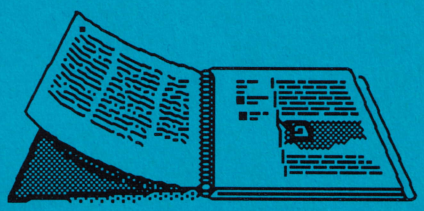
Forschungsschwerpunkt Allgemeine Sprachwissenschaft,
Sprachtypologie und Universalienforschung
der Förderungsgesellschaft Wissenschaftliche Neuvorhaben mbH

FAS Papers in Linguistics

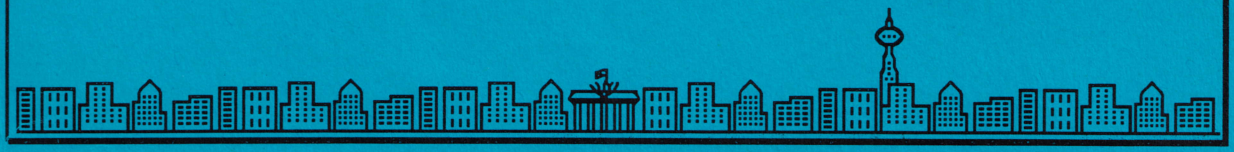
Volume 1
April 1995



Edited by

Artemis Alexiadou
Nanna Fuhrop
Paul Law
Sylvia Löhken

ISSN 0947-7055



The Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (FAS, Research Center for General Linguistics, Typology and Universals) was founded in 1992 under the auspices of the Förderungsgesellschaft Wissenschaftliche Neuvorhaben mbH (Agency for the Promotion of New Scientific Research), a subsidiary of the Max Planck Gesellschaft (Max Planck Society). The Center's 19 professionals and 7 doctoral candidates conduct original research in many subdisciplines of linguistics, including syntax, semantics, morphology, lexicon, pragmatics, phonology, phonetics, creolization, language change and language acquisition in a typologically diverse range of languages. In addition to providing a forum for the exchange of ideas in the academic community in the Berlin area through lectures, seminars, workshops and conferences, the Center has cooperation projects with other universities in Germany, and sponsors visits by scholars from Europe and America.

For further information on the Center, please write to:

Ewald Lang, Director
FAS
Jägerstr. 10/11
D-10117 Berlin
Germany

Telephone: +49 30 20 19 24 01
Fax: +49 30 20 19 24 02

E-mail: sprach@fas.ag-berlin.mpg.de

Table of Contents

Prefatory Note

On 1 January 1992, the Forschungsschwerpunkt Allgemeine Sprachwissenschaft, Typology und Universalienforschung (= FAS) [Research Centre for General Linguistics, Typology, and Universals] was set up as one of seven newly founded research centres for the humanities. In line with the recommendations of the Scientific Council of Germany, the FAS is conceived of as an independent institute of innovative basic research in linguistics having close contacts with nearby universities.

The name of the centre is its programme. The *FAS - Papers in Linguistics* is an informal publication intended to serve as a forum for presentation and rapid dissemination of current research by the staff members and by guest researchers of the FAS. Occasionally, we will also include current work by linguists from abroad who are associated with FAS projects in one way or other.

Due to irregular schedule of publication, we cannot offer subscriptions. We are interested in exchanging *FAS - Papers in Linguistics* with other linguistic and related publications, especially with the working papers of other linguistic departments. Prospective exchange partners are asked to contact the editors.

The second issue of *FAS - Papers in Linguistics* will appear in June 1995 and will contain papers on optimality theory, the third issue is planned for fall 1995 and will deal with morphology matters.

Berlin, Spring 1995

Ewald Lang
Provisional Head of Staff

The Typology of Structural Deficiency

On the Three Grammatical Classes*

January 1993 - May 1994

Revised: MIT, October 1994

Anna Cardinaletti
(Univ. of Venice)
cardin@unive.it

&

Michal Starke
(Univ. of Geneva / MaxPlanck Berlin)
starke@uni2a.unige.ch

1. On the Study of Pronouns

1.1. The Notion of "Classes of Pronouns"

It is a general property of language that words fall into classes. Among the many relevant oppositions (verbs/adjectives, transitives/ergatives, etc.), one distinguishes itself from all others: that instantiated by the opposition between different classes of pronouns.

This opposition is unique in regularly contrasting *synonymous pairs*; in cutting across all components of grammar; in having no systematic correlation with any interpretive characteristic (semantic or phonetic); in determining a large set of (apparently) absolute universals; and in cutting across lexical classes, §1.1.1-5.

The fundamental goal of the present inquiry is to uncover the primitive underlying these exceptional classes.

1.1.1. Unmarkedly, one and the same pronoun (semantically / functionally defined) falls into distinct classes. The third person plural feminine nominative Italian pronouns, for instance, divide into two distinct classes with respect to coordination and reference:

					<+human>	<-human>
(1)	a.	**Esse**	(*e quelle accanto)	sono troppo alte.	✓	✓
	b.	**Loro**	(e quelle accanto)	sono troppo alte.	✓	*
		3.pl.fm.nom	(and those besides)	are too tall/high		

One class of pronouns ("class 1") may be coordinated, but it is limited to human referents, while the other ("class 2") cannot be coordinated and may refer to both human and non-human entities. In many cases, the two classes are not only functionally but also phonetically non-distinct: the French translation of (1), for instance, reproduces exactly the same pattern without morphological variation.

(2)	a.	**Elles**		sont trop grandes.	✓	✓
	b.	**Elles**	et celles d'à côté	sont trop grandes.	✓	*

In (2), the non-human reading vanishes in coordination. The mystery of this correlation between coordination and interpretation reduces if the formal parallelism between (1) and (2) is taken into account: despite phonetic

* We thank the organisers of the Incontro di grammatica generativa (Trento, February 1993), Glow (Lund, April 1993), ESF Clitic group Meeting (Trondheim, June 1993), Potsdam Encounters (December 1993), Comparative Germanic Syntax Conference (Harvard, January 1994) and Linguistic Symposium on Romance Languages (Los Angeles, March 1994) for providing us with an occasion to present this to a receptive audience. Several lines of argumentations stem from discussions with Guglielmo Cinque, Liliane Haegeman, Morris Halle, Riny Huybregts, Anthony Kroch, Marina Nespor, David Pesetsky, Henk van Riemsdijk, Luigi Rizzi, Dominique Sportiche, Jean-Roger Vergnaud and Chris Wilder.

Finally, none of this would have been possible without previous studies, such as those of: R. Kayne, for the syntactic placement of clitic personal pronouns, CL-PL (1975, 1991), A. Holmberg, for first bringing to light clitic-like non clitics, (1986), and J. Schroten for the new emphasis on the old observation concerning the interaction between deficiency and ability to refer to non-human entities, (1986/ 1992).

Although the whole paper is a joint enterprise, for all academic purposes Anna Cardinaletti takes responsibility for Sections 1.-5.1., Michal Starke takes responsibility for Sections 5.2.-10.

identity, (2) features both classes of pronouns: as before, the class which may be coordinated can only refer to human entities.

