Formal and Functional Aspects of Abstract Analysis

Gereon Müller, IDS Mannheim

Workshop on Functionalism and Formalism, Leipzig, May 19, 2004

1. Introduction

Four dichotomies:
(i) Formal vs. functional principles
(ii) Language-specific vs. language-independent principles
(iii) Abstract vs. surface-oriented principles
(iv) Innate principles vs. principles that are learned

Main claims:
(i) An abstract analysis can reveal intriguing patterns in the data that cannot be discovered under a surface-oriented analysis.
(ii) The principles underlying these patterns must are highly general and abstract; but they may be language-specific or language-independent, they may be of a purely formal nature of functionally grounded, and they may be innate or not: These issues are orthogonal.
(iii) A case can be made for abstract, language-independent principles that are ultimately functionally grounded but still innate.

(1) *Two case studies*:
   a. Morphology: The analysis of fusional nominal inflection markers in German and Russian reveals an abstract *Iconicity* principle.
   b. Syntax: The analysis of displacement phenomena in English, Icelandic, Danish, German, and Japanese reveals an abstract *Minimality* principle.

Observation:
Iconicity and Minimality are both (a) highly abstract, (b) innate (i.e., unlearnable on the basis of empirical evidence alone), (c) language-independent, and (d) ultimately functionally grounded (motivated by computational efficiency).
(Under this view, it is in fact not the case that "strictly speaking, [functional] justifications are irrelevant in a theory that assumes innate constraints", as Haspelmath (1999) puts it.)

2. Background

Principles and Parameters Theory (Chomsky (1981)ff.):
Core principles are assumed to be abstract, innate, langage-specific and not functionally motivated.

(2) *Empty Category Principle*:
   A trace is properly governed (i.e., antecedent-governed or lexically governed).

(3) *Superiority effects in English*:
   a. (I wonder) who$_1$ t$_1$ bought what$_2$
   b. *(I wonder) what$_2$ who$_1$ bought t$_1$

Assumption:
LF movement of *who*$_1$ in (3-b) leaves a trace t$_1$ that is neither antecedent-governed nor lexically governed; in contrast, the S-structure trace t$_1$ in (3-a) is antecedent-governed.

Minimalist program (Chomsky (1995, 2000, 2001, 2004):
The language faculty is optimally designed so as to meet the demands imposed by the semantic (conceptual-intentional) and phonological (more generally, sensorimotor) interfaces: "We need no longer assume that the means of generation of structured expressions are highly articulated and specific to language. We can seriously entertain the possibility that they might be reducible to language-independent principles, whether or not there are homologous elements in other domains. We can, in short, try to sharpen the question of what constitutes a principled explanation for properties of language, and turn to one of the most fundamental questions of the biology of language: To what extent does language approximate an optimal solution to conditions that it must satisfy to be usable at all, given extra-linguistic structural architecture?" [...] We can regard an explanation of properties of language as *principled* insofar as it can be reduced to properties of the interface systems and general considerations of computational efficiency and the like." (Chomsky (2004, 7-8)).

Consequence:
Grammatical principles must be motivated by (a) properties of the conceptual-intentional or sensorimotor interfaces, or (b) computational efficiency. Principles like the Empty Category Principle are not motivated in either way, and must be abandoned.

Note:
The potential tension of explicitness and economy, germane to much functionally oriented work, is acknowledged in Chomsky (2004, 10). In the context of motivating the copy theory of movement, Chomsky notes:
"If language is optimized for communicative efficiency, we would expect all [copies in a movement chain] to be spelled out: that would overcome many of the filler-gap problems faced by processing models. If language is optimized for satisfaction of interface conditions, with minimal computation, then only one will be spelled out, sharply reducing phonological computation."

3. Iconicity in Morphology

(4) *Iconicity*:
   Similarity of form implies similarity of function.

Note:

In Haspelmath's (2003) taxonomy involving seven kinds of iconicity, this qualifies as 'Iconicity as correspondence of markedness/complexity'.

3.1. Pronominal Inflection in German

(5) *Pronominal inflection in German*:

    a. NOM: dieser Tee, dieses Bier, diese Milch, diese Gläser

    b. ACC: diesen Tee, dieses Bier, diese Milch, diese Gläser

    c. DAT: diesem Tee, diesem Bier, dieser Milch, diesen Gläsern

    d. GEN: dieses Tees, dieses Bieres, dieser Milch, dieser Gläser

P_1: *Standard paradigm of pronominal inflection in German*

dies ('this')	MASC.SG.	NEUT.SG.	FEM.SG.	PL.
NOM	r	s	e	e
ACC	n	s	e	e
DAT	m	m	r	n
GEN	s	s	r	r

Observation (Wiese (1996, 1999)):

An abstract analysis of pronominal (determiner) inflection in German that relies on a decomposition of case (Jakobson (1962)) and gender/number features (such that natural classes of cases and natural classes of genders/numbers can be formed that inflection markers can refer to) accounts for many instances of syncretism in a systematic way. Moreover, an underlying pattern of iconicity is revealed.

