of Semantic Interpretation will admit the structure as permitting V to qualify the coindexed arguments: V(DP$_i$, CP$_i$). This is exactly the structure of a sentence like (4):

(4) Jan zal het$_i$ betreuren [dat Piet Marie heeft ontmoet]$_i$
John will it regret that Pete Mary has met
'John will regret it that Pete met Mary'

Barbiers notes that his analysis does provide an explanation for the fact that when the coindexed pronominal is present the CP is a factive island, but when the pronominal is absent the island effect is weakened.

It is perhaps possible that Barbiers' Principle of Semantic Interpretation treatment of this type of sentence provides a nice explanation for a peculiar contrast between these sentences with and without the coindexed pronominal. Normally, the embedded CP in this type of sentence behaves as a regular constituent, permitting topicalization and clefting. However, when the coindexed pronoun is present, topicalization and clefting of the CP become impossible (cf. Bennis 1996; Vikner 1995 as cited by Brander 1995). This is illustrated in (5a,b), and is true in German and English as well as in the Dutch examples given here (Angeliek van Hout, p.c.):

(5)

a. Topicalization
[dat Piet Marie heeft ontmoet] zal Jan (*het) zijn
that Pete Mary has met will John it his
leven lang betreuren.
life long regret.
'That Pete has met Mary John will regret all his life.'

b. Clefting
Wat Jan (*het) zijn leven lang betreuren zal is
what John it his life long regret will is
dat Piet Marie heeft ontmoet
that Pete Mary has met
'What John will regret all his life is that Pete has met Mary.'

The shape of the solution that suggests itself is that Barbiers' Principle of Semantic Interpretation forces movement to be local: no move is permissible that does not establish a qualification relation. In topicalization constructions, as well as in the construction of the subject question in the cleft constructions, the complement to V (CP in (5a), *wat* in (5b)) must move from its position as complement to SpecCP. It will be unable to do so in one fell swoop, as I understand the approach, or else a qualification relation will not be established between C and the moved element in SpecCP. Movement must take place via an intermediate head, e.g. a high-up Agr head below CP (cf. Watanabe 1993). If a condition like Shortest Move is in effect in Barbiers' framework, perhaps the moving constituent will be forced through **all** the possible moves which establish its necessary qualification relations on the way to CP, being unable to skip any. If this is the case, the moving element will be forced to move through SpecVP, establishing V as its qualifier, and there will be no available site in which the pronoun *het* may appear. (Note that (5a) is good with the pronominal on a Topic–Comment type reading, with a comma intonation between the CP and the rest of the sentence; such an intonation makes a movementless derivation available.)

A final note with respect to Barbiers' adjunction structure in (2b): later in the thesis, he motivates the adoption of a split-VP verbal structure, with a non-verbal root selecting internal arguments and later adjoining to a higher VP-type projection. In order for the qualification relations to work out, the adjoined focused argument must be attached to the lower RootP, rather than the verbal vP, with the consequence that the participle *gezegd* appears overtly in the Root head. More refined tests hence should be possible with respect to the position of *maar*, it should appear (on the numeral-associated reading) to the right of vP and TP adjoined adverbs, and perhaps to the left of RootP adjoined adverbs (manner adverbs, cf. Travis 1988). I would also be interested to see these stranding tests applied to unaccusative constructions: Does *maar* appear to the right of manner adverbs when associated with a moved internal argument in a passive, e.g.? Do different readings result when stranded *maar* is on different sides of a manner adverb with a variable behavior (unaccusative with a PP, unergative without) verb like *jump (in the ditch)* (cf. Borer 1996)?

PP Extraposition and the large-scale mirror effect

Using the stranded focus particle test developed in Chapter 3, Barbiers investigates the structure of Dutch clauses with extraposed PPs, and arrives at an extremely thorough and compelling account of the legitimate PP structures. In Dutch, PPs may adjoin to VP, and adjunct PPs may appear either to the left or the right of the V. When they appear to the right of the V, in the extraposed configuration, they appear in the inverse order (given the same interpretation) that they appear in to the left of V. Barbiers invokes an unusual aspect of his c-command definition to permit movement into the

specifier of an adjunct, and derives all permissible orders, excluding all illegitimate orders, by moving the whole VP into the specifier of a PP which has adjoined to it. This results in the order VP,PP, "extraposing" the PP by moving the verb to its left, and with iteration for multiple PPs, the effect is exactly analogous to the derivation of the Mirror Principle with head movement outlined above.

This produces an unusual configuration, where a finite V head or complements of V must be allowed to be extracted from their (possibly multiply embedded) position in the specifier of an adjunct, similar to the extraction of the Num″ in the derivation of (2a) above. Again, Barbiers' definitions permit such extraction. The result (a good characterization of PP extraposition) is significant enough that unless a better alternative arises, movement to the specifier of an adjunct needs to be included in the theory, as well as extraction from such specifiers. The implications of this reformulation, however, really need to be fully worked out.

Modals, lattices and DPs

Modal verbs in Dutch have multiple interpretations, just as do modal auxiliaries in English: three subject-oriented root readings as well as the non-subject oriented epistemic reading. Barbiers argues that in fact the root/epistemic distinction cannot arise from a control/raising distinction, as there exist root readings for clauses in Dutch for which a control syntax is impossible; in fact, only a raising syntax is possible for these clauses (a modal verb with a non-verbal DP, PP or AP complement), but the epistemic interpretation is not available here. Barbiers proposes that the root/epistemic distinction with verbal complements arises from embedding the modal verb in different portions of the VP complex (providing more support for the split-VP analysis alluded to above), and the lack of such a distinction with non-verbal complements is the result of the lack of complex VP structure.

This result in itself is quite interesting. Barbiers also (almost) shows that it is not necessary to impose lexical selectional restrictions on modals to describe the distribution of the non-verbal complements. For example, a modal verb may take an adjectival complement such as *empty*, but may not take an adjectival complement such as *available* or *intelligent*; a modal verb may take a prepositional complement such as *to Mary* but not a prepositional complement such as *at home*; a modal verb may take an indefinite nominal complement like *a dog* but not a definite complement like *Mary*. Barbiers observes that the possible PP and AP complements to modal verbs all introduce a scale, a bounded lattice, which permits the hearer to deduce the direction of the change in that value implied by the modal. (e.g. *De fles moet leeg* 'The bottle must empty' implies that the current value of the emptiness scale is higher than 0, and the modal "must" denotes the requirement that the value move down the scale to zero). Thus, adjectives and PPs which do not introduce a scale may not appear as complements to modals (PPs must be directional, providing a spatial scale, rather than locative).

Barbiers extends this notion to treat modals with DP complements, but perhaps less successfully. DP complements, he observes, do not introduce a bounded lattice. Predicative DPs, in particular, are ungrammatical in this context (**Jan kan dokter*). However, as noted above, some DPs are grammatical: indefinites like *a dog* or *two things* may be complements to a modal verb like *moet*. In that context, Barbiers claims, the bounded lattice is introduced by the numerical scale which comes along with the indefinite. Definite NPs like *Mary* are grammatical with modals such as *mag* 'may' and *moet* 'must'; in that context, however, they no longer have their modal meaning but rather mean 'like' or 'admire', as you can see in his examples in (6):

(6)
a. Jan moet Marie *(niet)
John must Mary not
'John doesn't like Mary.'
b. Jan mag Marie (wel)
John may Mary well
'John likes Mary.'

Ideally, Barbiers would like to subsume these readings under his generalization about modals requiring a bounded lattice as well. He claims that the mandatory addition of the negation in (6a) and the optional addition of affirmative *wel* in (6b) serves to introduce a lattice on the negative/positive scale, permitting an interpretation of the modal. This account, however, runs into problems. First, the optionality of overt affirmation/negation in (6b) and the necessity of overt negation (not affirmation) in (6a) remain unexplained. (Another modal verb, *hoeft* 'need' is clearly a negative polarity item in this context, as Barbiers notes). Second, the interpretation of the verb as 'like' rather than 'need/want' as with the indefinite nominal complements (e.g. 'John definitely doesn't want Mary' in (6a)) is unexplained (although Barbiers does note that the 'like' use in (6b) is possible with auxiliary *mag* as well as main verb *mag*). Third, if the affirmative/negative scale permitted the use of the modal in this way, the predicative NPs (e.g. *dokter*) should become grammatical in this construction when followed by negation, as should the predicative non-scalar PPs and APs mentioned above. Something remains to be captured about this construction.

Even with the indefinite DP complements, it seems to me that some difficulties arise for Barbiers. His claim is that the numerical scale introduced by the indefinite determiner or numeral creates a bounded lattice, with the numeral in question being the upper bound. The scalar nature of this lattice seems doubtful to me (*Jan moet een hond*, with *een* the upper bound, creates a scale of dogs from zero to one?), but even when an indefinite is a mass noun, or a plural, with no natural upper bound, these sentences are grammatical (*Jan moet water/bonen*). Finally, Barbiers' account of the acceptability of definite DPs with some focus (*Jan moet de HOND/DIE hond*) seems flawed. He maintains that when focus arises, a presupposed contrastive set is created, and the superset/subset relation exists. This introduces a bounded lattice, again from 0 to 1. This is plausible, but there is a disturbing exception: focused proper names do not create grammaticality in this construction (**Jan moet MARIE*). Barbiers claims that this is because the subset/superset relation may not be created with proper names, but it is perfectly reasonable to assume that the discourse might provide a group of individuals which could contrast with any other individual ('John wants Mary, (not Sue or Bill or Alex)'); indeed this is exactly the interpretation of a cleft construction: *Het is Marie die Jan leukvind* 'It is Mary that John likes (not Sue or Bill)' (ex. above again from Angeliek van Hout, p.c.).

Note that none of these points invalidate Barbiers' primary conclusions about the bounded lattice requirement or the derivation of the root/epistemic distinction. It strikes me that the crucial difference between the PPs and APs which work very well and the DPs which are more problematic is the attributive nature of the PPs and APs, and the non-attributive nature of the DPs: in none of the DP constructions is the DP a predicate to the subject. For this reason, I have reservations about Barbiers' final proposal about the structure of the split VP; he claims that the upper VP projection can in fact be interpreted as a D head, and the lower VP projection as the N head; that, in fact, DP is the fundamental structure in UG. (The modal may adjoin either to DP or NumP to derive the root/epistemic distinction). While the scope and aim of the proposal is admirable, and the comparison of possible subject relations to possible possessor relations (source, goal, causer, possessor, etc.) is intriguing, for me it remains a speculation.

Conclusion

Barbiers' contention that morphological/PF requirements play no role in the motivation of syntactic movement is a strong and perhaps arbitrary claim. However, his Principle of Semantic Interpretation, enforcing the requirement that semantic relations be triples, proves to be essentially workable as a constraint, leading to some quite significant results, as does his revised definition of c-command. The points raised above with respect to his case studies do not cast serious doubt on any of his analyses. It is, perhaps, possible to adopt the syntactic analyses he proposes without adopting the Principle of Semantic Interpretation as the sole motivation for syntactic movement (although the definitions of c-command and immediate c-command are integral parts of that analysis). However, the complete picture hangs together well enough to merit serious consideration as a whole approach.

Precisely because of its viability as a framework, however, it would have been appropriate to consider at least briefly some of the larger questions raised by discarding morphology as a motivation for movement. In particular, the empirical basis for choosing morphology as a locus for movement motivation is the obvious parallel between cross-linguistic morphological variation and cross-linguistic movement variation. If all languages are the same at LF, and all movement is semantically motivated, why do languages differ in their movement possibilities? Further, within languages, why are some movements necessarily overt, some necessarily covert, and some optional (like PP extraposition)? What does the strongT/weakT parameter look like in such a framework (indeed, why is the establishment of a semantic relation between T and a DP necessary at all)? It is exactly because of the arbitrariness of morphology that minimalist syntax chooses to ascribe these seemingly arbitrary differences between languages to the morphological component. Admittedly, it's a cheap trick. On the other hand, it captures the observed facts. It is reprehensible to reproach an author who has produced a rigorous and broad analysis with a wish that he had written more, but a sketch of movement parameterization would have been very welcome indeed. In any case, that just means that I look forward very much to further work from this author.