(3)

		occurs in coordination	only human referents
class 1	*loro, elles*$_1$	+	+
class 2	*esse, elles*$_2$	–	–

One and the same semantically / functionally defined pronoun (third person plural feminine nominative unstressed) is the surface reflex of two distinct underlying grammatical elements. The existence of regular synonymous (and often homophonous) pairs, is a rare, if not unique, characteristic of the class 1 / class 2 distinction.

1.1.2. Not only is the class 1 / class 2 distinction exceptional in triggering homonymy and homophony, but it also triggers a large array of surface asymmetries, distributed across syntax, morphology, semantics and prosody. Again, it is a virtually unique characteristic in grammar that asymmetries of such different components, often considered strictly disjoint, all cluster around the same class-opposition, §2.

1.1.3. Although the class 1 / class 2 distinction is linked to several interpretive properties, both phonetic and semantic, none of these links is systematic. As seen above, there is for instance no strict covariation between class membership and human reference, only asymmetric (and overlapping) possibilities. The class 1 / class 2 distinction is purely grammatical, i.e. abstract, again an unusual state of affairs.

1.1.4. This unique abstract and pervasive distinction also seems to be an absolute universal. It is for example always true that a coordinated personal pronoun cannot refer to a non-human entity. As an example of the cross-linguistic invariance of class 1 and class 2, the following languages all have an asymmetry identical to that in (1)-(2). [1]

---

[1] Hungarian, Hebrew and Gun data courtesy of, respectively, Gabriella Toth, Ur Shlonsky and Enoch Aboh.

Some remarks, however: Hungarian speakers divide into two groups w.r.t. *öket*, those who use it as in (6) (the majority of our informants), and a second group who treats it as a pure class 1 pronoun on a par with Italian nominative *loro*, i.e. only referring to animate entities (the second group is irrelevant to this paradigm). The difference between the two groups is somewhat unclear, although the second is sometimes deemed "conservative".

Examples (5) and (6) also show the invariance of this paradigm w.r.t. the subject/object asymmetry. Somewhat more trivially, the same paradigm applies to English, where *it* patterns with Italian *esse*, and *he* patterns with Italian *loro*:

		<+human>	<- human>			<+>	<->
[i]	a. It is big	n.a.	✓	[ii]	a. He is big.	✓	n.a.
	b.*It and the other one are nice.	n.a.	*		b. He and the other one are nice.	✓	n.a.

The relevant fact being that exactly the pronoun which refers to a [- human] entity cannot be coordinated, whereas its human counterpart can. *It* is a class 2 pronoun restricted to [- human] referents, whereas *he* may act as a class 1 pronoun, though it is highly plausible that a class 2 counterpart exists. Due to the lack of morphological distinction and the absence of grammatical/semantic gender distinction a. o., English will not be discussed here in any depth. Cf. Cardinaletti & Starke (1994a) for more details.

Let us note however, that, somewhat paradoxically, English provides the only example, to our knowledge, going against the putative absolute universal that coordinated personal pronouns cannot refer to non-human entities: for a majority of speakers, with some variation both across speakers and constructions, coordinated *they*, *them* may still refer to non-human entities. This fact may however be irrelevant: the above generalisation holds only of personal pronouns. Demonstratives, for instance, may refer to non-human entities when coordinated. But the apparent exceptions involve exactly those pronouns which have an initial demonstrative morpheme, *th-*. English plural might thus be similar to Scandinavian languages, in which third person personal pronouns have demonstrative morphology, and no counterexample arises. Given the wealth of indications provided by morphology (§4-7), this path seems very plausible.

			<+human>	<- human>
(4)	German (∈ Germanic)			
a.	**Sie**	sind groß	✓	✓
b.	**Sie und die daneben**	sind groß	✓	*
	they and those besides	are tall/big		
(5)	Slovak (∈ Slavic)			
a.	Vidiel som **ich**		✓	✓
b.	Vidiel som **ich a tých druhých**		✓	*
	seen I.am them and these others			
(6)	Hungarian (∈ Finno-Ugric)			
a.	Láttam **öket**		✓	✓
b.	Láttam **öket és a mellettük levöket**		✓	*
	I.saw them and those besides			
(7)	Hebrew (∈ Semitic)			
a.	**Hi**	gvoha	✓	✓
b.	**Hi ve-zot le-yad-a**	gvohot	✓	*
	she and-that.one to-side-her	tall/big		
(8)	Gun (∈ Kwa)			
a.	**Yélè**	yon wankpè	✓	✓
b.	**Yélè kpo yélè kpo**	yon wankpè	✓	*
	she and she and	know beauty		

1.1.5. Finally, not only personal pronouns, but also quantifiers, adverbs, adjectives, etc. divide into class 1 / class 2, here Greek adverbs and French bare quantifiers (§9): [2]

(9) a. To **sigo** (*ke **kalo**) évrasa.
it slowly and well I.boiled

b. Jean a **tout** (*et encore plus) vu.
Jean has all (and even more) seen

1.1.6. The conjunction of such exceptional properties (regular synonymy, (homophony,) link between all components of grammar, no link to any interpretive characteristic, absolute universal) makes this distinction one of the most profound and mysterious properties of human grammar.

The goal of this study is to uncover the source of these asymmetries, that which makes a pronoun be a class 1 / class 2 pronoun:

- What is γ, the underlying (universal) trigger of (1) which provokes a wide array of distributional, semantic, prosodic and morphological asymmetries between two forms of one and the same pronoun?

1.2. Methodology: On Generalising and Idealising

1.2.1. In doing systematic research directed towards the formulation of an abstract model, facts (or asymmetries) are not interesting in and by themselves. What is to be explained by the model are (genuine) generalisations. In such research, it is usual that some facts resist generalisation, and some generalisation resist integration into the model. In these cases, idealisation is necessary: resisting facts are consciously evacuated, in hope of subsequent reintegration.

---

[2] There is no intrinsic impossibility in (9b), the class 1 version of the quantifier is perfectly acceptable in the same sentence: *Jean a vu tout et encore plus* "John has seen all and even more". The same holds of the Greek adverbs, where the counterpart is *To évrasa sigá ke kalà*.

Although these two guidelines are contradictory (generalisation dictates integration, idealisation provokes elimination), no contradiction results: idealisation is valid only as a 'last resort', when generalisation cannot be reasonably pursued further. [3]

1.2.2. In studies of pronouns this basic point is rarely respected: many a model seeks to derive a generalisation which eliminates an unnecessarily vast amount of facts. For this reason, a large part of what follows is devoted to a preliminary step: establishing what there is to be explained, i.e. what the surface reflexes of γ are (§2-3).

1.2.3. Extending a generalisation can mean one of three things. Generalisations being of the form "all elements of the set λ fall into N non-overlapping sets $\mu_1 .. \mu_N$ with respect to the set of properties π", either the *basis* of the generalisation, λ, the *classes*, μ, or the *contrasts* of the generalisation, π, can be extended.

The simple generalisation (3) can be extended in all three directions:

(a) w.r.t. π : the contrasting properties are not limited to coordination and human referents (§2)
(b) w.r.t. μ : the πs divide the λs into three, not two, classes (§3)
(c) w.r.t. λ : the elements submitted to the generalisation are not limited to personal pronouns (§9).