(6) *Case features*: [±oblique], [±governed] (cf. Bierwisch (1967)):

    NOMINATIVE: [–obl,–gov]

    ACCUSATIVE: [–obl,+gov]

    DATIVE: [+obl,+gov]

    GENITIVE: [+obl,–gov]

(7) *Gender and number features*:

    MASCULINE: [+masc,–fem]

    FEMININE: [–masc,+fem]

    NEUTER: [+masc,+fem]

    PLURAL: [–masc,–fem]

Note:

Wiese calls the decomposed gender and number features [±standard] and [±special] (not [±masc] and [±fem]).

(8) *Underspecified paradigm of pronominal inflection markers in German*

    1. /m/ ↔ [+masc,+obl,+gov]   (DAT.MASC.SG./NEUT.SG.)

    2. /s/ ↔ [+masc,+obl]   (GEN.MASC.SG./NEUT.SG.)

    3. /s/ ↔ [+masc,+fem]   (NOM./ACC.NEUT.SG.)

    4. /n/ ↔ [+masc,+gov]   (ACC.MASC.SG.)

    5. /r/ ↔ [+masc]   (NOM.MASC.SG.)

    6. /r/ ↔ [+obl,+fem]   (DAT./GEN.FEM.SG.)

    7. /n/ ↔ [+obl,+gov]   (DAT.PL.)

    9. /r/ ↔ [+obl]   (GEN.PL.)

    10. /e/ ↔ [ ]   (NOM./ACC.FEM.SG./PL.)

Note:

Competition arises because of underspecification of markers. Choice of the right marker for a given context follows from the Specificity Principle.

(9) *Specificity Principle*:

    For any given fully specified syntactic context, choose the most specific marker that is compatible with it.

    a. A marker M_i is more specific than a marker M_j if the set of morpho-syntactic features of M_j is a proper subset of the set of morpho-syntactic features of M_i.

    b. If the sets of morpho-syntactic features of two competing markers have the same cardinality, M_i is more specific than M_j if it has higher-ranked features on the hierarchy [+masc] > [+obl] > [+fem] > [+gov].

P_2: *Interaction of pronominal inflection markers in German*

dies ('this')	MASC.SG.	NEUT.SG.	FEM.SG.	PL.
NOM	$\mathbf{r}^5$, e^9	$\mathbf{s}^3$, r^5, e^9	$\mathbf{e}^9$	$\mathbf{e}^9$
ACC	$\mathbf{n}^4$, r^5, e^9	$\mathbf{s}^3$, n^4, r^5, e^9	$\mathbf{e}^9$	$\mathbf{e}^9$
DAT	$\mathbf{m}^1$, s^2, n^4, r^5, n^7, r^8, e^9	$\mathbf{m}^1$, s^2, s^3, n^4, r^5, r^6, n^7, r^8, e^9	$\mathbf{r}^6$, n^7, r^8, e^9	$\mathbf{n}^7$, r^8, e^9
GEN	$\mathbf{s}^2$, r^5, r^8, e^9	$\mathbf{s}^2$, s^3, r^5, r^6, r^8, e^9	$\mathbf{r}^6$, r^8, e^9	$\mathbf{r}^8$, e^9

Observation:

The paradigm of pronominal inflection in German obeys iconicity: Increasing specificity of a marker correlates with increasing phonological weight.

Three blocks in (8):

(i) There are super-heavy markers in the first block: /m/, /s/.

(ii) There are heavy markers in the second block: /n/, /r/.

(iii) There are light markers in the third block: /e/.

Result:

(i) Light markers show up with extremely non-specific feature specifications.

(ii) Heavy markers show up with more specific feature specifications.

(iii) Super-heavy markers show up with highly specific (but still underspecified) feature specifications.

Further observation:
Both the first and the second block contain a nasal and a non-nasal marker each. Choice of nasal or non-nasal marker can also be explained in a general and simple way: The feature [+gov] co-occurs with nasals (/m/, /n/), absence of the feature correlates with non-nasals (/r/, /s/).