Acknowledgement

Thanks very much to Angeliek van Hout for judgments and discussion; standard disclaimer applies in spades.

References

Baker, M. (1985). The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* **7,** 265–300.

Barbosa, P. (1995). *In defense of right-adjunction for head movement.* Ms., MIT.

Bennis, H. (1986) *Gaps and dummies.* Dordrecht: Foris.

Bobaljik, J. (1995). *Morphosyntax: The syntax of verbal inflextion*, Diss., MIT.

Borer, H. (1996). *Argument structure, aspect, and variable-behavior verbs.* Talk given at University of Pennsylvania, Fall 1996.

Brander, E. (1995). Review of *Dependent nexus: Secondary predication in English and the Scandinavian languages* by Peter Svenonius, in *Glot* **1,** 9/10.

Carnie, A. (1995). *Non-verbal predication and head movement*. Diss. MIT.

Chomsky, N. (1993). A minimalist program for linguistic theory. In *The view from building 20*, K. Hale & J. Keyser (eds.), 1–52. Cambridge, Mass.: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Kayne, R. (1995). *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Miyagawa, S. (1989). *Structure and case marking in Japanese*. New York: Academic Press.

Rooryck, J. (1995). *Focus as a primitive of raising verbs.* Paper presented at the 18th GLOW colloquium, University of Tromsø. *GLOW newsletter* **34.**

Travis, L. (1984). *Parameters and effects of word order variation*. Diss., MIT.

Travis, L. (1988). *The syntax of adverbs.* Paper presented at the 4th Comparative Germanic Syntax Workshop, McGill University.

Vikner, S. (1995) *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Watanabe, A. (1993) Agr-based Case Theory and its interaction with the A-bar system. Diss., MIT.

MOVEMENT AND PARAMETRIC VARIATION IN SYNTAX

by *Yuji Takano*
Reviewed by *Hiroto Hoshi*

Summary by *the author*

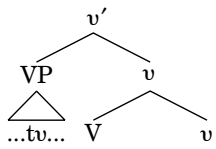
This thesis concerns issues centering around parametric syntax, with special focus on movement, linear order, and possible parameters in syntax. Within the general framework of the Minimalist Program, I investigate in detail the nature of parameters responsible for ordering variation between head-initial and head-final languages, and argue that head movement and scrambling play an essential role in deriving typological similarities and differences between the two types of languages, in particular, between English and Japanese.

Kayne (1994) proposes a theory of phrase structure and linear order in which hierarchical structure uniquely determines linear order. Under Kayne's proposal, asymmetric c-command relations among nonternimals of a given phrase structure invariably map into linear precedence relations among terminals of the phrase structure, and, as a result, UG allows only the S(pecifier)–H(ead)–C(omplement) order. The latter result in turn leads to the conclusion that the C-H order as found in languages like Japanese must result from overt movement of C to some functional domain.

In this thesis, I propose an alternative theory that attributes ordering variation to the presence/absence of head movement. Specifically, adopting the Bare Phrase Structure Theory of Chomsky (1994, 1995), I propose that a set theoretic syntactic object Σ formed by applications of Merge in the core computation undergoes Linearization in the phonological component, which consists of an operation that breaks down Σ into two parts by "reversing" the effects of Merge, and a second operation that concatenates the resulting parts. The former operation is called Demerge, and the latter Concatenate. Note that by virtue of the nature of Merge, Σ always consists of two constituents, one maximal and the other nonmaximal. Assuming that Demerge, just like Merge, can only "see" maximal categories, Demerge breaks down Σ into two elements by detaching the maximal constituent from Σ. To illustrate, suppose that Σ = VP, consisting of the two constituents α and V′. If Linearization applies to Σ, Demerge detaches α from Σ, given that α is maximal, whereas V′ is nonmaximal, and yields two separate elements α and V′, which is now maximal, according to Chomsky's (1994, 1995) view of phrase structure. The resulting elements α and V′ are ordered by Concatenate, which makes crucial reference to an asymmetric property inherent to the application of Demerge: V′ becomes an independent maximal element only after α is detached from Σ, and not vice versa. I propose that Concatenate map this abstract "precedence" relation to the precedence relation in temporal order at PF. Thus, in the case of α and V′, Concatenate determines the order α+V′ (α precedes V′). Linearization then applies recursively to each of α and V′, and then to smaller elements resulting from Linearization of α and V′, and so on, until Σ is completely linearized into a sequence of terminals.

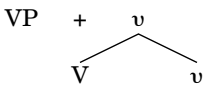
In this theory, if Σ involves no movement, the S–C–H order results, whereas the H–C order results when Σ involves movement, a conclusion exactly opposite to Kayne's. I further argue that the H–C order derives from raising of H to a higher position in Σ (in fact, this is one case; there is another way of deriving the H–C order that involves direct merger of H rather than movement of H, which I put aside here). Under the standard analysis of head movement, the head H adjoins to another head, creating a "two-segmented" structure. Capitalizing on the nature of the two-segmented structure thus created, I derive the H–C order in the following way. Suppose that V raises to *v*, yielding the following (sub)structure (order irrelevant):

(1)



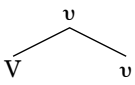
In this structure, because of the adjunction structure formed by head movement, *v*′ has three constituents: VP, [*v*, *v*] (two-segmented category), and V. Of the three constituents, VP and V are maximal and hence are "visible" to Demerge. Thus there are two options for Demerge: detaching VP and detaching V. Suppose that VP is detached first. This yields the following result (after concatenation):

(2)



Now the question is how the root adjunction structure in (3) can be linearized:

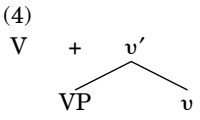
(3)



Noting that V is a constituent of the upper segment *v*, but not a constituent of the two- segmented category [*v*, *v*], I suggest that the root adjunction structure as in (3) can never be linearized, on the assumption that Demerge always detaches a maximal constituent of Σ, Σ a maximal category (rather than segment). As a result, the option of detaching VP from the structure in (1) fails (the derivation carrying out this option crashes at PF, with the unlinearized structure in (3) remaining at PF).

The same problem does not arise if V is detached from the structure in (1). After detachment of V, we have the following:

(4)



Further application of Linearization yields the sequence V+VP+*v*, which ensures that the raised V precedes all VP-internal elements on the surface, namely the "head-initial" order.

I argue that this theory resolves the problems that Kayne's original theory faces, while maintaining its basic insights into the nature of the relation between hierarchy and linearity (especially, the "leftness property" of syntax, namely that the specifier of the head H always precedes the complement of H, and that upward movement always results in leftward movement). In particular, this theory need not posit functional categories of a rather unclear nature to account for word order in Japanese, and is also consistent with the "traditional" intuition about Japanese syntax, expressed explicitly by Fukui (1986, 1988) (see also Kuroda 1988), that Japanese lacks "active" functional categories triggering overt movement and agreement. Furthermore, the proposed theory, like Kayne's theory, eliminates ordering parameters from UG entirely and thus contributes to restricting the theory of parameters in such a way as to reduce all language variation to parameters associated with the properties of functional categories.

I also show that this proposal opens up a new way of looking at the typological differences between English and Japanese, and between VO and OV languages in general. Under this proposal, English and Japanese differ in that English has verb raising, while Japanese does not (the difference being ultimately attributed to the different properties of *v* in the two languages). I demonstrate that this single difference derives a wide range of typological differences between English and Japanese, including those related to head-complement order, overt Case particles, multiple objects, adjacency effects, scrambling, overt *wh*-movement, overt question particles, complementizers, and head–complement order in synthetic compounds. This type of unification has been impossible in traditional approaches. The proposal also has consequences and implications for complex predicate formation, the theory of formal features, various aspects of the Japanese Case system, and the "leftness property" of syntax.

The present proposal about the relation between hierarchy and linearity leads to the conclusion that both English and Japanese have short scrambling (scrambling within VP) in the double PP structure (e.g. *talk to Bill about Mary*). On the basis of extensive examination of the double complement constructions in English and Japanese, I extend this claim to a broader empirical domain and show that English has short scrambling in a wide range of cases involving the V–DP–PP structure (e.g. *give a book to John*). I argue that in the English V–DP–PP structure, if the DP is a theme and the PP is another internal argument of the verb, the DP is always moved from a position lower than the PP. I further claim that the relevant movement of the theme DP is "partial" in the sense that it moves within VP and not to a position for Case assignment/checking, and hence should be analyzed as short scrambling. Despite the optional character of scrambling, partial object shift in English is obligatory. I show that the apparent obligatoriness of short scrambling follows from the nature of Attract-F and convergence under the Minimalist Case Theory of Chomsky (1995). Extending this analysis to cases involving partial movement of nominative DPs, I also show that nominative DPs undergo short scrambling in the existential construction in English (e.g. *There is a dog in the garden*). These considerations indicate that English has partial shift of accusative and nominative DPs, both analyzed as involving short scrambling. In contrast to the situation in English, short scrambling in Japanese is always optional. This fact supports the conclusion reached in Chapter 2 of the thesis that Japanese employs a non-checking system for Case. These results also lend strong empirical support to our claim that the VP-internal elements in English and Japanese have exactly the same hierarchical and linear properties, except the position of the verb.

Finally, I discuss the nature of scrambling. Scrambling is optional movement involving no feature checking (Fukui 1993; Fukui & Saito 1996). This view of scrambling raises nontrivial questions in the Minimalist Program. In particular, if Last Resort forces all movement to be driven by the necessity of feature checking, scrambling cannot be an instance of movement. Nor does its optionality fall out easily in this

framework. On the other hand, scrambling does show the properties of movement. I address these issues from the perspective of the Minimalist Program. First, on the basis of certain conceptual considerations, I propose to modify Chomsky's (1995) model of the computational system. The revised model of the computational system considerably reduces the role of the covert component and requires the whole syntax to be strictly cyclic. This proposal necessitates reconsideration of the nature of head movement. In this approach, head movement cannot be movement of a head to another head, as is standardly assumed. I will propose an alternative view in which the effect of head movement results from a combination of Copy and pure Merge ("generalized transformation" in the terms of earlier frameworks). On this analysis, "head movement" no longer involves Move (or Attract) in Chomsky's (1995) sense. Then I extend this analysis of head movement to scrambling, arguing that the effect of scrambling also results form Copy and pure Merge, rather than Move. This Copy + pure Merge approach resolves the paradoxical situation noted above with scrambling. Last Resort is irrelevant to scrambling because it does not involve Move. From this, its optionality also follows. On the other hand, scrambling analyzed as Copy + pure Merge does form a chain. Thus the movement properties of scrambling follow, to the extent that they are attributed to chain properties. I also show that this analysis explains why scrambling without "visible effects" is impossible and why Japanese allows scrambling out of VP, whereas English does not.

References

Chomsky, N. (1994). Bare Phrase Structure. *MIT Occasional Papers in Linguistics* **5,** Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.

Chomsky, N. (1995). *The minimalist program*, Cambridge, Mass.: MIT Press.

Fukui, N. (1986). *A theory of category projection and its applications*. Diss., MIT, Cambridge, Massachusetts. [A revised version published in 1995 as *Theory of Projection in Syntax*, CSLI, Stanford, California.]