[3] This is a somewhat simplified version of facts. In practice, the 'last resort' nature of idealisation is blurred by an additional factor: tolerance to uncertainty. Since it is rarely clear whether a generalisation is valid or spurious, a limit to reasonable doubt/uncertainty has to be fixed. This limit is ideally relatively low, so as not to work with potentially spurious generalisations. On the surface, this may give the impression that idealisation takes precedence over generalisation. A more correct statement is that idealisation is a last resort when a generalisation cannot be extended, where 'cannot' is understood as incorporating the accepted limit to uncertainty.

PART I. WHAT IS THERE TO BE ACCOUNTED FOR?

2. ON BEING DEFICIENT

2.1. Morphology

When (2) is transposed to a masculine subject two morphologically distinct, though related, pronouns appear. The same obtains with objects, here illustrated for Italian and Slovak:

						< +human >	< - human >
(10)	a.		**Il**		est beau	✓	✓
	b.	*	**Il** et celui de Jean		sont beaux	*	*
	c.		**Lui**		est beau	✓	*
	d.		**Lui** et celui de Jean he and the.one of John		sont beaux is/ are pretty	✓	*
(11)	a.		Non metterò	mai **loro**	il cappuccio	✓	✓
	b.	*	Non metterò	mai **loro** e loro	il cappuccio	*	*
	c.		Non metterò	mai	il cappuccio **a loro**	✓	*
	d.		Non metterò not I.will.put	mai never	il cappuccio **a loro** e a quelle altre the cap/pen-top (to) them (and to those others)	✓	*
(12)	a.		Vidím **ho**			✓	✓
	b.	*	Vidím **ho** a tých druhých			*	*
	c.		Vidím **jeho**			✓	*
	d.		Vidím **jeho** a tých druhých I.see it/ him (and these others)			✓	*

Minimally, the fact that the morphological differences exactly correlate with coordination possibilities and with possibilities w.r.t. human reference, confirms the correctness of the class 1 / class 2 distinction. But morphology not only confirms the existence of an abstract γ, it also reveals another property associated to it: the morphological difference is asymmetric. If transparently distinct, class 2 personal pronouns are systematically reduced with respect to class 1 personal pronouns: [4]

(13) *Morphological asymmetry*
morphology (class 2) ≤ morphology (class 1)

Terminology The abstractness of the two classes is no impediment to more intuitive terminology. Drawing on the clear orientation of the morphological asymmetry, class 2 elements will be called "deficient", and class 1 elements "strong".

2.2. Distribution

When the initial paradigm, (2), is embedded under *trouver* 'find', strong and deficient personal pronouns surface in different positions: [5]

---

[4] The proviso to *transparent* distinctness is necessary due to the existence of the third case of the three possible morphological relations: (a) the two lexemes are identical <elles ; elles>, (b) the two lexemes are different, one is a proper subset of the other, transparent morphology, <jeho ; ho>, <a loro ; loro>); and (c) the two lexemes are different, no (proper) subset relations obtains, opaque morphology, <lui ; il>.

If opaque class 1/ class 2 relationships are due to the class 1 element being a porte-manteau morpheme for the distinct morphemes of a transparent class 1 pronoun, then the text generalisation is correct underlyingly but will be statistical at the surface: some surface counterexamples should exist due to the surface impredictibility of portemanteau morphemes.

[5] The c-example is not acceptable as it is. It becomes natural if *elle* is understood as contrastive, cf. §2.3.

					< +human >	< - human >
(14) a.		Jean **les**	trouve	belles	✓	✓
b.	*	Jean **les** et celles d'à côté	trouve	belles	*	*
c.		Jean	trouve **elles**	belles	✓	*
d.		Jean	trouve **elles** et celles d'à côté	belles	✓	*
		John them.fem (and those besides)	finds them.fem (and those besides)	pretty		

Again, this asymmetry strictly correlates with those discussed above (coordination, human referents, morphology) and such a perfect correspondance of four properties legitimates the postulation of two abstract classes.

But again, not only is there a *difference* between the two classes, but there is an *asymmetric* difference: one class has an impoverished distribution w.r.t. the other. While strong pronouns have the distributional liberty of a corresponding noun-phrase (a full noun-phrase must occur in post-verbal position in (14)), there are three types of positions a deficient pronoun cannot occupy (cf. Kayne (1975) for an early systematization of the distributional properties of pairs such as *les / elles* in French).

2.2.1. θ-Positions.

Differently from strong personal pronouns and noun-phrases, deficient pronouns cannot occur in what might be taken to be the base, or θ-position. The following examples illustrate the base position of subjects, indirect objects and direct objects, respectively, in Italian: [6] [7]

(15) a.		{essa$_D$; lei$_S$; Maria}	forse l'ha fatto	{*essa$_D$; lei$_S$; Maria}	da sola
		{it$_D$; she$_S$; Mary}	maybe it-has done		DA alone
b.	Non dirò mai	{loro$_D$; *a loro$_S$; *a Gianni}	tutto	{*loro$_D$; a loro$_S$; a Gianni}	
	not I.will.say never	{them$_D$; to them$_S$; to Gianni}	everything		
c.	Gianni	{li$_D$; *loro$_S$; *questi studenti}	stima	{*li$_D$; loro$_S$; questi studenti}	
	Gianni	{them$_D$; them$_S$; these students}	estimates		

2.2.2. Peripheral Positions.

Differently from strong personal pronouns and noun-phrases, deficient pronouns cannot occur in a series of peripheral positions (counting isolation as peripheral, maybe as a subcase of dislocation). Literally, the same constraint holds of any other deficient pronoun, be it Dutch *het* "it", Slovak *mi* "to me" or English *it*: [8]

(16) a.	E'	{*essa$_D$; lei$_S$;Maria }	che è bella.		(*cleft*)
	It is	{*3.sg.fm$_D$; 3.sg.fm$_S$; Mary }	that is pretty		
b.		{*essa$_D$; lei$_S$;Maria },	lei è bella.		(*left dislocation*)
		{*3.sg.fm$_D$; 3.sg.fm$_S$;Mary },	she/it is pretty		
c.			*pro* arriverà presto,	{*essa$_D$; lei$_S$;Maria }	(*right dislocation*)
			She/it will arrive soon,	{*3.sg.fm$_D$; 3.sg.fm$_S$;Mary}	
d.			Chi è bella?	{*essa$_D$; lei$_S$; Maria}	(*isolation*)
			Who is pretty?	{*3.sg.fm$_D$; 3.sg.fm$_S$;Mary }	

2.2.3. C-Modification / Coordination

Noun-phrase internal modifiers cannot modify strong personal pronouns, (17a). Adverbs that modify the whole noun-phrase (c-modifiers) may however do so, (17b,c). But even c-modifiers cannot modify deficient pronouns, (17b',c').

(17) a.	*	{beau; rapide; ... }	lui	a'.	*	{beau; rapide; ... }	il
b.	✓	{vraiment; seulement; ... }	lui	b'.	*	{vraiment; seulement; ... }	il
c.	✓ lui	{seul; aussi; ...}		c'.	* il	{seul; aussi; ...}	

The ban on c-modification and coordination holds even if the complex occupies an otherwise licit position:

---

[6] D- and S- indices correspond to 'deficient' and 'strong'. The restrictions on the placement on *essa* (or equivalently *egli* "he", *essi* "they") and dative *loro* "to.them" are particularly interesting due to absence of any adjacency effect with the verb, contrary to other Italian deficient pronouns.