3.2. Noun Inflection in Russian

P_3: *Standard paradigm(s) of noun inflection in Russian, singular*

	I_m	$II_{f,m}$	III_f	IV_n
NOM	Ø	a	Ø	o
ACC	Ø/a	u	Ø	o
DAT	u	e	i	u
GEN	a	i	i	a
INST	om	oj	ju	om
LOC	e	e	i	e

Observation:
The system of noun inflection in Russian exhibits many instances of syncretism, across cases as well as across inflection classes. Both instances of syncretism can be accounted for by assuming decomposed case and inflection class features that underspecified marker entries can refer to (Müller (2004)).

(10) *Case features*: [±subject], [±governed], [±oblique]

NOMINATIVE:	[+subj,–gov,–obl]
ACCUSATIVE:	[–subj,+gov,–obl]
DATIVE:	[–subj,+gov,+obl]
GENITIVE:	[+subj,+gov,+obl]
INSTRUMENTAL:	[+subj,–gov,+obl]
LOCATIVE:	[–subj,–gov,+obl]

(11) *Inflection class features*: [±α], [±β]

I:	[+α,–β]	$zavod_m$ ('factory')
II:	[–α,+β]	$komnat_f$ ('room'), $muž\check{c}in_m$ ('man')
III:	[–α,–β]	$tetrad'_f$ ('notebook')
IV:	[+α,+β]	$mest_n$ ('place')

(12) *Underspecified paradigm of singular noun inflection markers in Russian*:
1. /oj/ ↔ [–α,+β,+subj,–gov,+obl]
2. /ju/ ↔ [–α,–β,+subj,–gov,+obl]
3. /om/ ↔ [+α,+subj,–gov,+obl]
4. /e/ ↔ [–α,+β,–subj,+obl]

5. /e/ ↔ [+α,–subj,–gov,+obl]
6. /o/ ↔ [+α,+β,–obl]
7. /Ø/ ↔ [–β,–obl]
8. /i/ ↔ [–α,+obl]
9. /u/ ↔ [–subj,+gov]
10. /a/ ↔ [ ]

Some assumptions about further operations affecting vocabulary items:
(i) There is a morphophonological rule that realizes /om/ as *em* after a soft ([–back]) consonant, and as *om* otherwise.
(ii) There is a morphophonological rule that realizes underlying /oj/ as *ej* after a [–back] consonant, and as *oj* otherwise.
(iii) There is a morphophonological rule that realizes underlying /i/ as *y* after a hard ([+back]) consonant, and as *i* otherwise.

Note:
The (partial) feature hierarchy relevant for (9-b) is [class] > [case].

P_4: *Interaction of noun inflection markers in the singular in Russian*

	I: [+α,–β]	II: [–α,+β]	III: [–α,–β]	IV: [+α,+β]
NOM: [+subj,–gov,–obl]	**/Ø/**$_7$ (/a/$_{10}$)	**/a/**$_{10}$	**/Ø/**$_7$ (/a/$_{10}$)	**/o/**$_6$ (/a/$_{10}$)
ACC: [–subj,+gov,–obl]	**/Ø/**$_7$ (/u/$_9$, /a/$_{10}$)	**/u/**$_9$ (/a/$_{10}$)	**/Ø/**$_7$ (/u/$_9$, /a/$_{10}$)	**/o/**$_6$ (/u/$_9$, /a/$_{10}$)
DAT: [–subj,+gov,+obl]	**/u/**$_9$ (/a/$_{10}$)	**/e/**$_4$ (/i/$_8$, /u/$_9$, /a/$_{10}$)	**/i/**$_8$ (/u/$_9$, /a/$_{10}$)	**/u/**$_9$ (/a/$_{10}$)
GEN: [+subj,+gov,+obl]	**/a/**$_{10}$	**/i/**$_8$ (/a/$_{10}$)	**/i/**$_8$ (/a/$_{10}$)	**/a/**$_{10}$
INST: [+subj,–gov,+obl]	**/om/**$_3$ (/a/$_{10}$)	**/oj/**$_1$ (/i/$_8$, /a/$_{10}$)	**/ju/**$_2$ (/i/$_8$, /a/$_{10}$)	**/om/**$_3$ (/a/$_{10}$)
LOC: [–subj,–gov,+obl]	**/e/**$_5$ (/a/$_{10}$)	**/e/**$_4$ (/i/$_8$, /a/$_{10}$)	**/i/**$_8$ (/a/$_{10}$)	**/e/**$_5$ (/a/$_{10}$)

Observation:
Given an abstract analysis based on underspecified markers that is independently motivated (syncretism), the system of Russian noun inflection in the singular reveals a correlation of form and function, i.e., iconicity:
(i) From top to bottom, specificity decreases in (12).
(ii) From top to bottom, sonority increases in (12).