Fukui, N. (1988). Deriving the differences between English and Japanese: A case study in parametric syntax, *English Linguistics* **5,** 249–270.

Fukui, N. (1995). The principles-and-parameters approach: A comparative syntax of English and Japanese. In *Approaches to language typology*, M. Shibatani and T. Bynon (eds.), 327–371. Oxford: Clarendon Press.

Fukui, N. & M. Saito (1996). Order in the theory of phrase structure and movement. Ms., University of California, Irvine and Nanzan University.

Kayne, R. (1994). *The antisymmetry of syntax*, Cambridge, Mass.: MIT Press.

Kuroda, S.-Y. (1988). Whether we agree or not: A comparative syntax of English and Japanese. In *Papers from the Second International Workshop on Japanese Syntax*, W. Poser (ed.), 103–144. CSLI, Stanford, California.

Review: Accusative case and complex predicate formation in Japanese by *Hiroto Hoshi*

1. Introduction

In this valuable work on parametric syntax, Takano investigates, within the framework of Chomsky's (1993, 1994, 1995) Minimalist Program, such issues as movement, linear order and possible parameters in syntax, and he proposes an elegant theory which derives the typological differences between head-initial and head final-languages, in particular, between English and Japanese, from a single parameter, i.e. the presence/absence of head movement.

Two of the most important theoretical points Takano makes are the following: First, he rejects parameters associated with ordering variation like Kayne (1994), and argues that only functional categories are subject to parametric variation, thereby simplifying the theory of parameters (cf. Fukui 1986, 1988, 1995; Fukui & Saito 1996). Secondly, following Kayne, he also argues that syntax has the "leftness" property: (i) what is structurally higher precedes what is lower; (ii) upward movement results in leftward movement. To capture this property, Kayne proposed that UG allows only Specifier–Head–Complement order and the Specifier–Complement–Head order found in languages like Japanese must result from overt movement of Complement to some functional domain above Head but below Specifier. However, Takano convincingly shows that there is no strong evidence which indicates that Japanese has such functional categories or such obligatory overt movement of Complement which are crucially assumed to be present in Kayne (1994) (Fukui 1986, 1988; Kuroda 1988) and he concludes that we cannot maintain Kayne's theory. Takano, then, proposes an ingenious alternative theory which explains the leftness property of syntax and the ordering variation between head-initial and head-final languages in a different way. Specifically, Takano develops a theory in which head movement plays a crucial role in determining the word order variation in the two types of languages. Because overt head movement obligatorily takes place in head-initial languages such as English, Head precedes Complement in those languages. Japanese, a head-final language, does not have such obligatory overt head movement, and thus, Head follows Complement. (See Takaon's summary above for the detailed discussion about the interaction between head movement and "linearization," which consists of "demerge" and "concatenate".)

Furthermore, Takano derives, from this single parameter, i.e. the presence/absence of head movement, in particular, verb raising, a very wide range of typological differences between English and Japanese, including those related to overt Case particles, multiple objects, adjacency effects, scrambling, overt *wh*-movement, overt question particles, complementizers, and head–complement order in synthetic compounds.

Considering the importance of the theoretical points the work under review makes and the wide range of data which his theory accounts for, I believe that it will become one of the most important pieces of work on parametric syntax. In this review, however, I will focus on discussing issues such as verb raising in Japanese and point out that there is some empirical evidence which would pose potential problems for Takano's claims that accusative Case is inherent in Japanese and complex predicate formation takes place in the phonological component, but not in the core computation. In Section 2, after briefly summarizing some of the important proposals regarding verb raising, linearization and accusative Case feature licensing which are made in Chapter 2 of Takano's dissertation, I will show some data of which Takano's theory cannot provide a straightforward analysis. In Section 3, I will argue that although Takano convincingly defends the leftness property of syntax in Chapter 3, there is some evidence which poses potential problems for Kayne's and Takano's antisymmetrical approach to syntax in which there does not exist upward rightward movement. In Section 4, I will conclude the discussions of this review.

2. Verb raising, linearization, and Case assignment/checking

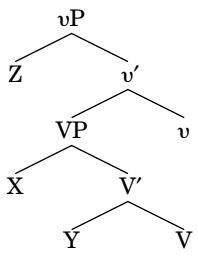
Takano follows Chomsky (1994, 1995) in claiming that a set theoretic object Σ formed by application of merge in the core computation undergoes linearization in the phonological component and that linear order plays no role in the N(umeration) $\rightarrow$ λ computation. To derive the typological differences such as head–complement/complement–head order in English and Japanese, he proposes the following important parameter:

(1)
Light head $\varkappa$ has [−X] in English but not in Japanese.

Here, Takano again follows Chomsky to claim that [−X] of light head $\varkappa$ is a strong feature, which requires overt movement of lexical affixal head X.

Thus, if we have the structure for the core proposition of a ditransitive verb given in (2) in Japanese, V does not raise to light head υ. (Just for ease of exposition, I adopt informal tree diagrams such as (2) rather than set-theoretic objects here. The order of terminals and non-terminals in tree diagrams is irrelevant.)

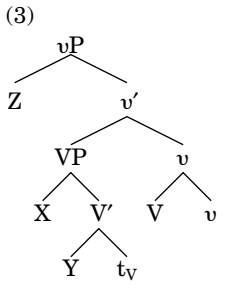
(2)



This is because light head υ in Japanese does not have [−V] feature. In PF, linearization, i.e. demerge and concatenation, applies to this object, and assigns structure (2) the surface order Z–X–Y–V–υ, i.e. Specifier–Complement–Head order, in this theory. (See Takaon's summary above for the explanation of why structure (2) is linearized in this way.)

In contrast, if we are given the same syntactic object as (2) in English, V obligatorily raises and adjoins to υ in the overt syntax as shown below,

(3)



because if V raises to υ, it enters into a checking relation with υ, successfully eliminating [−V]. However, if the verb raising of V does not take place, [−V] of light head υ remains at LF and hence the derivation crashes at LF because [−V] is not only uninterpretable but also strong. So, it must be eliminated even before Spell-Out. In PF, (3) is mapped into Z–V–X–Y–t_V–υ, i.e. Specifier–Head–complement order, because under Takano's theory, the raised verb is forced by linearization to precede all VP-internal elements.

As is explained in some detail in the author's summary above, under Takano's theory, a set-theoretic syntactic object Σ formed by merge in the core computation undergoes linearization in PF. This operation, linearization, consists of two natural PF operations: demerge and concatenate. Demerge simply reverses the effects of merge, and concatenate maps the abstract precedence relation inherent in demerge to the precedence relation in temporal order at PF. Based on these natural assumptions, Takano proposes an elegant theory which not only accounts for word order variation, but also captures the leftness property of syntax by appealing to the presence/absence of head movement, in particular, verb raising. As was illustrated above, in Takano's theory, if a language does not involve verb raising as in (2), we have Specifier–Complement–Head order. If, on the other hand, a language involves verb raising as in (3), we have Specifier–Head–complement order. Importantly, by doing this, Takano successfully eliminates parameters associated with ordering variation like Kayne (1994), and shows a different way from Kayne's to reduce all language variation to

parameters associated with the properties of functional categories (cf. Fukui 1986, 1988, 1995; Kuroda 1988). Takano also convincingly argues that his theory has huge advantages over Kayne's theory by showing that Japanese does not have functional categories or obligatory movement of Complement to those functional domains which are assumed to play a crucial role in deriving word order variation under Kayne's theory. Furthermore, Takano showed that his theory accounts for a wide range of typological differences between English and Japanese with a single parameter given in (1).

Given the assumptions above, it follows from Takano's proposals above that Japanese, a head-final language, does not have any sort of verb raising, contrary to claims by Otani & Whitman (1991) and Koizumi (1995), among others. Further, complex predicates in Japanese such as *yom-ase* in (4) cannot be formed in the overt syntax, contrary to proposals by Kuroda (1965, 1993), Kuno (1973), Shibatani (1973), among others.

(4)
John-ga Mary-ni hon-o yom-ase-ta.
John-NOM Mary-to book-ACC read-CAUSE-PST
'John made Mary read a book.'

Specifically, basing himself on empirical facts, Takano proposes that causative sentences such as (4) maintain "biclausal" structure in the N → λ computation which does not involve any verb raising, and that complex predicates such as *yom-ase* above are formed in the phonological component. That is, the two syntactically separate verbal morphemes, *yom* and *ase*, in (4) undergo "phonological merger" (or "cliticization") carried out by phonological rules under adjacency which is assumed to be a necessary condition for this phonological process.

Given the condition that V enters into a checking relation with the accusative DP only if V raises to υ (Chomsky 1995), it also follows from Takano's theory that the accusative Case feature of the DP cannot be eliminated by checking in Japanese because verbs simply do not raise to υ, as illustrated in (2). In this way Takano makes the important proposal that Case particles such as *o* make the Case feature of DPs visible to Spell-out and that Spell-out removes those Case features from Σ — the derivation forming Σ yielding λ without crashing at LF. Takano further claims that accusative Case features of transitive verbs cannot be eliminated by checking in Japanese either (because of the lack of V-raising), and as a result, accusative Case in Japanese must be inherent (Takahashi 1993), that is, it must always be linked to a particular θ-role, making the Case feature of transitive verbs interpretable at LF.

By making these proposals, Takano claims that his theory accommodates the fact that Japanese can have multiple objects with accusative Case/postposition *o* in a sentence, since Spell-out can eliminate multiple accusative/postpositional features as long as they are realized by overt particles. Consider the examples in (5).

(5)
a. ??Mary-ga John-**o** hoho-**o** butta.
Mary-NOM John-**ACC** cheek-**ACC** hit
'Mary hit John on the cheek.'
b. Mary-ga hoho-**o** butta no-wa John-**o** da.
Mary-NOM cheek-**ACC** hit COMP-TOP John-**ACC** COP
'It is John that Mary hit on the cheek.'
(Harada 1973; Kuroda 1978, 1988)

(5b), where *John* is marked with accusative Case *o* and *hoho* is marked with the postposition *o*, is perfectly grammatical. The marginality of (5a) is attributed to Harada's (1973) language particular constraint, the "surface" double-*o* constraint, which mildly prohibits multiple occurrence of *o* in a single clause.

Takano further claims that his theory can also account for the ungrammaticality of English example (6) in contrast with the grammaticality of (5b),

(6)
*Mary hit John his/the cheek.

because in English, a transitive verb checks the accusative Case feature of a DP but the "one-to-one convention" is required for checking/agreement. The contrast between (5) and (6) thus lends support to Takano's proposal that to eliminate Case features, English adopts the checking system, whereas Japanese utilizes Spell-out.

Consider now the ungrammatical example in (7).

(7)
*John-ga [Mary-**o** hon-**o** yom]-ase-ta.
John-NOM Mary-**ACC** book-**ACC** read-CAUSE-PST
'John made Mary read a book.'

(7) is another instance of the multiple object construction in Japanese, but (7) is far worse than (5a). Harada (1973), Kuroda (1978, 1988), Poser (1981), Saito (1985), among others, ruled out examples such as (7) by appealing to Harada's (1973) "abstract" double-*o* constraint, which Saito (1985) interprets as follows: a verb can assign accusative Case to at most one NP in Japanese. This traditional account of (7) crucially assumes that in Japanese causatives, the embedded verb raises to the matrix causative verb and forms one complex verb in the overt syntax and examples such as (7) are ruled out at SS/LF. (Notice that a complex predicate counts as a single verb for the double-*o* constraint.) Importantly, however, Takano cannot rule out examples such as (7) at SS/LF. This is so, because he proposes that due to the lack of verb raising in Japanese, the embedded verb *yom* and the matrix causative verb *ase* in (4) and (7) are separated throughout the derivation from N to λ, and that they are merged into a single verb only in PF. Takano subsequently reinterprets the abstract double-*o* constraint as a PF constraint, as stated in (8).