[7] There are so many interesting interactions between "being deficient" and "being complement of a preposition", that we reserve this topic for a different article. No mention of the interaction between pronouns and prepositional phrases will be made here (modulo "dummy prepositions, §5).

[8] As expected, in all these constructions, the French *elle* "she" may only refer to human entities.

(18) a. Anche/Solo {*essa$_D$; lei$_S$; Maria } è bella *(c-modification)*
b. Lei e(d) {*essa$_D$; lei$_S$; Maria } sono belle *(coordination)*
She and / Also / Only {*3.sg.fm$_D$; 3.sg.fm$_S$; Mary } is/ are pretty

2.2.4. Overview

(19) *Syntactic Asymmetry.*
A deficient, but not a strong, personal pronoun cannot occur at surface structure in:
a. θ–/ base positions
b. peripheral positions
c. {c-modification, coordination}

As a generalisation on distributional asymmetries between the deficient and strong pronouns, (19) is redundant. The first two clauses are reformulable as special cases of a more general positive constraint which forces deficient pronouns to occur in a given (functional) projection:

(20) *Syntactic Asymmetry*
A deficient, but not a strong, personal pronoun:
a. must occur in a special derived position
b. is incompatible with {c-modification, coordination}

2.3. Choice

As noted in fn. 5, (14c) is strongly idealised. The relevant paradigm is ("☞" denotes ostension): [9]

(21) a.	Je	✓ **l'**	ai aidé	* **elle**
b.	J(e)	* **LA**	ai aidé	✓ **ELLE**
c.	J(e)	* ☞**la**	ai aidé	✓ ☞**elle**
d.	J(e)	* **la** et l'autre	ai aidé	✓ **elle** et l'autre
e.	J(e)	* seulement **la**	ai aidé	✓ seulement **elle**
	I		have helped	her / her and the other / only her

That the post-participial variant of (21a) is impossible is *a priori* unexpected since the postverbal position is adequate for a strong pronoun. The comparison with (21b-e) brings a clear generalisation: the strong form is impossible where the deficient form is possible, and the strong form is possible where the deficient form is independently excluded: by contrastive stress (§2.4.1.), by an accompanying pointing gesture (§2.4.1.) or by coordination or c-modification (§2.2.3). Descriptively (cf. §7 for a more formal version):

(22) *Choice of a pronoun*
Choose the most deficient possible form.

2.4. Semantics: Description

2.4.1. Prominent Discourse Referents

In turn, (21b-c) is somewhat idealised. It is not the case that deficient pronouns can never be contrastively focussed. (23a) for instance, severly contrasts with (23b-c):

(23) a. * Jean **LA** voit.
✓ Jean voit **ELLE**.
John sees her

b. A: On a dit que **je** mangerai ce gateau demain.
A: we have said that I will.eat this cake tomorrow

✓ B: Non, que **JE** mangerai ce gateau demain.
B: no, that I will.eat this cake tomorrow

---

[9] (21b) is more marked than the corresponding Italian (15c). Such variation is independent of the theory of pronouns: the same preferences obtain with contrasted full DPs. Cf. also fn. 35.

✓A: Mais, non, que **JE** mangerai ... (etc.)
A: but, no, that I will.eat

c. A: Je **te** casserai la gueule!
A: I you will.break the face

✓ B: Ah ouais? tu veux dire que je **TE** casserai la gueule! (ad lib.)
B: oh yeah? you want to.say that I YOU will.break the face!

This state of affairs is not particular to prosody: the same holds with ostension, under "flat" intonation. In a limited range of contexts, a deficient pronoun may accompany ostension: [10]

(24) a. * J'ai vu Marie puis je ☞**l'** ai vu.
✓ J'ai vu Marie puis j' ai vu ☞**elle**.
I have seen Mary then I her have seen her

b. ✓ Mets-toi içi et regardes cette maison. Tu ☞**la** vois bien maintenant?
come here and look at this house. You it see well now?

c. ✓ Mais, tu ne vois donc pas ce livre? Bien sûr que je ☞**le** vois
But, you don't see this book? Of course that I it see

In both cases the generalisation is the same: *the deficient elements are permissible with {contrastive stress; ostension} only if they refer to an entity which is "already prominent in the discourse".* [11] [12]

(25) *Semantic Asymmetry #1*
Deficient personal pronouns must have an antecedent prominent in the discourse.

---

[10] The same holds of third person pronouns w.r.t. focus:
A: Jean a dit que Pierre arrivera en premier. J. has said that P. will.arrive as first
B: Non, Jean a dit qu' IL arrivera en premier No, J. has said that HE will.arrive as first
When the contrast is realised however a deficient pronoun becomes impossible:
B: * Non, Jean a dit qu' IL (, pas son frère,) arrivera en premier (, pas son frère) .... HE (,not his brother) ...
We hypothesize that overt contrast is a case of c-modification: the contrastive phrase modifies the pronoun and thus systematically excludes deficient pronouns. The apparently discontinuous constituent, i.e. extraposed contrasted phrase, is then similar to the English contrast [i]-[ii] (with *only* modifying the pronoun):
[i] John has only seen [t him] [ii] * John has only seen [t it]

[11] "Discourse" should not be restricted to linguistic events. It is possible to introduce an entity by gesture (ostension) and then refer to it by a deficient indexical.

For the dicussion of the recoverability conditions on the antecedent, cf. Tasmowski-De Ryck & Verluyten (1982), who arrive at the same conclusion that "true pronouns [i.e. deficient pronouns, in our terminology, A.C. & M.S.] can only refer to something that is already familiar" (p. 341). It is however clear that much remains to be done to define what the conditions on "prominence" are.

[12] Several recent studies capitalise on a similar generalisation: deficient personal pronouns are "specific" (e.g. Sportiche (1992), Uriagereka (1992)). There is unfortunately a lot of terminological confusion around this term. On the one hand, proponent of this view seem to understand "specificity" as Enç (1991), i.e. a term closely related to "old information" (among others, Uriagereka (op. cit.) explicitly relates it to notions such as "information already introduced in the discourse" (p. 8), "familiarity" (p. 14), "being anaphoric on [...] in the discourse" (p. 13), the "subject's point of view" (p. 22), etc.; Diesing (1991) undestands specificity in terms of "presuppositionality"). On the other hand, "specificity" has widely been understood *literally* (i.e. x is specific iff x is unique and x is "well-defined"), maybe due to the semantics of personal pronouns *per se*, which tend to be definite, irrespectively of their strong / deficient status.

The latter (literal) understanding brings a wrong generalisation about deficient personal pronouns: it is not the case that deficient pronouns always refer to an entity which is both unique and well-defined, i.e. literally specific. Counterexamples abound, among which non-referential pronouns (i.e: *les Siciliens a peine ils te voient ils t'embrassent*, "the Sicilians, as soon as they see you, they kiss you" and other cases discussed in §2.4.3., §2.5.), and non-definite pronouns (*des touristes, à Venise, j'**en** ai vu plein* "tourists, in Venice, I have seen plenty", *intelligent? Pierre l'est sans aucun doute* "intelligent? Pierre it is without doubt", *la bière, s'y digère-t-elle mal?* "the beer, *se* there digests badly?", etc.).