Generalization:
The more specific an inflection marker, the lower its rank on the sonority hierarchy (Hankamer & Aissen (1974)).

(13) *Sonority hierarchy status and specificity of markers*:
    /oj/, /ju/, /om/ > /e/ > /o/, /Ø/, /i/ > /u/ > /a/

Potential problem 1:
Why should /e/ count as less sonorous than the other vocalic markes?
Solution:
/e/ is in fact usually (except after consonants like /c/, /š/, /ž/) realized as *je*, which makes this marker quasi-consonantal. This may also partly explain why the syncretism with /e/ is not fully resolved in the present approach.

Potential problem 2:
Why does the null marker /Ø/ occupy an intermediate position with respect to specificity, given that the sonority hierarchy would seem to support an edge position for this marker?
Solution:
Halle (1994) argues that the "null marker" /Ø/ is in fact an abstract yer vowel /O/, which is independently motivated in the morphophonology of Russian. Halle argues that an abstract /O/ vowel has otherwise the same features as /o/; there is a general rule that deletes abstract vowels unless they immediately precede a syllable with another abstract vowel (which, of course, they never do if they are inflection markers at the end of a word). Thus, /O/ can be assumed to replace the the null marker /Ø/ assumed so far. And, of course, given that /O/ has a similar form as /o/, and given that optimal grammar design maximizes form/function correspondence, we expect it to have a similar function (degree of specificity of insertion context) as well.

Potential problem 3:
The sonority-based order /u/ > /o/ predicted by (12) seems incompatible with the sonority-based order /o/ > /u/ argued for in Matthews (1974), Ross (1980), Kenstowicz (1994), Müller (1997), and Crosswhite (2000).
Solution:
This might indicate some language-particular variability, or minor imperfections of some of the systems involved here.

Note:
The assumption that the sonority hierarchy plays a role in the system of Russian noun inflection is not new. Shapiro (1969) and Plank (1979) correlate a Jakobsonian hierarchy of cases H_c (nom > inst > gen_2 > loc_2 > acc > dat > gen_1 > loc_1) and a sonority hierarchy H_s (a > o, e, u, i > v, j, m > x); Plank states the following generalization: "The higher-ranked a case is in [H_c], the more sonorous is the set of phonological segments used for its expression." However, such a surface-oriented correlation can at best be assumed to be true as a weak tendency; it is only under an abstract analysis in terms of underspecified markers that the correlation becomes (near-) perfect, and Plank's and Shapiro's basic insight can be properly expressed.

3.3. The Status of Iconicity

Conclusion:
Iconicity restricts the shape of fusional inflectional systems (among other things).

Question:
What kind of principle is Iconicity?

(i) Iconicity as a functional principle:
If the child acquiring an inflectional system knows that the system obeys Iconicity, this drastically restricts the hypothesis space arising under the task of assigning inflection markers to (typically underspecified) specifications of morpho-syntactic features. Thus, Iconicity significantly simplifies the learning of inflectional systems. Hence, a grammar with this principle is computationally more efficient than a grammar without it (economy considerations; cf. Haspelmath (2003)).

(ii) Iconicity as a language-independent principle:
The principle does not refer to language-specific notions.

(iii) Iconicity as an abstract principle:
Under a surface-oriented analysis that refers to fully specified rather than underspecified feature specifications of inflection markers, Iconicity cannot be satisfied by the system of pronominal inflection in German, or by the system of noun inflection in Russian.

(iv) Iconicity as an innate principle:
Of course, Iconicity could be extracted by a child from an inflectional paradigm that it has learned. However, by assumption, the task of Iconicity is to narrow down the number of possible analyses that must be considered by a child acquiring an inflectional system. So, if the principle is to guide acquisition of an inflectional system, it seems that it must be in place before acquisition takes place. This may suggest that the principle is innate. (Iconicity contributes to computational efficiency and thus conforms to economy, but it is not clear whether it can be made to *follow* from economy considerations entertained by the language learner – there may be many other ways to ensure computational efficiency.)

4. Minimality in Syntax

(14) *Minimality* (Chomsky (1993, 1995, 2001); also see Rizzi (1990), Fanselow (1991)):
    α can only form a dependency of type δ with β if (a) and (b) hold:
    a. β has property δ.
    b. There is no γ such that (i) and (ii) hold:
        (i) γ is closer to α than β.
        (ii) γ has property δ.

Note:
For movement dependencies, closeness can be understood in terms of minimality of path length.