(8)
A single verb can have at most one [assign accusative Case].

Under the assumption that accusative Case feature of a transitive verb is visible at PF, example (7), after word formation in the phonological component, contains a single verb *yom-ase* that has two accusative Case features, one from the embedded V and the other from the matrix V. Thus, the example violates Takano's version of the abstract double-*o* constraint (8) at PF. Importantly, on Takano's theory, we can correctly distinguish the grammaticalilty of the examples in (5a,b) and (7). (5a,b) are not in violation of constraint (8) because *but* 'hit' has only one accusative Case feature, and not (5b) but (5a) violates the surface double-*o* constraint. Hence, (5a) is marginally ungrammatical, while (5b) is perfectly grammatical. (7) is totally out, violating PF constraint (8) and the surface double-*o* constraint.

Although Takano's theory elegantly accounts for the word order difference in English and Japanese and explains the data in (5), (6) and (7) in an innovative way, questions arise as to the validity of some of the basic assumptions under his theory. Given his attractive theory, I believe that it would be desirable for Takano to come up with solutions for the following questions to develop his theory further. First, Kuroda (1981, 1991) convincingly argued that the causative verb *sase* can be detached from the embedded verb by showing examples such as (9).

(9)
Masao-ga Hanako-ni mondai-no arika-o
Masao-NOM Hanako-to issue-GEN point-ACC
wakar-anaku sase-ta.
understand-not CAUSE-PST
'Masao made Hanako not understand the point of the issue.'
(Kuroda 1981, 1991)

And as Mamoru Saito points out in a personal communication, examples like (10) are ungrammatical and have the flavor of the abstract double-*o* constraint effects.

(10)
*Taroo-ga Hanako-**o** LI-**o** yomi-mo sase-ta.
Taroo-NOM Hanako-**ACC** LI-**ACC** read-also CAUSE-PST
'Taroo made Hanako read LI, too.'

Notice that the embedded V *yom* and the matrix V *sase* are detached from each other in example (10) but it still violates the abstract double-*o* constraint. The ungrammaticality of this example thus implies that the abstract double-*o* constraint cannot be a PF constraint like the one stated in (8) by Takano. To avoid this problem, we might want to adopt, under Takano's theory, Koizumi's (1995) version of the abstract double-*o* constraint which requires that there should be no more than one accusative phrase within a single TP in Japanese. However, if we adopt this, a question immediately arises as to why T is relevant to licensing for accusative Case in Japanese. Under Takano's theory, it is crucially assumed that accusative Case features of transitive verbs need not be eliminated by checking in Japanese, because those features are inherent and are linked to particular θ-roles, making the Case features of transitive verbs interpretable at LF. (See Koizumi 1995 for his argument that the abstract double-*o* constraint can be derived from Watanabe's 1993 Case Theory in which accusative Case in Japanese is assumed to be abstract/structural Case.)

Secondly, Chomsky (1981) and Jaeggli (1986), among others, claim that in general, the "Case absorption" applies only to structural/abstract Case and not to inherent Case. Hence, if accusative Case in Japanese is inherent as Takahashi (1993) and Takano argue, we predict that there does not exist any Case absorption phenomenon in Japanese. Consider now the examples in (11). (11a) is from Poser (1981), and (11b,c) from Saito (1982).

(11)
a. kono hon_i-wa John-ga Hanako-***o**/
this book$_{TOP}$ John-NOM Hanako-**ACC**/
ni $[e]_i$ $[_V$ yom-ase-ta].
to read-CAUSE-PAST
'This book, Mary made John read it.'
b. kono hon_i-ga John-***o**/ni $[e]_i$ $[_V$ yom-ase-yasu-i].
this book$_{NOM}$ John-**ACC**/to read-CAUSE-easy-PRS
'This book is such that it is easy to make John read it.'
c. Mary-ga $John_i$-**o**/ni damatte Tom-ni $[e]_i$
Mary-NOM John$_{ACC}$/to silently Tom-by
$[_V$ sika-rare-sase-ta].
scold-PASS-CAUSE-PST
'Mary made John be scolded by Tom without saying anything.'

Contrary to the above prediction, Saito (1982) observed the Case absorption phenomenon in the passive example in (11c). In the topic construction (11a) and the *tough*-construction (11b), the causees *Hanako* and *John* cannot be marked with *o*, whereas in passive-causative example (11c), the causee *John* can be attached with the accusative Case *o*. This is because in (11a,b), the gap which is coindexed with *kono hon* requires accusative Case. If the causee were marked with *o* in examples (11a,b), these examples would be in violation of the abstract double-*o* constraint. The passive example (11c), on the other hand, does not show the abstract double-*o* constraint effects, because the empty category coindexed with *John* does not require Case due to the Case absorption of the passive morpheme *rare*. Thus, Saito convincingly argued that accusative Case in Japanese is structural/abstract Case. This conclusion also poses a potential problem for Takano's theory because the conclusion implies that V moves up to υ to eliminate [−V] of υ and to check accusative Case features of DPs in Japanese as well as in English. This flatly contradicts Takano's claims.

Thirdly, Miyagawa (1986) observed the contrast in (12a,b).

(12)
a. *Taroo-ga [Hanako-ga kuruma-o soogankyoo-de
Taroo-NOM Hanako-NOM car-ACC binoculars-with
nusunda no]-o mi-ta.
stole COMP-ACC see-PST
'Taroo saw with binoculars that Hanako stole a car.'

b. Taroo-ga Kanda-ni [PRO hon-o zitensya-de
Taroo-NOM Kanda-to book-ACC bicycle-by
kai-ni] itta.
buy-to went
'Taro went to Kanda by bicycle to buy a book.'

(12a) is anomalous because the PP *soogankyoo-de* is in the embedded clause although the PP must be construed with the matrix verb for the sentence to be semantically well-formed. (12b), on the other hand, allows the PP *zitensya-de* to be construed with the matrix verb, which is the appropriate interpretation, in spite of its apparent placement inside the infinitival clause. By showing this contrast, Miyagawa argued that both (12a) and (12b) start out as a "bisentential" structure, but only (12b), the purpose expression, underwent "restructuring," becoming a "simplex" structure. Therefore, the PP *zitensya-de* is construed with the matrix verb *it-ta* in (12b).

Given this, consider also the following causative examples from Miyagawa (1986):

(13)
a. John-ga kodomo-**o**/ni [PRO hon-o kai-ni]
John-NOM child-**ACC**/to book-ACC buy-to
ik-ase-ta.
go-CAUSE-PST
'John made the child go out to buy a book.'

b. Taroo-ga kodomo-***o**/ni [PRO hon-o zitensya-de
Taroo-NOM child-**ACC**/to book-ACC bicycle-by
kai-ni] ik-ase-ta.
buy-to go-CAUSE-PST
'Taro made the child go to buy a book by bicycle.'

The causee *kodomo* can be marked with accusative Case *o* in (13a) but not in (13b). Miyagawa argued that in (13b), the PP *zitensya-de* in the lower clause requires restructuring, which forces *kai-ni ik-ase* to form a complex predicate in the core computation. Consequently, the abstract double-*o* constraint disallows the causee to be marked with *o* in (13b) at SS/LF. It is not clear how Takano would treat restructuring under his system, but given Matsumoto's (1996) proposal that in examples such as (13b), *kai-ni* and *ik-ase-ta* are morphologically separated, it seems to be clear that Takano cannot rule out (13b) by PF constraint (8). To the extent that Miyagawa's arguments are correct, Takano is thus required to come up with a way to account for the interaction between restructuring and the abstract double-*o* constraint effects, shown in (13b).

Finally, consider the following object control constructions:

(14)
a. watasi-wa Michael-ni [PRO suupaa-de nanimo
I-TOP Michael-to supermarket-in anything
kaw]-ase-nakat-ta.
buy-CAUSE-not-PST
'I did not make Michael buy anything in the supermarket.'

b. ?*watasi-wa Michael-ni [PRO suupaa-de nanimo
I-TOP Michael-to supermarket-in anything
kaw-yoo-ni] meizi-nakat-ta.
buy-to order-not-PST
'I did not order Michael to buy anything in the supermarket.'

c. ?*watasi-wa Michael-o [PRO suupaa-de nanimo
I-TOP Michael-ACC supermarket-in anything
kaw-yoo-ni] settokusi-nakat-ta.
buy-to persuade-not-PST
'I did not persuade Michael to buy anything in the supermarket.'

There is a contrast in grammaticality between (14a) and (14b,c). Namely, the negative polarity item *nanimo* in the embedded clause is licensed by the negative morpheme in the matrix clause in Japanese causative (14a), whereas this is not the case in (14b,c). Nemoto (1993) proposed based on a different set of data that verb rasing helps a negative polarity item in the embedded clause be licensed by the negative morpheme in the matrix clause. Given her analysis, we can straightforwardly account for the contrast in (14a/b,c), assuming that only (14a) involves verb raising. Under Takano's theory, however, this type of contrast does not seem to be expected in (14a–c), if we assume that negative polarity items are licensed in the core computation. Recall that on his proposal, Japanese lacks verb raising and complex predicates such as *yom-ase-nakat-ta* are not formed from the N to λ computation but are formed by phonological merger in PF.

3. Leftness property of syntax and upward rightward movement

In Chapter 3, Takano extensively examines the properties of the double complement constructions in English and Japanese, and he defends the leftness property of syntax (cf. Hoji 1985; Larson 1988). Namely, within VP, what precedes is structurally higher than what follows in English and Japanese (Kayne 1994). Furthermore, by showing examples such as the ones in (15), which display connectivity (or reconstruction effects),

(15)
a. ?I gave each other$_i$'s babies to the mothers$_i$.
(Kitagawa 1994)
b. ?I borrowed each other$_i$'s pictures from the boys$_i$.
(Pesetsky 1995)
c. ?Mary puts its$_i$ label on every box$_i$.
(Pica & Snyder 1995)

Takano argues that in the English V–DP–PP structure, if the DP is a theme and the PP is another internal argument of the verb, the DP always undergoes "partial" object shift, i.e. "short scrambling," overtly from a position lower than the PP.

The structure Takano proposes for example (16a) is given in (16b).

(16)
a. John gave a book to Mary
b. [$_{TP}$ John$_i$ T [$_{vP}$ t_i [$_{v'}$ gave$_j$-v [$_{VP}$ a book$_k$ to Mary [$_{V'}$ t_j t_k]]]]

Here, the theme DP *a book* undergoes partial object shift, i.e. short scrambling, moving only within VP and not to a position for Case assignment/checking.

Despite the fact that scrambling in Japanese is optional, the partial object shift in English is obligatory, as shown below:

(17)
a. *John gave to Mary a book
b. *[$_{TP}$ John$_i$ T [$_{vP}$ t_i [$_{v'}$ gave$_j$-v [$_{VP}$ to Mary [$_{V'}$ t_j a book]]]]

Takano provides an ingenious analysis of the contrast between (16a) and (17a) by appealing to the nature of Attract-F and convergence under the Minimalist Case Theory of Chomsky (1995). Notice that in (17b), the accusative Case feature of the DP *a book* cannot be attracted by *gave$_j$-v*, since there is another DP *Mary* that is closer to *gave$_j$-v*. Since the Case feature of *a book* can never be checked, the derivation yielding (17b) crashes at LF. On the other hand, the derivation yielding (16b) does not have the same problem. Here *a book* is scrambled over *to Mary* before the derivation reaches the stage where *gave$_j$-v* attracts the accusative feature. In (16b), this scrambling makes it possible for the Case feature of *a book* to be attracted by *gave$_j$-v*. Therefore, although scrambling itself is optional, only the derivation involving scrambling of the accusative DP converges, because of the nature of Attract-F.