On the other hand, when understood correctly (i.e. non-literally), "specificity" of deficient pronouns is identical to the text generalisation, (25).

The a-example of (23) and (24) are impossible to the extent that contrastive stress and ostension usually refer to an entity non-prominent in the discourse, while the c-examples are constructed such that the referent of the deficient pronoun is the prominent topic of discourse.

It is thus not the case that strong pronouns have any special ability to be stressed or used in ostension. Both deficient and strong pronouns are identically stressable and usable in ostension. Strong pronouns are more frequent than deficient pronouns in these constructions only because they are able to introduce refer to a non-prominent discourse referent. [13]

2.4.2. Expletives

Expletive and quasi-expletive constructions always require personal pronoun subjects to be deficient. Strong pronouns are uninterpretable in these non-referential positions. [14]

(26) a. ✓ Il est arrivé un grand malheur.
** Lui (il) est arrivé un grand malheur.
he (he) is arrived a big disaster

b. ✓ Il pleut.
** Lui (il) pleut.
he (he) rains

2.4.3. Impersonal Constructions

The same holds of impersonal interpretation both with the deficient *on* in (27) (which has no strong counterpart), and with third person plural pronouns, (28). Again, only the deficient form is possible in a non-referential context, and strong forms are uninterpretable, either as doublers of the deficient subject or by themselves: [15]

			impersonal interpr.	referential interpret.
(27)	**On**	t'a vendu un livre pas cher	✓	✓
	$they_{non\text{-}ref}$ / we_{ref}	you have sold a book not expensive		
(28) a.	**Ils**	m'ont vendu un livre pas cher.	✓	✓
b.	**Eux** ils	m'ont vendu un livre pas cher.	*	✓
c.	**Eux**	m'ont vendu un livre pas cher.	*	✓
	they	me have sold a book not expensive		

2.4.4. Non-Referential Datives

Contrary to other pronominal objects, non-argumental datives such as the boldfaced French and Slovak pronouns in (29):

(29) a. Je vais **te** lui foutre une de ces claque !
b. Ja **ti** mu dám takú facku !
I will you him give such a smack ! = "By Joves, I'll give him a blow he'll remember !"

---

13 This is strikingly shown by the fact that when the referent of the focalised pronoun is prominent in the discourse, the strong form is NOT possible, in accordance with the choice-principle (22): only [ia] is a possible continuation in the "dialogue" below.

[i] Je te casserai la gueule.
a. Tu parles, je ✓TE casserai la gueule.
b. Tu parles, je casserai la gueule *A TOI.
You bet, I YOU will.break the face

This rather clearly illustrates that there is no preference to stress strong forms, but rather that two independent factors intervene: (a) deficient pronouns are limited w.r.t. their referent, (b) whenever possible, a deficient pronoun is chosen over a strong one.

14 The same holds of deficient subjects in Northern Italian dialects (P. Benincà (p.c.)), cf. §3.1.

15 This is one of the many cases in which deficient pronouns are restricted to [+human] reference (see fn. 59 for an account of this particular case).

do not have any referent. They are rather similar to "discourse-particles". Such an interpretation is totally impossible with strong pronouns: [16]

(30) a. * Je vais lui foutre une de ces claque **à toi** !
b. * Ja mu dám takú facku **tebe** !
I will him give such a smack to you!

2.4.5. [±Human]
The differing behaviour of deficient and strong pronouns w.r.t. human referents is amply illustrated in the introductory examples (§1): strong forms, contrary to deficient forms may not refer to non-human entities (the reverse is not true, cf. also fn. 15). [17]

2.4.6. Summary
The surface interpretive asymmetries involving deficient pronouns are: [18]

(31)

		semantics				
		must have D-antecedent (i.e. ostension, contrast, etc.)	expletive	impersonal	non-referential dative	possibly non-human
personal pronouns	strong	–	–	–	–	–
	deficient	+	+	+	+	+

2.5. Semantics: Range

Although descriptively correct (to the best of our knowledge), the preceding generalisation (31) is redundant.

2.5.1. The ban on strong pronouns as expletives and as arbitrary subjects of impersonals repeats twice the same fact: a strong element is incapable of being a semantically vacuous subject, it must be referential. Deficient elements on the other hand do not need to be referential and can be semantic dummies.

2.5.2. Similarly, strong pronouns are capable of being referential without being associated to an antecedent prominent in the discourse. Deficient pronouns cannot refer unless they are associated to such an antecedent. Again, strong pronouns are referential in a way in which deficient pronouns are not.

To capture the uniform asymmetrical behaviour of the two classes of pronouns with respect to "referentiality", unifying expletives, impersonals, and the need for a prominent discourse antecedent, some notion of "referential deficiency" is needed. Deficient pronouns are, in some sense to be defined, "less"

---

[16] The non-referential datives are to be kept apart from benefactive (/ethical) datives, which are equally non-argumental but which are referential to the same extent as other deficient pronouns. They always refer to a "benefactor". The two constructions are often found in minimal pairs such as:
[i] a. Je vais me manger un pomme. (benefactive)
I will me eat an apple = "I will eat myself an apple"
b. Je vais te manger une (de ces) pommes! (either benefactive or non-referential)
I will you eat an apple = "I will eat one of your apples" (benef.) / "I tell ya, I'm gonna eat an apple like..." (non-ref.)
The gloss of the non-referential examples is misleading: in the non-referential reading, these examples involve no second-person addressee. There is no referent to these pronouns, even derivatively.

[17] The asymmetry between some pronouns being able to refer only to human entities and other being able to refer to non-humans is noted from the earliest stages of grammatical research. Cf. for instance the *Grammaire Générale et Raisonnée de Port-Royal*, Arnauld & Lancelot (1846:319) quoting Reignier "*lui, elle, eux , elles...* , avec des prépositions, ne se disent guère que de personnes. Car quoiqu'un homme dise fort bien d'un autre *qu'il se repose sur lui de cette affaire*, ... on ne dira pas cela d'un lit ou d'un baton".
This asymmetry has then repeatedly been noted 'in passing', Damourette & Pichon (1911/1952), Perlmutter & Oresnik (1973:439), Kayne (1975), Jaeggli (1982:41), Rizzi (1982), Zwart (1992), Haegeman (1994) and has only recently received closer attention: cf. Berendsen (1986), Schroten (1992), Corver & Delfitto (1993).

[18] As a further semantic property, idioms often distinguish two series of pronouns (i.e. the two series are not interchangeable in idioms). This does not add anything beyond (re-) making the point that the distinction between the two classes is valid. Cf. Berendsen (1986), quoted in Zwart (1992), for Dutch.

referential than strong pronouns. They do not need to refer, and upon doing so, are dependent on the presence of an antecedent. [19]

2.5.3. Non-referential datives are one more instance of the same pattern: only deficient pronouns can be non-referential. Strong pronouns, as with expletives and impersonals, are incapable of occurring in referentially vacuous contexts.