(15) *Path* (Pesetsky (1982), Collins (1994), Müller (1998)):
The path from X to Y is the set of categories Z such that (a) and (b) hold:
a. Z is reflexively dominated by the minimal XP that dominates both X and Y.
b. Z dominates X or Y.
The length of a path is determined by its cardinality.

Observation:
Minimality has two main consequences:
(i) γ blocks movement of β to α because γ c-commands β.
(ii) γ blocks movement of β to α because γ dominates β.

(16) *Minimality configurations*:
a. *... β_1 α ... [... γ ... [... t_1 ...] ...] ...
b. *... β_1 α ... [$_\gamma$ ... t_1 ...] ...

4.1. C-Command Effects

Observation:
Many cases for which Minimality is relevant involve c-command.

(17) *Wh-movement in English (Superiority)*:
a. (I wonder) who$_1$ t_1 bought what$_2$
b. *(I wonder) what$_2$ who$_1$ bought t_1
c. Whom$_1$ did John persuade t_1 [to visit whom$_2$] ?
d. *Whom$_2$ did John persuade whom$_1$ [t'_2 to visit t_2] ?

(18) *Non-pronominal object shift in Icelandic*(Collins & Thráinsson (1996)):
a. *Ég lána bækurnar$_2$ ekki Maríu$_1$ t_2
   I lend the books not Maria
b. Ég lána Maríu$_1$ ekki t_1 bækurnar$_2$
   I lend Maria not the books

(19) *Pronominal object shift in Danish* (Vikner (1994)):
a. *Peter viste den$_2$ jo Marie$_1$ t_2
   Peter showed it indeed Marie
b. Peter viste hende$_1$ jo t_1 bogen$_2$
   Peter showed her indeed the book

4.2. Domination Effects

Observation:
Some cases for which Minimality is relevant involve domination (Koizumi (1995), Kitahara (1997), Müller (1998), Fitzpatrick (2002)).

(20) *Infinitive scrambling in German* (Fanselow (1991), Haider (1993)):
a. *(Ich denke) dass [t_1 zu lesen]$_3$ [das Buch]$_1$ keiner t_3 versucht hat
   I think that to read the book$_{acc}$ no-one tried has

b.(?)(Ich denke) dass [das Buch zu lesen]$_3$ keiner t_3 versucht hat
   I think that the book$_{acc}$ to read no-one tried has
c. (Ich denke) dass [das Buch]$_1$ keiner [t_1 zu lesen]$_3$ versucht hat
   I think that the book$_{acc}$ no-one to read tried has

(21) *Wh-movement across a wh-island in English* (Barss (1986)):
a.??[$_{NP}$ Which book about Chomsky]$_2$ don't you know [$_{CP}$ whether to read t_2] ?
b. *[$_{NP}$ Which book about t_1]$_2$ don't you know [$_{CP}$ who$_1$ to read t_2] ?

(22) *Scrambling in Japanese* (Kitahara (1997)):
a. [Mary-ga [sono hon-o]$_1$ yonda-to]$_2$ Bill-ga [John-ga t_2 itta-to]
   Mary$_{nom}$ that book$_{acc}$ read-COMP Bill$_{nom}$ John$_{nom}$ said-COMP
   omotteiru (koto)
   think fact
b. *[Mary-ga t_1 yonda-to]$_2$ [sono hon-o]$_1$ John-ga t_2 itta (koto)
   Mary$_{nom}$ read-COMP that book$_{acc}$ John$_{nom}$ said fact

4.3. The Status of Minimality

Conclusion:
Minimality provides a unified account of restrictions on movement that emerge in contexts where there is an initial ambiguity in rule application.

Question:
What kind of principle is Minimality?

(i) Minimality as a functional principle:
Grammars that obey Minimality are computationally efficient, for at least two reasons. From the point of view of economy considerations, shorter dependencies are preferred over longer dependencies. Furthermore, given Minimality, indeterminacies in the application of operations that would otherwise arise are avoided.

(ii) Minimality as a language-independent principle:
The principle does not refer to language-specific notions.

(iii) Minimality as an abstract principle:
Minimality needs abstract, hierarchical syntactic structure in order to be able to account for data like those in (17)–(22). The principle restricts operations in the derivation; on the surface, it cannot discriminate between legitimate and illegitimate movement dependencies.

(iv) Minimality as an innate principle:
It is difficult to see how Minimality can be learned (on the basis of linguistic evidence); the problem is related to the argument for innate principles based on the Poverty of the

Stimulus.

4.4. Minimality and Poverty of the Stimulus

Background:
Chomsky assumes that language acquisition would be impossible if there were no innate principles. ('Plato's problem: How can we know so much, given that we have so little evidence?')