Takano convincingly argues for Kayne's claim for the leftness property of syntax in Chapter 3, but there is some evidence which indicates that we cannot deny the existence of upward rightward movement. Therefore, we seem to be required to find a way to revise the Kayne/Takano antisymmetrical approach to syntax so that we can accommodate upward rightward movement under it. To finish the discussion of this section, I will show three pieces of evidence for upward rightward movement below. First, consider Saito's (1994) examples in (18). (Saito attributes the observation to Daiko Takahashi.)

(18)
a. *Mary wanted [PRO to meet [the men who had been accused of the crime]$_i$ until each other$_i$'s trials.
b. Mary wanted [PRO to meet t_i] until each other$_i$'s trial [the men who had been accused of the crime]$_i$

By showing the contrast in (18a,b), Saito argued that heavy NP shift raises the object NP [*the men who had been accused of the crime*] out of the embedded non-tensed clause so that the NP binds anaphors such as *each other* in the matrix clause as shown in (18b).

Second, consider the following parasitic gap constructions:

(19)
a. Which articles$_i$ did you file t_i without reading e_i?
b. *Who$_i$ t_i expected Bill to send a picture of e_i?
c. Which men$_i$ did the police warn t_i [$_{CP}$ that they were about to arrest e_i?

To account for the contrast between (19a) and (19b), the "anti-c-command requirement on parasitic gaps" in (20) was adopted in Taraldsen (1981), Chomsky (1982) and Engdahl (1984).

(20)
A parasitic gap is licensed by a variable that does not c-command it.

On the other hand, to account for the grammaticality of (19c), Chomsky (1986), Safir (1987), Saito (1991), among others, proposed that in (19c), the complement CP is raised rightward, i.e. extraposed, to a position outside the c-command domain of the matrix object to satisfy requirement (20). According to this proposal, (19c) is thus assigned the following structure:

(21)
Which men$_i$ did the police [$_{VP}$ warn t_i t_j] [$_{CP}$ that they were about to arrest e_i]$_j$?

Notice that in (21), after the CP extraposition, the parasitic gap is not c-commanded by the variable in the matrix object position thanks to the upward rightward movement of the CP.

Finally, by showing the contrast between (22a) and (22b), Lasnik (1993) argued that for infinite regress resolution in antecedent contained deletion (henceforth, ACD), the option of CP extraposition, originally proposed by Baltin (1987), is necessary.

(22)
a. *Mary stood near Susan, who Emily did not/as well [$_{VP}$ e].
b. ?Mary stood near everyone Emily did [$_{VP}$ e].

As shown in (23a), CP extraposition is available to a restrictive relative clause. However, as in (23b), it is not available to a non-restrictive relative clause.

(23)
a. A man arrived who was wearing a red hat.
b. *John arrived who was wearing a read hat.

Thus, Lasnik argues that we can apply CP extraposition to (22b), which has the restrictive relative clause, but not to (22a) with the non-restrictive relative clause. This is illustrated in (24).

(24)
a. *Mary [$_{VP1}$ stood near Susan] [$_{CP}$ who Emily did not/as well [$_{VP2}$ e]]
b. Mary [$_{VP1}$ stood near everyone] [$_{CP}$ Emily did [$_{VP2}$ e]]

In (24b), by CP-extraposition, the CP is crucially raised to the right out of VP_1. Then, after applying QR to *everyone* in (24b), we can successfully copy VP_1 onto the elided VP, resolving an infinite regress. The final well-formed LF representation that Lasnik proposed for (22b) is given below:

(25)
[$_{IP}$ [$_{NP}$ everyone]$_i$ [$_{IP}$ Mary [$_{VP1}$ [$_{VP1}$ stood near t_i]
[$_{CP}$ OP [Emily did [$_{VP2}$ stood near t_i]]]]]]

As Jun Abe points out in a personal communication, under Takano's theory, it is possible for us to account for the grammaticality of example (19c) and example (22b) by appealing to CP-extraposition in the covert component. This is because movement operations taking place in LF do not change the surface word order of a sentence in his theory. However, it is not immediately obvious under Takano's theory why movement operations involved for infinite regress resolution

in ACD display the properties of rightward movement, as shown below. Crucially, under his theory, linear order does not play any role in the N to λ computation.

Consider the following example:

(26)
John [$_{VP1}$ wanted PRO to [$_{VP2}$ criticize Susan, who Emily did not [$_{VP}$ *e*]]].

This example is ambiguous in that the elided VP can be interpreted either as the matrix VP as in (27a) or the embedded VP as in (27b).

(27)
a. John wanted to criticize Susan, who Emily did not [$_{VP}$ want to criticize]
b. John wanted to criticize Susan, who Emily did not [$_{VP}$ criticize]

Abe & Hoshi (1997b) observed that in contrast with (26), example (28a) is not ambiguous and that the elided VP can be construed as only VP_2 as illustrated in (28b):

(28)
a. John [$_{VP1}$ wanted Mary to [$_{VP2}$ criticize Susan, who Emily did not [$_{VP}$ *e*]]].
b. John wanted Mary to criticize Susan who Emily did not [$_{VP}$ criticize]

In examples (26) and (28a), a nonrestrictive relative clause is involved and thus, CP-extraposition cannot be used for infinite regress resolution. Abe & Hoshi thus proposed that in these two ACD cases, heavy NP shift of [$_{NP}$ *Susan, who Emily did not*] is used to resolve an infinite regress. The structures which Abe & Hoshi assigned to example (26) and (28a) are given in (29) and (30), respectively.

(29)
a. John [$_{VP1}$ wanted PRO to [$_{VP2}$ criticize t_i] [$_{NP}$ Susan, who Emily did not [$_{VP}$ *e*]]$_i$]
b. John [$_{VP1}$ wanted PRO to [$_{VP2}$ criticize t_i]] [$_{NP}$ Susan, who Emily did not [$_{VP}$ *e*]]$_i$

(30)
a. John [$_{VP1}$ wanted Mary to [$_{VP2}$ criticize t_i] [$_{NP}$ Susan, who Emily did not [$_{VP}$ *e*]]$_i$]
b. *John [$_{VP1}$ wanted Mary to [$_{VP2}$ criticize t_i]] [$_{NP}$ Susan, who Emily did not [$_{VP}$ *e*]]$_i$

In (29a), [$_{NP}$ *Susan, who Emily did not*] moves out of VP_2 by heavy NP shift, and thus, VP_2 can be copied onto the elided VP. In contrast, in (29b), [$_{NP}$ *Susan, who Emily did not*] undergoes rightward movement, crossing a nontensed clause. Hence, either the matrix VP or the embedded VP can be successfully copied onto the elided VP in (29b), which makes the matrix VP reading or the embedded VP reading available for the elided VP in (26).

On the other hand, in (30a), [$_{NP}$ *Susan, who Emily did not*] moves out of VP_2 and this movement operation allows VP_2 to be copied onto the elided VP. Importantly, however, (30b), where [$_{NP}$ *Susan, who Emily did not*] is moved out of a nontensed clause by heavy NP shift, is an ill-formed LF representation. This is so, because the movement operation involved in (30b) violates Postal's (1974) constraint which prohibits an element from moving rightward crossing a lexical subject. The effects of this constraint are witnessed in (31).

(31)
a. John wanted PRO to buy t_i until yesterday [$_{NP}$ a new Cadillac]$_i$.
b. ?*John wanted Mary to buy t_i until yesterday [$_{NP}$ a new Cadillac]$_i$.
cf. [$_{NP}$ A new Cadillac]$_i$ John wanted Mary/PRO to buy t_i until yesterday.

Showing structures in (29) and (30), Abe & Hoshi thus accounted for the differences in interpretation between (26) and (28a). (See Abe & Hoshi 1997b for more evidence that upward rightward movement is involved for infinite regress resolution in ACD. See also Jayaseelan 1991 and Abe & Hoshi 1997a, b for evidence which shows that English gapping also involves upward rightward movement to create a structure for copying.)

4. Conclusions

In this review, I focused on discussing some of the major proposals made in Takano's work, and I pointed out that there is some empirical evidence which shows that contrary to Takano's claims, accusative Case is abstract/structural Case in Japanese and that complex predicate formation in Japanese takes place in the core computation. Nonetheless, there is no doubt that Takano's work will be one of the most important pieces of work on parametric syntax because of the theoretical importance of his proposals and because of the fact that his theory accounts for many typological differences between English and Japanese with one single parameter. That is, light head ϰ has [−X] in English but not in Japanese.

Acknowledgments

I wrote this review when I was staying at Nanzan University as a guest academic researcher in 1997. I would like to thank Nanzan University, which provided me with such a valuable opportunity. I am very grateful to Jun Abe, Lisa Cheng, Yuji Takano and Mamoru Saito for their comments on an earlier version of this review.

References

Abe, J. & H. Hoshi (1997a). Gapping and P-stranding. *Journal of East Asian Linguistics* **6,** 101–136.

Abe, J. & H. Hoshi (1997b). A generalized rightward movement analysis of antecedent-contained deletion. *SOAS Working Papers in Linguistics*, SOAS, University of London.

Baltin, M. (1987). Do antecedent-contained deletions exist? *Linguistic Inquiry* **18,** 579–595.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1982). *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1986). *Barriers*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1993). A minimalist program for linguistic theory. In *The view from building 20*, K. Hale and S. J. Keyser (eds.), 1–52. Cambridge, Mass.: MIT Press.

Chomsky, N. (1994). Bare phrase structure. *MIT Occasional Papers in Linguistics* **5,** Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.

Chomsky, N. (1995). *The minimalist program*. Cambridge, Mass.: MIT Press.

Engdal, E. (1984). Why some empty subjects don't license parasitic gaps. In *The proceedings of the Third West Coast Conference on Formal Linguistics*. Stanford, Calif.: CSLI.

Fukui, N. (1986). *A theory of category projection and its applications*. Diss., MIT, Cambridge, Massachusetts. [A revised version published in 1995 as *Theory of projection in syntax*, Stanford, Calif.: CSLI.

Fukui, N. (1988). Deriving the differences between English and Japanese: A case study in parametric syntax. *English Linguistics* **5,** 249–270.

Fukui, N. (1995). The principles-and-parameters approach: A comparative syntax of English and Japanese. In *Approaches to language typology*, M. Shibatani and T. Bynon (eds.), 327–371. Oxford: Clarendon Press.

Fukui, N. & M. Saito (1996). Order in the theory of phrase structure and movement. Ms., University of California, Irvine and Nanzan University.

Harada, S.-I. (1973). Counter Equi-NP-deletion. *Annual Bulletin* **7,** Research Institute of Logopedics and Phoniatrics, University of Tokyo.

Hoji, H. (1985). *Logical form constraints and configurational structure in Japanese*. Diss., University of Washington.

Jaeggli, O. (1986). Passive. *Linguistic Inquiry* **17,** 587–622.

Jayaseelan, A.K. (1990). Incomplete VP deletion and gapping. *Linguistic Analysis* **20,** 64–81.

Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Kitagawa, Y. (1994). Shell, yolks, and scrambled e.g.s. In *Proceedings of NELS 24*, University of Massachusetts.

Koizumi, M. (1995). *Phrase structure in minimalist syntax*. Diss., MIT.

Kuno, S. (1973). *The structure of Japanese*. Cambridge, Mass.: MIT Press.

Kuroda, S.-Y. (1965). *Generative grammatical studies of the Japanese language*. Diss., MIT, Cambridge, Massachusetts.

Kuroda, S.-Y. (1978). Case-marking, canonical sentence patters, and Counter Equi in Japanese. In *Problems in Japanese syntax and semantics*, J. Hinds & I. Howard (eds.), 30–51. Tokyo: Kaitakusya.