2.5.4. The notion of "referential deficiency", or "being less referential than" is obviously far too vague as such. The comparison of impersonal and generic pronominal subjects however allows a much more precise characterisation of the semantic difference between deficient and strong pronouns.

Impersonal and generic pronominal subjects are similar in not being strictly referential (without being expletives), but minimally differ in that generic subject pronouns contrary to impersonal ones, may be strong:

(32) a. (***eux**) ils m'ont vendu des livres écornés (impersonal 3.pl.pron)
they they me have sold some books rotten

b. *les temporaires*, (**eux**) ils me vendent toujours des livres écornés. (generic 3.pl.pron)
the temporaries they they me sell always books rotten

c. *à NY*, **toi t**'es / **vous vous** êtes à peine arrivé(s), que les autres y sont d'ja tous à la sortie.
in NY you(impersonal) are just arrived, that the others are already all in the exit (generic 2.sg/pl.pron)

d. **les carottes** sont bonnes pour tes yeux (lexical generic)
carrots are good for your eyes

2.5.5. Since there is no clear sense in which the boldfaced pronouns of (32b-c) are more referential than that of (32a), non-referentiality as such cannot be the reason for the inacceptability of (32a). The impossibility of strong pronouns as impersonals must be linked to some other property distinguishing generic from impersonal constructions. There are (at least) five such differences: [20]

(i) impersonal subjects are existentially quantified, generic subjects universally
(ii) impersonal reading requires specific time-reference while genericity forbids it
(iii) impersonal but not generic subjects must be underlying subjects (non-ergatives)
(iv) impersonal but not generic subjects forbid inclusion of the speaker in their reference
(v) impersonals forbid but generics requires a range-restriction on the subject
(either by a dislocated noun-phrase, (32b), an adverbial, (32c), or from the lexical content of the generic itself, (32d))

The exclusion of strong pronouns as impersonal but not generic subjects cannot be due to one of the first four properties of impersonals. Strong pronouns may have existential import, are not incompatible with specific time-reference, are not restricted to underlying subjects, and may refer to the speaker.

2.5.6. The fifth property describes the fact that a generic sentence is acceptable only if some property / range (other than that contained in the predicate) is associated to its subject: (32b-c) are not acceptable as generics if the italicised phrases are absent (lexical generics of the type (32d) trivially always have a range). On the other hand, impersonals do not require any such range restriction: no other property than that of having sold a cheap book is associated to the subject of (32a).

Not only *can* impersonal subjects be rangeless, but they *always* are so:

(33) a. They have cleaned a cow today in Switzerland.
b. They usually clean cows in Switzerland.

---

[19] Specificity has often been attributed to the *presence* of a feature (cf. Sportiche (1992), Uriagereka (1992) among others). The fact that there should be one common explanation to the possibility of deficient (as opposed to strong) pronouns as expletives and to their need of a discourse antecedent, renders it improbable that these properties be due to the presence of some feature. It is not very likely that the capacity to occur as an expletive subject is rendered possible by the presence of a feature. It thus follows that specificity, *viz* needing a discourse antecedent (cf. fn. 12), should be rather attributed to the *absence* of some feature / property in deficient pronouns. If there is some feature / property in strong pronouns which forces referentiality, it is its *absence* in deficient elements that allows them to be non-referential and forces them to seek an antecedent in order to be referential.

[20] Cinque (1988) notes the first four differences. The fifth difference is discussed for lexical generics and for 2nd person singular generic pronouns by Brugger (1990).

In its impersonal reading, starting a discourse with (33a), or its French, Italian or Slovak counterparts, does not imply anything about the cleaners: anybody could have done the cleaning, i.e. "somebody, whoever, cleaned a cow today and this event of cow-cleaning-today took place in Switzerland"). But (33b) requires the cleaners to be inhabitants of Switzerland (in a broad sense of the term). In other words, the generic subject but not the impersonal subject is associated to the range-restriction in Switzerland. *No restriction is ever associated to the subject of impersonal constructions*, even if adverbials are present. [21]

2.5.7. Since no range-restriction is associated to (quasi-)expletives either, the property *having a range* correctly draws the line between those constructions which exclude strong pronouns (expletives and impersonals) and those which allow them (generics and referential contexts).

(34) subjects of impersonal and expletive constructions are never associated to a range-restriction
subjects of generics and referential subjects are always associated to a range-restriction.

2.5.8. Thus, *not being associated to range* seems to be the appropriate formulation of *being referentially deficient*. Strong pronouns, since they cannot be expletive or impersonal, must always be associated to some range. Deficient pronouns, since they cannot be interpreted without a (non-deficient) antecedent, never have their own range-restriction, but rather associate to that of their antecedent. [22]

The following generalisation now correctly brings together the four initial asymmetries (regarding prominent antecedents in discourse, expletives, impersonals, non-referential datives), without overgenerating:

(35) *Semantic Asymmetry #2*
a. Deficient pronouns are incapable of bearing their own range restriction
(and are therefore either rangeless (expletives, impersonals, non-referential datives), or associated to the range-restriction of an element prominent in the discourse)
b. Strong pronouns always bear their own range-restriction.

2.5.9. More speculatively, the fact that strong pronouns always require a range could be extended to the last semantic property distinguishing the two series, ±human reference, thus extending the generalisation fully. If strong pronouns must always have a range, independently of that of an antecedent, they are faced with a contradiction: having no nominal head including a range, they must contain a range but do not contain one. In this case, +human may simply be the default-range of human language. Again, the meaning of "having a range independently of that of an antecedent" can only be clarified within a formal proposal, §5.4.

In sum, all semantic properties distinguishing strong from deficient pronouns reduce to a single primitive: having a range or not. The precise grammatical representation of these generalisations is taken up in §5.

2.6. Phonology and Prosody

Phonological processes such as sandhi rules distinguish strong from deficient elements. French liaison seems to apply only with deficient elements. It is grammatical in the simple sentence (36a), where the pronoun *elles* may be analysed as deficient, but ungrammatical in the preverbal position of complex inversion, which requires strong pronouns (cf. *Lui/ *Il a-t-il dit la verité?* 'he has-he said the truth?') (underlining in (36) indicates that the final consonant is pronounced):

---

[21] Existential bare plurals are an intermediate object between impersonal pronouns and generics: they share with impersonals all above properties except that of (a) not being restricted to deep subject position, (b) always having a restrictor (the lexical element itself). Under a broader view of "strong" elements not restricted to pronouns (§9), they are strong elements. From the minimal pair formed by impersonals and existential bare plurals, only two properties qualify for the ban on strong pronouns: deep subjecthood and restrictors. Again, only the latter is plausible.

Past tense cannot be the restrictor in (32a) since it is incapable of being the restrictor of a generic (while being compatible with genericity): **(in NY), you couldn't walk alone.*

[22] A formulation in terms of "range" is also empirically much superior to one in terms of "reference": the former but not the latter correctly subsumes all non-referential deficient elements which nevertheless require a discourse antecedent, such as the partitive pronoun *en / ne* of Romance, or predicative deficient pronouns (cf. fn 12).

(36) a. ✓ **Elles** ont dit la verité.
b. * Quand **elles** ont-elles dit la verité?
they have-they said the truth?