(23) *Argument from Poverty of the Stimulus* (Chomsky (1980, 33-36)):
    a. "My own suspicion is that a central part of what we call 'learning' is actually better understood as the growth of cognitive structures along an internally directed course under the triggering and partially shaping effect of the environment."
    b. "If it were proposed that we 'make' our physical constitution, or are 'taught' to pass through puberty, or 'learn' to have arms rather than wings, no one would take the suggestion very seriously, even in the present state of ignorance concerning the mechanisms involved. Why is this so? Presumably, the reason derives from the vast qualitative difference between the impoverished and unstructured environment, on the one hand, and the highly specific and intricate structures that uniformly develop, on the other. In essence, this is a variant of a classical argument in the theory of knowledge, what we might call 'the argument from the *poverty of the stimulus*'."
    c. "Note that the argument is of course nondemonstrative. It is what is sometimes called an inference to the best explanation, in this case, that what the stimulus lacks is produced by the organism from its inner resources."

(24) *Argument from Poverty of the Stimulus* (from Pullum & Scholz (2002, 18)):
    a. Human infants learn their first languages either by *data-driven learning* or by *innately-primed learning*.
    b. If human infants acquire their first languages via data-driven learning, then they can never learn anything for which they lack crucial evidence.
    c. *But infants do in fact learn things for which they lack crucial evidence*.
    d. Thus human infants do not learn their first languages by means of data-driven learning.
    e. Conclusion: Human infants learn their first languages by means of innately-primed learning.

An argument for innately-primed learning based on Poverty of the Stimulus (Chomsky (1971)):

(25) *Auxiliary-initial clauses in English, 1*:
    a. The dog in the corner is hungry.
    b. Is the dog in the corner hungry?

(26) *Auxiliary-initial clauses in English, 2*:
    a. The dog that is in the corner is hungry.
    b. Is the dog that is in the corner hungry?
    c. *Is the dog that in the corner is hungry?

(27) *Principles that might restrict auxiliary fronting*:
    a. Front the first auxiliary.
    b. Front the auxiliary in the matrix Infl (the structurally highest auxiliary).

Argument:
(i) The child does not find sentences of the type in (26-b) in the input.
(ii) Still, (27-b) is the 'acquirendum', (27-a) is not the rule underlying auxiliary fronting in English (we know this because (26-c) is ungrammatical).
(iii) Therefore, innate knowledge helps the child to find out that (27-b) is correct ((27-b) may or may not be innate as such).

Counter-argument in Pullum & Scholz (2002) (adopted in Tomasello (2003, 288-289)):
Assumption (i) is untenable: The child has sufficient access to data of this type. In fact, simple four-word sentences exhibit the relevant construction:

(28)   a. Is what's left mine?
    b. *Is what left's mine?

Argument against Pullum & Scholz (2002) in Legate & Yang (2002):
(i) Most of the statistical evidence given by Pullum and Scholz comes from sources that do not correctly mirror the child's natural learning environment: Wall Street Journal, the works of Oscar Wilde, etc.
(ii) Only the NINA corpus (CHILDES) cited by Pullum and Scholz provides realistic acquisition data (natural mother-child interaction). However, here they only report counts from one file, "which happens to be the file that has the most number of critical sentences, out of all 56 files."
(iii) The number of input sentences needed to determine that the (possibly non-first) highest auxiliary is moved is too small anyway. Here is why:

(29)   a(*)Like him. ('I like him.')
    b. There are two men in the room.

Generalization:
Children have acquired the right auxiliary-inversion rule by 3;2 years. Children acquiring English also stop producing pro-drop sentences (like (29-a)) at roughly the same developmental age. The triggering data for the latter are 'there' constructions, which, under at least some theories, are incompatible with the pro-drop property.

Strategy:
Count the 'there' sentences in CHILDES files, and compare them with the number of non-first auxiliary inversion sentences. Result: There are much fewer instances of the latter. This shows that these input data do not suffice to rule out the wrong auxiliary

fronting hypothesis.

Reaction in Scholz & Pullum (2002):
It is highly unlikely that 'there' sentences are incompatible with pro-drop. E.g., Modern Hebrew is a pro-drop language with lexical expletives. And: Pro-drop does to some extent exist in English:

(30) a. Had a talk with the boss this morning.
      b. Got milk?