Kuroda, S.-Y. (1981). Some recent issues in linguistic theory and Japanese syntax. *Coyote Papers* **2,** 103–122.

Kuroda, S.-Y. (1988). Whether we agree or not: A comparative syntax of English and Japanese. In *Papers from the Second International Workshop on Japanese Syntax*, W. Poser (ed.), 103–144. Stanford, Calif.: CSLI.

Kuroda, S.-Y. (1991). *Fundamental facts about Japanese causatives: Implications for morphology*. Presented at the Japanese Syntax Workshop, University of Rochester, New York.

Kuroda, S.-Y. (1993). Lexical and productive causatives in Japanese: An examination of the theory of paradigmatic structure. *Journal of Japanese Linguistics* **15,** 1–81.

Larson, R. (1988). On the double object construction. *Linguistic Inquiry* **19,** 335–391.

Lasnik, H. (1993). Lectures on minimalist syntax. *MIT Occasional Papers in Linguistics* **1,** Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.

Matsumoto, Y. (1996). *Complex predicates in Japanese: A syntactic and semantic study of the notion 'word'*. Stanford, Calif.: Kuroshio Publishers and CSLI.

Miyagawa, S. (1986). Restructuring in Japanese. In *Issues in Japanese linguistics*, T. Imai and M. Saito (eds.). Dordrecht: Foris.

Nemoto, N. (1993). *Chains and Case positions*, Diss., University of Connecticut.

Otani, K. & J. Whitman (1991). Verb-raising and VP-ellipsis. *Linguistic Inquiry* **22,** 345–358.

Pesetsky, D. (1995). *Zero syntax: Experiencers and cascades*. Cambridge, Mass.: MIT Press.

Pica, P. & W. Snyder (1995). Weak crossover, scope, and agreement in a minimalist framework. In *The Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. Stanford, Calif.: CSLI.

Poser, W. (1981). The 'Double-*o* Constraint': Evidence for a direct object relation in Japanese? Ms., MIT, Cambridge, Massachusetts.

Postal, P. (1974). *On raising*. Cambridge, Mass.: MIT Press.

Safir, K. (1987). The Anti-C-Command Condition on parasitic gaps. *Linguistic Inquiry* **18,** 678–683.

Saito, M. (1982). *Case marking in Japanese: A preliminary study.* Ms., MIT, Cambridge, Massachusetts.

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. Diss., MIT, Cambridge, Massachusetts.

Saito, M. (1991). Extraposition and parasitic gaps. In *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 467–486. Dordrecht: Kluwer.

Saito, M. (1994). Improper adjunction. In *Formal approaches to Japanese linguistics* 1. *MIT Working Papers in Linguistics* **24,** M. Koizumi & H. Ura (eds.), 263–293. Department of Linguistics and Philosophy, MIT, Cambridge, Massachusetts.

Shibatani, M. (1973). Semantics of Japanese causativization. *Foundations of Language* **9,** 323–373.

Takahashi, D. (1993). *On antecedent-contained deletion.* Ms., University of Connecticut.

Taraldsen, K. (1981). The theoretical interpretations of a class of marked extractions. In *Theory of markedness in generative grammar*. A. Belletti, L. Brandi, & L. Rizzi (eds.), 475–516. Pisa: Scuola Normale Superiore di Pisa.

Watanabe, A. (1993). *Agr-based Case theory and its interaction with the A-bar system*. Diss., MIT.

DRAWING THE LINE

by *Keren Rice*

Review of **Moraic for Macintosh** by *Cascadilla Press*

Moraic is a font designed to "create simple phonological structures right in your word processor!" (phrase on front of the user manual). The version of Moraic that is under review here is for the Macintosh. An IBM version of Moraic also exists, but the IBM version does not have the same range of possibilities that the Macintosh font does.

Moraic is, according to the user manual, "the phonological structure font" (page 1). This could mean a number of different things, as various kinds of structures exist under the general rubric of "phonological structure". In order to see just what is meant by "phonological structure", I will first examine just what it is that this font can do and then what it does not do.

In order to test the font, I decided to attempt to reproduce diagrams from *The Handbook of Phonological Theory* (John A. Goldsmith, editor; 1995). This book contains articles on a wide range of topics in phonological theory, and I thought that it would most likely contain the full range of structure types used by phonologists.

1. What can it do?

Moraic allows the user to make solid lines, dashed lines, and barred lines. The range of each of these is shown in (1).

(1)

solid lines:

(single keystroke 1 2 3 g 8 9 0)

dashed lines:

(shift — keystroke 1 2 3 g 8 9 0)

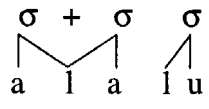
barred lines:

(single keystroke q w e b i o p)

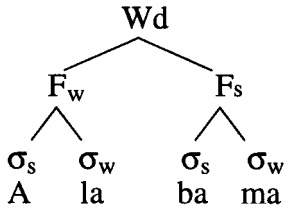
These symbols represent left and right branches at different angles and a vertical line. This latter is found with most Macintosh fonts, but it is a nice addition to Moraic as vertical lines and slanted lines often occur together in the same diagram (e.g. as in the examples in (2) and (3) below), and it is not necessary to change fonts to make two angled lines with a vertical line between. Using these keystrokes alone, it is possible to create on the Macintosh many of the diagrams found in the *Handbook*. I will begin with some relatively simple ones, shown in (2).

(2)

from Broselow 177



from Broselow 186

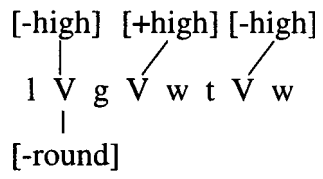


from Kager 369

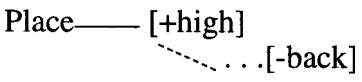
These representations are very similar to the originals in the *Handbook*. They are created in a simple way, using just the lines shown on the first line in (1).

Some feature geometry diagrams and phonological rules are equally easy to prepare, as in (3).

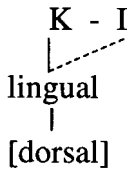
(3)



from Steriade 133



from Steriade 135



from Clements & Hume 291

These representations are all easy to create, and are very attractive.

Equally easy to do is crossing lines. These involve a different type of keystroke than I have introduced so far, non-advancing lines. The non-advancing lines essentially create an overlay so that the cursor does not advance.

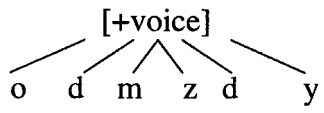
(4)

from Clements and Hume 266

(key 7 followed by 2)

Non-advancing lines serve another purpose. They allow one to create structures in which there is more than one branch on a side of a diagram. Using just the keystrokes given in (1), I designed the following structure.

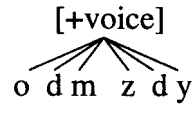
(5)



cf. Steriade 122

This is quite unlike the structure in the *Handbook*, however. By using non-advancing lines, I was able to create a structure like that found in the *Handbook*.

(6)

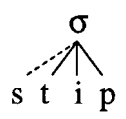


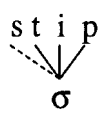
from Steriade 122

There is thus a considerable amount that can be done using Moraic.

In addition to the font called Moraic, another font is included, Moraic Flip. As the user manual points out, this font allows the user to change the orientation of the font. I made (7a) using Moraic and then created (7b) by highlighting the lines in Moraic and changing the font to Moraic Flip. I then copied and pasted the other lines.

(7)

a. 

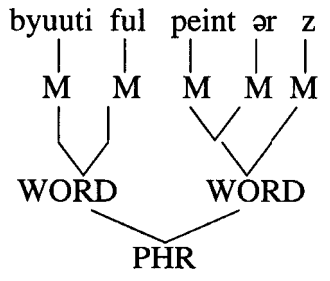
b. 

2. What can't it do?

While Moraic is a very flexible font, it is not without its limitations; I now discuss some of these.

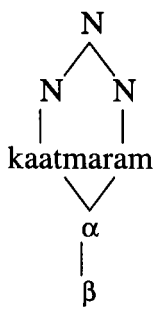
First, Moraic only allows for a few number of angles to be drawn. In some cases, these angles are not exactly right. This can be seen in the structures in (8) and (9).

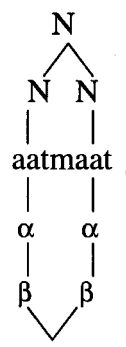
(8)



(after Mohanan 30)

(9)



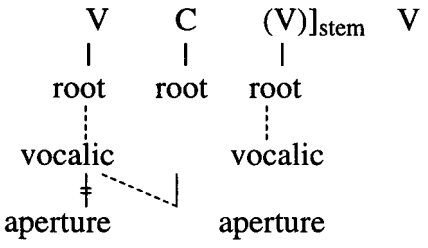


(after Mohanan)

These are not bad looking structures, but they are not perfect — the branches do not line up exactly over the symbols. I tried changing the size of the fonts; this may work in some cases, but not necessarily in all cases. Thus, Moraic is good for things that are relatively standard sizes (e.g. syllable structure, moraic structure), but not when, as in the examples above, words are included that necessitate angles that fall between those given. It may be possible to solve this problem by using Arboreal, another Cascadilla font which is designed to create syntactic structures.

A feature geometry structure shows the problems that arise from this inflexibility in creating angles even more dramatically. Note that in the diagram in (10), taken from Clements and Hume, there should be a vertical line connecting the right "vocalic" with the right "aperture" and there should be a dashed line linking the first "vocalic" with the second "aperture." The dashed line in (10) represents the longest dash line that it is possible to make with Moraic — it is impossible, to my knowledge, to make a long enough line to achieve the appropriate linking.

(10)



after Clements and Hume 282

This is not just a problem in this diagram, but in many other structures where a fair amount of material intervenes between two elements to be linked up by a line.

Feature geometric representations present two additional problems. First, in many of the papers in the book that employ these structures, the lines starting at a particular mother node are not equal in length: one might take up four lines while one takes up six or more lines, for instance. A drawing program is necessary to create these tree structures; it cannot be done with Moraic, or with Arboreal.

The second problem comes in diagrams with more than three diagonal lines on one side. Consider the diagram in (11).

(11)

after van der Hulst & van de Weijer 506

This is a very attractive diagram. However, it is not, in fact, exactly the structure that van der Hulst & van de Weijer have in their paper — their diagram has one additional diagonal line on the right hand side. I was unable to produce a diagram with four diagonal lines on the same side.

3. Learnability

It took me very little time to learn to use Moraic. Only part of the top two rows, shift, and the key "g" are used and the keyboard is set up in such a way that the diagonals are intuitive — keys on the left side of the keyboard produce lines starting at the left and moving up to the right; those on the right hand side produce lines starting at the left and moving down to the right. The most difficult part was probably getting the spacing right so that the lines match up with the elements with which they are associated, and this is a relatively simple problem to deal with so long as the terminals at the end of a line are not too long. I do not think that any user will complain about a steep learning curve with Moraic. After doing the diagrams that appear in the review, I feel like I no longer need to check the layout to remember where a particular symbol is.

4. The user manual and keyboard layout

The user manual provides instructions for installing Moraic with either System 6 or System 7 for Macintosh, instructions for using Moraic, the keyboard layout, some technical notes about mixing fonts, and some ideas for troubleshooting. In addition, a reference card summarizing the keyboard layout and the use of the non-advancing branches comes with the font. This is a handy guide for anyone who has trouble remembering where on the keyboard non-standard font symbols are found. I kept the reference card next to my computer while I learned the layout, and I will keep it handy in case I forget where something is or just how to combine advancing and non-advancing lines.