Contrary to strong pronouns and nouns, deficient pronouns are able to form a single prosodic unit with an adjacent lexical element. This is independent of the prosodic weight of the (pro)nominal: the proper name *Al* contrasts with the pronoun *il*. The relevant prosodic domains are taken to be as indicated by underlining: [23]

(37) a. Al mange beaucoup. a'. Jean voit Anna
b. ☞Lui mange beaucoup. b'. Jean voit ☞elle
c. Il mange beaucoup. c'. Jean la voit.
Al/he eats a lot John her sees Anne/her

Finally, reduction phenomena are only found with deficient pronouns; in English, for instance, strong pronouns (e.g. in a coordination) cannot undergo reduction: [24]

(38) a. ✓ I saw **'ya** in the garden.
b. * I saw **'ya** and John in the garden.
c. ✓ I saw **you** and John in the garden.

These asymmetries may be subsumed under (again a general notion to be clarified by the theory, cf. §5.5 for a tentative proposal): [25]

(39) *Prosodic asymmetry*
Deficient but not strong pronouns may prosodically restructure.

This is, to our knowledge, the only prosodic asymmetry between strong and deficient pronouns (cf. §2.8).

2.7. Summary

The asymmetry between those pronouns which can, and those which cannot, be coordinated is perfectly correlated to a large number of other asymmetries, both syntactic, semantic, prosodic and morphological, uniform across widely different languages. These asymmetries divide into two types: relational properties, which link the two series, (40), and monadic properties, holding of one series but not of the other series, (41):

(40) a. deficient pronouns *are reduced w.r.t.* strong ones, if a difference obtains (morphology)
b. where possible, deficient pronoun *are preferred over* strong ones (choice)

(41) a. only deficient pronouns must occur at S-structure in a special derived position (syntax)
-> cannot occur in base position, dislocation, cleft, etc.
b. only deficient pronouns cannot be coordinated and c-modified (syntax)
c. only deficient pronouns may prosodically restructure (prosody)
-> liaison, reduction processes, prosodic domains
d. only strong pronouns bear their own range-restriction (semantics)
-> prominent discourse-antecedents (ostension, contrastive focus), expletives, impersonals, non-referential datives, reference to human entities only

The trigger of these asymmetries is exceptional not only in having such wideranging and crosslinguistically uniform consequences, but also in being a purely grammatical, i.e. abstract, property not correlated to any interpretational feature. This last point is illustrated both (i) by the fact that none of the surface interpretational

---

[23] The link between the (syntactic) property of being strong/deficient and the (prosodic) property of destressing and of contraction, seem to be one of the rare very robust syntax-prosody correspondance. It is all the more interesting that this correspondance seems to be generally valid across languages.

[24] If only deficient pronouns may be destressed and contracted, then this is the strongest evidence for the existence of an otherwise quasi-untestable systematic homonymy of strong and deficient pronouns in English: pronouns such as *him* may both be coordinated (and are therefore strong) and may form a unique prosodic domain with a left-adjacent verb (and are therefore deficient), cf. fn. 1.

[25] We are here borrowing and slightly changing a term from Nespor & Vogel (1986). A more precise version of this constraint would require data about prosody which do not seem to be available.

asymmetries of (41) strictly covary with the class distinction (possibilities of the two classes overlap), and (ii) by the fact that both semantic and phonological features are present: given the strict disjointness of semantics and phonology, a trigger which is purely internal to one of the two would not explain the properties of the other (cf. also fn. 26).

2.8. Annexe: Against "Focus"

2.8.1. Due to what is probably a historical accident, the (inaccurate) generalisation that deficient pronouns cannot be stressed has come to be viewed as a fundamental property of deficient elements. Informally, the fact that deficient pronouns do not occur coordinated, modified or with ostension, has been linked with the fact that deficient elements mostly occur unstressed, resulting in the claim that deficient pronouns do not occur in these cases *because* deficient pronouns cannot be (contrastively) stressed. Given the historical importance of this view, some space is devoted here to show that under none of its instantiations can this view be sustained. [26]

2.8.2. Once explicited, the reasoning seems to be:

(i) deficient pronouns (contrary to strong pronouns) cannot bear contrastive stress.

Together with the implicit assumption that

(ii) strong pronouns occur only where the deficient form is impossible (= (22)),

(i) would unify all cases, provided that

(iii) all contexts excluding deficient pronouns assign/require contrastive stress (overt in (21b)).

2.8.3. Since the assumptions (ii-iii) imply that strong pronouns are always contrastively stressed, the entailed surface generalisations are that:

(42) a. deficient pronouns are never contrastively stressed.
b. strong pronouns are always contrastively stressed.

Because contrastive stress involves both prosody (*prosodic focus*) and semantics (*semantic focus*), (42) can be taken to be a generalisation either about semantics, or about prosody:

(43) a. deficient pronouns are never {semantically / prosodically} focussed.
b. strong pronouns are always {semantically / prosodically} focussed.

2.8.4. Prosodic judgments (§2.8.5), semantic judgments (§2.8.6), and distributional facts (§2.8.7), all invalidate (43). None of the four generalisations involved are correct statements about the prosody and semantics of personal pronouns. Ultimately, both the hypotheses (i) and (iii) above are too strong. [27]

2.8.5. Against Prosodic Focus

a) *Unstressed Strong*. The version of (43) which chooses prosodic focus as the primitive for (21) is the less defensible of the two.

It entails that all coordinated pronouns, modified pronouns, post-prepositional pronouns, clitic-left dislocated pronouns, pronouns with ostension, etc. are always prosodically focussed. But this does not seem the case. The most minimal pair is given by ostension and contrastive stress (i.e. 21b-c)): in *d'abord j'ai vu Jacques et ensuite j'ai vu* ☞*lui* "first I saw J. and then I saw him" the two objects may have similar flat prosodies, while still excluding deficient pronouns. Simpler examples making the same point include most modified pronouns such as *Jean a vu seulement lui* "John has seen only him". The (absence of) prosodic accentuation on such modified pronouns contrasts very clearly with the strong prosodic accentuation in

---

[26] In its more radical versions, this proposal seeks to derive all asymmetries linked to deficient pronouns from the unstressed nature of the latter. (This is most prominent in languages in which deficient elements are limited to roughly the second position of the clause, as in many Slavic languages. On the empirical inadequacy of this approach, cf. *inter alia* Toman (1993)). Such an account is inadequate *in principle*: if prosody and semantics are not directly linked, postulating a unique prosodic trigger would leave semantic properties unexplainable, and postulating a semantic trigger would leave prosodic properties without a possible explanation.

[27] The simplest (and weakest) argument of all against both the semantic and the prosodic version of the claim that deficient forms are non-focussed stems from the observation that the strong contrastive stress present in (21b) is uncontroversially not required in the other contexts excluding deficient pronouns. One is then forced to invoke the existence of a lighter form of focus which excludes deficient elements and is present in all other cases.

constructions such as (21b). To unify the fact that both prosody-neutral ostension and contrastive stress legitimate a strong pronoun, a primitive different from prosodic focus is needed. (Prosodically unstressed strong pronouns are also clearly found in Italian left dislocation, cf. (47) below.)

b) *Stressed Deficients*. Deficient pronouns are not always prosodically inert. They may bear both word-stress, and phrasal stress.