The crucial argument against Pullum & Scholz (2002) in Lasnik & Uriagereka (2002) (based on Freidin (1991)):
"Chomsky may have been too generous with the empiricist alternative he was arguing against: Not even the fact that [(26-b)] is grammatical proves that something with the effect of hypothesis [(27-b)] is correct (and the only possibility), hence does not lead to adult knowledge of English. More generally, no amount of positive evidence, 'exotic' or not, would suffice. The poverty of the stimulus is thus extreme." (p. 148).

Assumption:
Language acquisition depends on *positive evidence* (well-formed data); pieces of *negative evidence* (ungrammatical data, e.g., in the form of corrections) are ignored by the child and play no role in acquisition.

Consequence:
If data of the type in (26-b) are systematically available to the child, this implies that (27-a) is not maintained by the child; but it does not imply that (27-b) is acquired by the child. Alternatively, the child could acquire the (wrong) principles in (31), which are compatible with (26-b) (but not with the negative evidence (26-c)).

(31) *Further principles that might restrict auxiliary fronting*:
      a. Front any auxiliary.
      b. Front any finite auxiliary.

Reaction in Scholz & Pullum (2002):
(i) These issues are orthogonal to the main claim in Pullum & Scholz (2002), which concerns Chomsky's original argument only.
(ii) Apart from that, negative evidence might play a role after all in acquisition.

Tomasello (2003, 177):
"Adults [...] do respond differently to well-formed and ill-formed child utterances. [...] But most theorists do not consider this kind of indirect feedback sufficient to constrain children's overgeneralization tendencies, as it is far from consistent. It is also not clear that this type of feedback is available to all children learning all languages. Nevertheless, it is still possible that linguistic feedback from adults may play some role – although neither a necessary nor a sufficient role – in constraining children's overgeneralization tendencies."

Question:
Can Minimality be learned, given Poverty of the Stimulus?

Claim:
Assuming that the child has acces to the well-formed data in (17)-(22) (positive evidence), but not to the ill-formed data in (17)-(22) (lack of negative evidence), (32-a) can be learned whereas (32-b) cannot be learned. However, (32-c) is also fully compatible with the empirical evidence. Hence, given that negative evidence is not available, there is no reason why the child should not maintain a superset grammar based on (32-c) and fail to switch to the target grammar based on (32-a): The first grammar generates a superset of the sentences generated by the second grammar. Thus, there is again an extreme case of Poverty of the Stimulus, exactly like the one noted by Lasnik and Uriagereka.

(32) *Three hypotheses*:
      a. Move the matching category that is closest.   (= (14))
      b. Move the matching catgory that is most remote.
      c. Move a matching category.

5. Conclusion

Main conclusion:
It seems that the primary dividing line is not form vs. function, language-specific vs. language-independent principles, or innateness vs. learning by experience; it boils down to abstract vs. surface-oriented analysis of grammatical phenomena:
"Usage-based theorists do not search for the most general abstraction possible" (Tomasello (2003, 98)).
"The basic computational ingredients are considerably more abstract" [than, e.g., Ross's island constraints, or the Specified Subject Condition] (Chomsky (2004, 7)).

References

Barss, Andrew (1986): Chains and Anaphoric Dependence. Ph.d. thesis, MIT, Cambridge, Mass.

Bierwisch, Manfred (1967): Syntactic Features in Morphology: General Problems of So-Called Pronominal Inflection in German. In: *To Honour Roman Jakobson*. Mouton, The Hague/Paris, pp. 239–270.

Chomsky, Noam (1971): *Problems of Knowledge and Freedom*. Fontana, London.

Chomsky, Noam (1980): *Rules and Representations*. Blackwell, Oxford.

Chomsky, Noam (1981): *Lectures on Government and Binding*. Foris, Dordrecht.

Chomsky, Noam (1993): A Minimalist Program for Syntactic Theory. In: K. Hale & S. J. Keyser, eds., *The View from Building 20*. MIT Press, Cambridge, Mass., pp. 1–52.

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

Chomsky, Noam (2000): Minimalist Inquiries: The Framework. In: R. Martin, D. Michaels & J. Uriagereka, eds., *Step by Step*. MIT Press, Cambridge, Mass., pp. 89–155.

Chomsky, Noam (2001): Derivation by Phase. In: M. Kenstowicz, ed., *Ken Hale. A Life in Language*. MIT Press, Cambridge, Mass., pp. 1–52.

Chomsky, Noam (2004): Three Factors in Language Design. Ms., MIT, Cambridge, Mass.

Collins, Chris (1994): Economy of Derivation and the Generalized Proper Binding Condition, *Linguistic Inquiry* 25, 45–61.

Collins, Chris & Höskuldur Thráinsson (1996): VP-internal Structure and Object Shift in Icelandic, *Linguistic Inquiry* 27, 391–447.