5. Support

Cascadilla Press, the developer of Moraic, offers unlimited technical support. For instance, while Moraic exists for both the Macintosh and Windows, the developers of the font point out that the conversion between these formats is not trivial and they suggest that they be contacted for assistance in doing a conversion. They provide the user with their snail-mail address, their e-mail address, and their telephone number.

6. Conclusion

Moraic is a font that I am likely to make good use of to create a number of different phonological structures. Used together with Arboreal, a second font designed to create trees with lines at different angles, these fonts should meet a wide range of needs for both phonologists and syntacticians, although structures remain that will have to be made with a drawing program. The font is reasonable in cost, well-supported, and is another device that will help linguists create more professional looking camera-ready copy.

Reference

Goldsmith, John A. 1995. *The handbook of phonological theory*. Oxford: Blackwell.

CONFERENCE REPORTS

Comparative Germanic Syntax 13
by *Kleanthes K. Grohmann*

NELS 28
by *Jan-Wouter Zwart*

The Thirteenth Comparative Germanic Syntax Workshop

Cornell University, Ithaca
July 18–19, 1997

In the heat of the summer, the 13th Comparative Germanic Syntax Workshop took place at Cornell University. We coped with the temperatures in an air-conditioned auditorium that despite its size gave the event a family atmosphere, having plenty of cold refreshments. There were seven talks scheduled for the workshop (on Friday and Saturday afternoons) and we all had the opportunity to attend the four Germanic talks of the HPSG-conference that took place at the same time on Saturday morning. In the following, I will just review the CGSW-talks.

After a warm reception, Norvin Richards (MIT) presented his proposal of the "AgrEverythingPhrase" (which looks shorter in a proper typeset, replacing the lexical item with logical notation!). Implementing the idea of multiple specifiers (Kuroda 1988; Chomsky 1995) into a minimalist framework that also distinguishes nesting and crossing paths (Pesetsky 1982), Richards' proposal yields interesting results for a variety of syntactic constructions. In particular, he looked at Germanic object shift and multiple *wh*-movement and showed that movement to multiple specifiers takes place if (i) the paths obligatorily cross and (ii) the highest available mover moves first; otherwise, (iii) movement targets the specifiers of distinct heads if the paths nest. In short, points (i), (ii) and (iii) follow from featural Cyclicity and Shortest Move (Kitahara 1994). Interaction of these principles and a formalized theory of multiple specifiers could account for object shift (crossing paths of argument-movement resulting in movement to multiple specifiers of one projection: AgrEverythingP) as well as multiple *wh*-movement (Bulgarian *wh*-elements cross their paths too and target multiple specifiers of a single head).

Next were Artemis Alexiadou (ZAS Berlin) and Elena Anagnosopoulou (Tilburg/MIT) with an investigation of the behaviour of "Postverbal Subjects" in Germanic as opposed to Celtic and Romance. The different patterns that are observed (either subject or object or both raising out of VP) can be accounted for by postulating that in the presence of an object, arguments must raise overtly to check formal features. The cross-linguistic differences can be seen as interaction of different requirements of the LF- or PF-component, the TP-parameter and object shift. In particular, Alexiadou and Anagnostopoulou propose a ban on multiple covert movement for case-checking. Potential counter-evidence coming from Greek, Romanian and Spanish can be explained by the presence of clitic-doubling in these languages.

The Friday session ended with a double-bill on multiple interrogatives in German. Martina Wiltschko (British Columbia/Vienna) started off on the topic "Superiority in German" arguing against the standard claim that German lacks superiority effects. As is well known, D-linking (Pesetsky 1987) can counter such effects, and Wiltschko's proposal is that scrambling in German serves as an additional trigger for D-linking. Evidence that superiority effects do obtain in German comes from 'was-für NPs' that are inherently non-D-linked; secondly, 'wer' or 'was' (which can be either D-linked or not) show superiority effects just in case they appear in a scrambled position (using adverbs as diagnostics where a *wh*-argument following the adverb indicates non-D-linking). In short, superiority effects arise if the *wh*-element is not D-linked; furthermore, non-D-linking can be forced for testing. Thus scrambling triggers D-linking which in turn counters superiority effects.

Kleanthes K. Grohmann (Maryland) in contrast, defended the standard claim that there is no "German Superiority". It became apparent in discussion before, during and after both talks, that there is a crucial difference in German dialects which may be split into north and south: southern dialects (including or predominantly Austrian) do indeed show superiority effects, while northern do not. Grohmann (from the north) judges basically all of Wiltschko's examples (based on Austrian judgements) as fine. Thus, while Wiltschko's analysis seems to be a fine approach to one variety of German, Grohmann tries to fill in the need to account for the other. Grohmann shares with Wiltschko the intuition that scrambling can counter superiority effects. In particular, he claims that the MLC — which is arguably the minimalist version of the original superiority condition (Kitahara 1994) — is never violated under his assumption (Grohmann 1996) that scrambling over the subject in German is related to topicalization. He also observes that the interpretation of multiple *wh*-elements in German differs from single *wh*-expressions in a restricted quantified reading. He thus claims that multiple *wh*-phrases are topics, drawing from similar claims about Chinese *wh*-topics (Wu 1996) and other cross-linguistic data.

The second session was opened by Elly van Gelderen (Arizona State) who presented an account for unexpected behaviour of "Anaphora in Middle English" which do not show the complementarity with respect to pronouns but act as pronominal objects; furthermore, when used, reflexives often appear only as objects of prepositions. The account is based on the feature content of pronouns (phi-features of simple pronouns are less specified) and on inherent Case coupled with the change of 'self' from adjective to noun (reflexives may not appear in a structurally Case-marked position).

Ken Safir (Rutgers) fittingly followed with some thoughts on the "Symmetry and Unity in the Theory of Anaphora", looking at Modern English. Safir set out to distinguish reflexivity from binding, justify principles of reflexive and local binding interpretation and eliminate principles A and B. In this respect, he proposes a symmetric system of binding as well as a unification of the interpretation of semantic reflexivity. He also reformulates the conditions under which an anaphor may be employed as a logophor. Safir proposes syntactic and semantic principles of binding encouraged by missed generalizations on the one hand and internal inconsistencies on the other as proposed by Reinhart and Reuland (1993).

Concluding the the 13th CGSW, Peter Svenonius (Tromso) talked about "Particle–Verb Incorporation and Participle Agreement in Scandinavian". Extensive data from all Scandinavian languages (and a number of dialects) seem to support an analysis of particles as complements of verbs that incorporate (den Dikken 1995). Differences in behaviour, such as lack of incorporation in some dialects, follow from an intricate connection of A-movement and head movement. Overt participial agreement means that this language has an active Agr, forcing (possibly covert) specification of phi-features on Agr if NP is in SpecAgrP. Svenonius ties in possibilities of incorporation with participial agreement, predicting that active participial Agr makes overt incorporation more economical and hence the only grammatical option.

The small number of talks at the 13th CGSW was acceptable not only due to the option of listening to four more talks at the HPSG-conference but also because the content of all talks was very satisfying, and there were plenty of stimulating discussions during and after the talks. The team at Cornell did a good job organizing the workshop and deserve all our gratitude. We are all full of hopes that more abstracts of high quality will be submitted, allowing for more talks at the next workshop which will take place in Lund in January 1999.

This Conference Report is published simultaneously in the Germanic Generative Syntax Newsletter, Vol. 13.

NELS 28
University of Toronto
24–26 October 1997

Every year when I apply for a travel grant that will allow me to attend NELS, I tell my employers that NELS is the most important theoretical linguistics conference in North America, and the second most important one in the world. It does not always work. I've had to miss out on several conferences in the past few years, and of some of those I've heard that "NELS wasn't so good". I've even attended some that weren't so good. It seems that lately, that comment has become more frequent, as if to confirm the nagging suspicion that things aren't what they used to be. Well, this year I was lucky enough to be able to cross the Atlantic again and see and hear the conference with my own eyes and ears. And it was a fairly interesting conference, with at least 4 or 5 extremely interesting papers. There were also a number of stimulating poster sessions, and there were about 200 linguists gathered to discuss any aspect of anyone's ongoing research with. The organization was smooth, and the general atmosphere was extremely pleasant. So all the potential was there for a conference of global interest. Still I don't doubt that many came away from NELS 28 with the feeling that it wasn't like the old days. If so, that would be entirely due to the attendants that stayed home — especially the senior linguists and the between-junior-and-senior linguists. I was amazed to find that very few attendants of over 40 and/or with permanent positions were present. Of course, coming from Europe it's not my business, but I'd like to say something corny in the range of: NELS is what you make it, where *you* is not the speakers or the organizers or the selection committee or even the reviewers, but the attendants of all ranks and persuasions. That said, let me go on to discuss five important papers from NELS 28 in the remainder of this report.

Two of those were about binding. **Johan Rooryck and Guido Vanden Wyngaerd** discussed **The 'Self' as Other: a Minimalist Approach to *zich* and *zichzelf* in Dutch**. They made sense of subtle differences in Dutch involving the reflexive pronoun *zich* and the anaphor *zichzelf*, which is essentially a reflexive with a focus element *zelf* 'self'. The focus element prevents the reflexive from doing its normal work, which is to link up with Tense. Being linked up with tense, *zich* pairs the time slices of which the (animate) subject is composed (cf. Carlson 1977) with the time slices making up the situation expressed by the verb. The effect of this pairing is that all the time slices of the subject end up being simultaneous. The focus marker *zelf* blocks this process, with the result that *zichzelf* is represented as a pure object: an 'Other'. The temporal dissociation effected by the focus element explains the felicity of *Jan zag zichzelf op film* [John saw SE-self on film] vs. *?*Jan zag zich op film* [John saw SE on film] and many other subtle — but striking — effects of interpretation. (The analysis ties in very nicely with K. Jayaseelan's discussion of morphologically complex anaphors in *Studia Linguistica* 51.2, 1997.) **Carol Tenny** presented a paper on **Pronominal Coreference and Point of View**, discussing the contexts where pronouns appear to be locally bound, in violation of Principle B of the Binding Theory. These contexts are presented by locative and directional PPs (*The men found a smokescreen around them*) and representational NPs (*Martha doesn't mind a joke about her every now and then*). What these contexts have in common is that they both constitute a *Point of View Domain*, leading to the reformulation of Principle B as requiring that a pronoun be free in its local point of view domain. There was extensive discussion in the paper of alternative ways of redefining the binding domains, among which the approach based on argument structure, of Reinhart and Reuland, *Linguistic Inquiry* **24.4,** 1993. **Željko Bošković** rounded off a session on locality conditions within the minimalist framework (**LF-Movement and the Minimalist Program**). He explored the hypothesis, generated by Chomsky's conception of LF-movement as feature movement (hence, head movement), that LF-movement is actually more constrained than overt movement. This hypothesis was confirmed by phenomena in several languages (French, German, and English), where it could be shown that covert *wh*-movement is blocked by the presence of a complementizer and/or negation, whereas overt wh-movement is not (for example, French *Tu as vu qui?* [you have seen who] vs. *?*Marie croit que tu as vu qui?* [Marie thinks that you have seen who]). **Andrew Simpson** presented a highly convincing analysis of **Focus, Pre-supposition, and Light Predicate Raising in South-East Asian**. In the relevant languages (Thai, Vietnamese, Cambodian, Hmong, and Cantonese), a single modal verb meaning 'can/be able to' invariably appears VP-finally, unlike other modal verbs, which precede the predicate. Simpson showed that the modal verb in question is a Middle Chinese borrowing, taking a predicate complement to its right, but triggering raising of the predicate to its overt syntax position to the left of the modal verb. The complement of the modal verb may also contain a Focus Phrase, which focused material from the embedded predicate may raise to, with the result that the focused material is stranded by the raised predicate, yielding the attested word order *predicate–modal–focused material*. The analysis, which is fully compatible with Kayne's 1994 conjecture regarding the universal order of specifier, head, and complement, presents an interesting addition to the literature on ways of expressing focus. **Henry Davis and Lisa Matthewson** added another fascinating chapter to the increasing body of work on Salish (Sťát̓imcets) with their paper **Determiners, Tense, and the Entity/Event Parallel**. They proposed that although the syntax of noun phrases and verb phrases in Salish are different, the functional domain of verbs and nouns are totally identical (instead of merely parallel). Verbal and nominal arguments are introduced by phonologically identical elements (*ti..a* and *ku*), which yield the same semantic interpretations in each case, and both types of arguments have the same syntactic distribution. The functional domain in Salish is a DP, headed by a D which doesn't care whether its complement is an event or an entity (an *unselective D*). The paper also contained detailed discussion of morphological and semantic differences between events and entities in Salish which, in spite of the noted identical functional domains, nevertheless do exist. For example, verbal DPs may contain a prefix *s* associated with subject morphology. This prefix was argued to head an additional functional projection FP which *is* sensitive to the distinction between events and entities. The relevance of these observations is that the head of FP apparently represents certain features usually associated with Tense, but not Tense itself. This is makes it possible to study the question which of the features usually associated with Tense actually require a VP, and which are just accidentally connected with verbal syntax in Indoeuropean languages.