Crosswhite, Katherine (2000): Sonority-Driven Reduction. Ms., University of Rochester. To appear in *Proceedings of BLS 26*.

Fanselow, Gisbert (1991): Minimale Syntax. Habilitation thesis, Universität Passau.

Fitzpatrick, Justin (2002): On Minimalist Approaches to the Locality of Movement, *Linguistic Inquiry* 33, 443–463.

Freidin, Robert (1991): Linguistic Theory and Language Acquisition, *Behavioral and Brain Sciences* 14, 618–619.

Haider, Hubert (1993): *Deutsche Syntax – generativ*. Narr, Tübingen.

Halle, Morris (1994): The Russian Declension: An Illustration of the Theory of Distributed Morphology. In: J. Cole & C. Kisseberth, eds., *Perspectives in Phonology*. CSLI Publications, Stanford, pp. 29–60.

Hankamer, Jorge & Judith Aissen (1974): The Sonority Hierarchy. In: A. Bruck, R. Fox & M. Labaly, eds., *Papers from the Parasession on Natural Phonology*. Chicago Linguistic Society, University of Chicago, pp. 131–145.

Haspelmath, Martin (1999): Optimality and Diachronic Adaptation, *Zeitschrift für Sprachwissenschaft* 18, 180–205.

Haspelmath, Martin (2003): Against Iconicity and Markedness. Ms., MPI für evolutionäre Anthropologie.

Jakobson, Roman (1962): Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus. In: *Selected Writings*. Vol. 2, Mouton, The Hague and Paris, pp. 23–71.

Kenstowicz, Michael (1994): Sonority-Driven Stress. Ms., MIT.

Kitahara, Hisatsugu (1997): *Elementary Operations and Optimal Derivations*. MIT Press, Cambridge, Mass.

Koizumi, Masatoshi (1995): Phrase Structure in Minimalist Syntax. PhD thesis, MIT, Cambridge, Mass.

Lasnik, Howard & Juan Uriagereka (2002): On the Poverty of the Challenge, *The Linguistic Review* 19, 147–150.

Legate, Julie Anne & Charles Yang (2002): Empirical Re-Assessment of Stimulus Poverty Arguments, *The Linguistic Review* 19, 151–162.

Matthews, Peter (1974): *Morphology*. Cambridge University Press, Cambridge.

Müller, Gereon (1997): Beschränkungen für Binomialbildung im Deutschen, *Zeitschrift für Sprachwissenschaft* 16, 5–51.

Müller, Gereon (1998): *Incomplete Category Fronting*. Kluwer, Dordrecht.

Müller, Gereon (2004): A Distributed Morphology Approach to Syncretism in Russian Noun Inflection. In: O. Arnaudova, W. Browne, M. L. Rivero & D. Stojanovic, eds., *Proceedings of FASL 12*. University of Ottawa.

Pesetsky, David (1982): Paths and Categories. PhD thesis, MIT, Cambridge, Mass.

Plank, Frans (1979): Ikonisierung und De-Ikonisierung als Prinzipien des Sprachwandels, *Sprachwissenschaft* 4, 121–158.

Pullum, Geoffrey & Barbara Scholz (2002): Empirical Assessment of Stimulus Poverty Arguments, *The Linguistic Review* 19, 9–50.

Rizzi, Luigi (1990): *Relativized Minimality*. MIT Press, Cambridge, Mass.

Ross, John (1980): Ikonismus in der Phraseologie, *Zeitschrift für Semiotik* 2, 39–56.

Scholz, Barbara & Geoffrey Pullum (2002): Searching for Arguments to Support Linguistic Nativism, *The Linguistic Review* 19, 185–223.

Shapiro, Michael (1969): *Aspects of Russian Morphology*. Slavica Publishers, Cambridge, Mass.

Tomasello, Michael (2003): *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Harvard University Press, Cambridge, Mass.

Vikner, Sten (1994): Scandinavian Object Shift and West Germanic Scrambling. In: N. Corver & H. van Riemsdijk, eds., *Studies on Scrambling*. Mouton de Gruyter, Berlin, pp. 487–517.

Wiese, Bernd (1996): Iconicity and Syncretism. On Pronominal Inflection in Modern German. In: R. Sackmann, ed., *Theoretical Linguistics and Grammatical Description*. Benjamins, Amsterdam, pp. 323–344.

Wiese, Bernd (1999): Unterspezifizierte Paradigmen. Form und Funktion in der pronominalen Deklination, *Linguistik Online* 4. (www.linguistik-online.de/3_99).