Other presentations included: Eric **Reuland**, 'Deriving c-command in binding,' Brian **Agbayani**, 'Generalized Pied-Piping and Island Constraints,' Masanori **Nakamura**, 'Global Issues,' Jan-Wouter **Zwart**, 'Rethinking Subject Agreement in Swahili,' Heidi **Harley** and Rolf **Noyer**, 'Mixed Nominalizations, Short Verb Movement, and Object Shift in English,' Paul **Law**, 'A Unified Analysis of P-Stranding in Romance and Germanic,' Andrea **Massar** and LouAnn **Gerken**, 'Abstract output: an OT analysis of children's omissions from prosodically complex structures,' Mitsuhiko **Ota**, 'The emergence of the unmarked in early prosodic structure,' Brude **Tesar**, 'Using the mutual inconsistency of structural descriptions to overcome ambiguity in language learning,' Beverley **Goodman**, Ponapean weight and [+/- consonantal]', Trisha **Causley**, Variable Markedness and fixed hierarchies,' Daniel **Silverman**, 'Alveolar stops in English and the nature of allophony,' Stuart **Davis** and Gina **Toretta**, 'An Optimality-Theoretic account of vowel lengthening and geminate throwback in Trukese,' Eugene **Buckley**, 'Integrity and correspondence in Manam double reduplication,' Dan **Karvonen** and Adam **Sherman**, 'Opacity in Icelandic: a sympathy account,' Thomas **Ernst**, 'Scope based adjunct licensing,' André **Meinunger**, 'A monoclausal structure for (pseudo)cleft sentences,' Benjamin **Shaer**, 'Adverbials, functional structure, and restrictiveness,' Sungeun **Cho**, 'A new analysis of Korean inalienable possession constructions,' Roumyana **Izvorski**, 'Non-indicative wh-complements of possessive and existential predicates,' Yael **Sharvit**, 'Possessive wh-expressions and reconstruction,' Martha **McGinnis**, 'Locality and Inert Case,' Murat **Kural**, 'Passives without argument incorporation,' Simin **Karimi** and Anne **Lobeck**, Specificity effects in English and Persian,' Carson T. **Schütze**, Subject Case, temporal specification, and the typology of (Small) Clauses,' Marc **Authier**, 'When syntax overrules semantics,' Elena **Anagnostopoulou**, Ken **Hale**, Sabine **Iatridou**, and Roumyana **Izvorski**, 'On the morpho-syntax of the Perfect and how it relates to its meaning,' Jianxin **Wu**, 'Topic, floating quantifiers, and partitivity,' and Anna-Maria **Di Sciullo** and Carol L. **Tenny**, 'Modification, event structure and the word/phrase asymmetry.'

The proceedings will be published, as always, by the Graduate Linguistics Student Association of the University of Massachusetts at Amherst, and will include the papers of the conference, including the alternates (Wilfried **Lechner** and Arthur **Stepanov**), as well as a selection from the papers presented in the poster sessions.

A RESPONSE TO ALAN PRINCE

by *Morris Halle & William Idsardi*

"We write to respond to Alan Prince's letter in *Glot International* 2-6, and to continue the discussion of Optimality Theory (OT) following on Luigi Burzio's article in *Glot International* 1-6."

Continued from page 1

dialect is to insert coda r's in hiatus. The Elsewhere Condition is important because it provides specific restrictions on the class of possible grammars through intrinsic ordering principles, and also provides restrictions on the application of rules within a grammar. It is the core of a theory of rule-interaction in grammar. To the extent that examples of the Elsewhere Condition in different languages can be understood in this manner, these examples provide additional empirical support for the Condition.

According to Prince, a superiority of OT over rule-based theories is that OT "needs no Elsewhere Condition to guide it" (p.231), whereas "rule-ordering theory with an adjoined Elsewhere Condition forms a redundant and centaur-like composite." Its lapidary wording notwithstanding, this statement is not true. The fact is that both OT and rule-based theories must — and do — include machinery dealing with the interaction of their basic empirical statements — whether rules or constraints. The Elsewhere Condition is part of this additional machinery in rule-based theories. In OT the additional Elsewhere-Condition-like machinery can be found in various proposals such as constraint conjuction (Smolensky 1995). Smolensky states that two constraints A and B, can be locally conjoined to create a new constraint A&B, and that this new constraint A&B will always be ranked higher than either of its constituent constraints. This is clearly a principle of intrinsic ordering for constraints, one which is a direct analog of the Elsewhere Condition. Obviously the conjoined constraint A&B is necessarily more specific than either A or B, and therefore it outranks the single constraints A and B. One might think that Smolensky could claim, pace Prince, that UG simply contains all possible conjoined constraints, and that there is no need for intrinsic ordering principles, the more specific constraints being shadowed by more general ones when outranked. But Smolensky specifically intends that constraints can be conjoined with themselves, A&A, to make double violations count more than just two violations of a single constraint. In particular, he employs this technique in OT syntax to enforce Subjacency effects when crossing two or more bounding nodes, while allowing movement across a single bounding node. Since constraints can be conjoined with themselves there is an infinite number of possible constraints, and therefore it is impossible to claim that UG has simply pre-calculated all possible conjoined constraints. In addition, conditions on constraint evaluation analogous to the disjunctive application of rules have also been proposed, for example, by Hammond (1997). Hammond proposes in an OT model of speech processing that if constraint A is ranked above constraint B and it is impossible to evaluate A at some processing stage (from lack of information) then constraint B goes unevaluated as well, even if there is enough information available to evaluate B, leading to ties between candidates. In sum, it is clear that OT employs principles exactly like the Elsewhere Condition in constraint ordering and evaluation, and is better off because of it, exactly the opposite of what Prince claims.

While the preceding serves to clear up some misunderstandings which may arise from Prince's letter, it does not deal with the important issue of how the disagreement between proponents of OT and those of ruled-based theories might be resolved. In this concluding section we briefly comment on this matter.

In our view the situation is quite straightforward: the phonological literature includes a fair number of insightful accounts of facts of considerable complexity, and all of these accounts are rule-based. One such example is the above-mentioned lengthening and shortening of English stressed vowels and their interaction with other phenomena of English phonology such as stress assignment, vowel shift, velar softening, etc.; in short, the data discussed in Chomsky & Halle (1968), Halle & Mohanan (1985), etc. Another is spirantization in Tiberian Hebrew (Chomsky 1951) and its relationship with other phenomena in the language, many of which were discussed in Prince (1975), Rappaport (1984), Idsardi (1997, 1998) and elsewhere. Yet a third example is the treatment of the fundamental phenomena of Spanish phonology as discusssed in Harris (1993, 1995) and Harris & Kaisse (1997). For each of these there exist well worked out rule-based accounts, but it has not been shown that these facts can be dealt with properly within OT. Until this challenge is met, i.e., until it is shown that OT can deal with the really hard problems in phonology, OT's claim to be the theory of phonology is without much force.

It is significant that in spite of numerous references to the English lengthening and shortening facts in his letter, Prince does not get down in the trenches and attack particular facets of the rule-based analysis in order to demonstrate the superiority of an OT alternative. Although he cites Bakovic (1996) — an unpublished research paper not available to us as we write this letter — Prince does not deal with any of its details, does not indicate how it explains the interaction between shortening and lengthening, and cites no new facets of the analysis which go beyond what has already been discovered. In short, Prince adduces no new evidence showing that progress has been made in this domain using Optimality Theory.

In the light of the preceding, the text in the box on p. 23 of *Glot International* 2–6 is incorrect. Since Prince does not discuss the details of the rule-based solution, there is no basis for the assertion that "various misinterpretations and flaws of logic combine to nullify [Halle's] qualms." The qualms remain in full force until someone comes along and provides an OT account of complex facts such as those mentioned above, because as even the least experienced bettor knows, you can't beat something with nothing.

Morris Halle, MIT
William J. Idsardi, University of Delaware

References

Chomsky, Noam (1951). *Morphophonemics of Modern Hebrew*. MA dissertation, University of Pennsylvania. [New York: Garland Press 1979.]

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

Chomsky, Noam & Morris Halle (1968). *The sound pattern of English.* New York: Harper and Row. [Cambridge: MIT Press 1991.]

Halle, Morris & W.J. Idsardi (1997). r, Hypercorrection, and the Elsewhere Condition. In *Derivations and Constraints in Phonology,* Iggy Roca (ed.), 331–348. Oxford: Oxford University Press.

Halle, Morris & K.P. Mohanan (1985). Segmental phonology of Modern English. *Linguistic Inquiry* **16,** 57–116.

Hammond, Michael (1997). Optimality Theory and prosody. In *Optimality Theory: An overview,* Diana Anchangeli & D. Terence Langendoen (eds.), 33–58. Cambridge, Mass.: Blackwell.

Harris, James W. (1993). Integrity of prosodic constituents and the domain of syllabification rules in Spanish and Catalan. In *The view from building 20,* Kenneth Hale & Samuel J. Keyser (eds.), 177–193. Cambridge, Mass.: MIT Press.

Harris, James W. (1995). Projection and edge marking in the computation of stress in Spanish. In *A handbook of phonological theory,* John A. Goldsmith (ed.), 867–887. Cambridge, Mass.: Blackwell.

Harris, James W. & Ellen Kaisse (1997). *Palatal vowels, glides, and obstruents in Argentinian Spanish.* Ms., Department of Linguistics, MIT and Department of Linguistics, University of Washington.

Idsardi, William J. (1997). Phonological derivations and historical changes in Hebrew Spirantization. In *Derivations and constraints in phonology,* Iggy Roca (ed.), 367–392. Oxford: Oxford University Press.

Idsardi, William J. (1998). Tiberian Hebrew spirantization and phonological derivations. To appear in *Linguistic Inquiry*.

Prince, Alan (1975). *The phonology and morphology of Tiberian Hebrew.* Diss., MIT.

Prince, Alan & Paul Smolensky (1993). *Optimality Theory.* Ms., Department of Linguistics, Rutgers University.

Pullum, Geoffrey (1976). The Duke of York Gambit. *Journal of Linguistics* **12,** 83–102.

Rappaport, Malka (1984). *Issues in the phonology of Tiberian Hebrew.* Diss., MIT.

Smolensky, Paul (1995). *On the internal structure of the constraint component Con of UG.* Talk presented at UCLA, April 7, 1995.