(44) a. **Essi** vanno in chiesa. (word-stress)
they go to church

b. Non parlerò mai **loro**. (phrasal stress)
non I.will.speak never to.them

c. Mais regarde-**le** ! (phrasal stress)
but regarde-him

Examples discussed above (§2.4.1) show that deficient pronouns may also bear the strong prosodic focus associated to contrastive stress.

Since strong pronouns can be prosodically unaccented and deficient pronouns can be prosodically strongly accented, prosody cannot be the underlying factor guiding the distribution of deficient/ strong pronouns.

2.8.6. Against Semantic Focus

a) *Contrastive Deficients*. That deficient pronouns are never semantically focussed, is again incorrect. Examples (23) above show that deficient pronouns are perfectly compatible with contrastive focus, whenever the appropriate (independent) discourse conditions are satisfied. [28]

(45) a. ✓ B: Non, que **JE** mangerai ce gateau demain. (cf. (23b))
B: no, that I will.eat this cake tomorrow

b. ✓ B: .. je **TE** casserai la gueule ... (cf. (23c))
B: ... I YOU will.break the face

This alone falsifies the semantic version of (43).

b) *Non-contrastive Strong*. The claim that all strong pronouns are always semantically focussed, is slightly more difficult to disprove. This is due to the fact that it is always possible to construe a semantic contrast. In the absence of overt (observable) manifestations of such contrasts, the only possible direct argument against such claims is the equally untestable observation that many cases of coordination, clitic-left dislocation, etc. do not involve a greater dose of semantic focus than the usual use of a deficient element. The clearest case of all is that of prepositions. There is no sense in which a pronominal object of a preposition must always be semantically contrasted.

Under a flat intonation, the following example illustrates this point twice: the strong object of P, *eux*, and the strong coordinated subject, *lui*, receive no more semantic focus than the deficient, *le*.

(46) ✓ **Lui** et Marie **l**'avaient fait bien avant **eux**.
He and Mary it had done well before them

In Italian, left-dislocated strong pronouns may cooccur with a contrastively focussed constituent: given the generalisation that only one constituent per sentence may be contrasted through displacement to the left-periphery of its clause, the left-dislocated *lui* cannot be contrasted.

(47) ✓ **Lui**, QUESTO ha detto.
he, this has said

Again, semantic focus cannot be the primitive that excludes deficient pronouns from being objects of prepositions, occurring in coordination or in left-dislocation, since no semantic focus is involved. A primitive distinct from semantic focus is needed.

2.8.7. Strong pronouns are not focussed: GUN

A stronger argument to the effect that strong pronouns do not necessarily involve focus (semantic or prosodic) is provided by distributional facts from languages which overtly show both semantic and prosodic contrast through syntactic displacement of the contrasted element. In these languages, all focussed elements are

---

28 Another example of stressed deficient pronouns is reported in Zwart (1992, fn. 9).

displaced, but it is not the case that all strong pronouns are displaced. Strong pronouns therefore occur independently of focus (semantic or prosodic).

One such language is Gun, a rigid word-order African language of the Kwa family with a special focus-construction. In case an element is focussed, semantically or prosodically, a particle, *wè*, appears towards the front of the clause, and the focussed element obligatorily precedes it (the focussed argument is underlined). [29]

(48) a. ✓ N mon Mari
b. * N mon Mari
I saw Mari

c. ✓ Mari wè n mon
d. * Mari wè n mon
Mari FOC I saw

Constructions with coordination pattern exactly with those without coordination: an unstressed coordinated object, just as its non-coordinated counterpart, remains in situ (49a, c), while a stressed coordination must be placed in front of the focus-particle.

Now the vital fact is that focus on only ONE conjunct DOES trigger anteposition to the focus position, (49b, d). From this it follows that neither of the conjuncts of (49a) receives focus. Therefore coordination, in Gun, does not assign/require focus on the conjuncts, and the strong pronoun in (49a) is not focussed.

(49) a. ✓ N mon [Mari kpo éo kpo]
b. * N mon [Mari kpo éo kpo]
I saw Mary and him and

c. * [Mari kpo éo kpo] wè n mon
d. ✓ [Mari kpo éo kpo] wè n mon
Mary and him and FOC I saw

But, as a final stone to our demonstration, deficient elements in Gun still cannot be conjoined, cf. (50)-(51=8):

(50) a. ✓ N mon **éo**
b. ✓ N mo-é
I saw him

c. ✓ N mon [Mari kpo **éo** kpo]
d. * N mon [Mari kp(o)-é kpo]
I saw Mary and him and

(51) a. ✓ **Yélè** yon wankpè
b. * **Yélè kpo yélè kpo** yon wankpè
she[-human] and she and know beauty

In these two cases, the non-coordinable deficient elements cannot be excluded by semantic or prosodic focus, since the non-displacement overtly shows that no such focus exists. There must exist some γ distinct from semantic and prosodic focus which is capable of excluding deficient elements from coordination. [30]

2.8.8. The idea that deficient pronouns are somehow handicapped w.r.t. semantic or prosodic focus, popularised by the traditional account of the distribution of pronouns, is an artefact due the deficient pronoun's need of a prominent discourse antecedent, requirement mostly incompatible with the use of contrastive stress and ostension (cf. §2.4.1). As a result, both the premise and the conclusions of the traditional arguments are inaccurate. [31]

3. Two Types of Deficiencies

3.1. Regular Tripartitions

The partition of pronouns into two abstract classes, deficient and strong, is descriptively insufficient: regularly, pronominal systems divide into three distinct distributional patterns. The following are five among

---

[29] We owe these paradigms to the kindness of Enoch Aboh (who is not responsible of the use we make of them).

[30] Exactly the same argument holds of modification: modifiying a strong pronoun by c-modifiers such as *also* does not provoke anteposition and appearance of the focus-particle. But deficient pronouns are still excluded from such contexts.

[31] The Gun facts together with the French contrasted clitics (§2.4.1) lock up the back door which consists of postulating diverse types of focus and claiming that the above discussion is inconclusive because it fails to distinguish them. From the French facts it would follow that if there are two such types of focus, one of them, C, has the property of being compatible with deficient pronouns while being understood as contrastive. The other, F, is not contrastive but excludes deficient pronouns. Now in Gun, the C-type of stress would both trigger anteposition and be compatible with deficient pronouns. But this is a wrong conclusion: there is no stress which licences anteposition of deficient pronouns in Gun.

the numerous examples in which a language possesses three distributional paradigms. In each case, confronting the a- and the b-examples produces three patterns:

(52) OLANG TIROLESE (Oberleiter & Sfriso (1993))

a.	✓	**E:r**	isch	intelligent	b.	✓	**E:r und si:**	sain intelligent.
		he	is	intelligent			he and she	are intelligent
	✓	**Es**	isch	toire		*	**Es und es**	sain toire.
	*	**S**	isch	toire		*	…, daß **z und z**	toire sain.
		it	is	expensive			… that it and it	expensive are

(53) ITALIAN (Cardinaletti (1991))

a.	Non	**a lui*	dirò mai	**a lui*	tutto	**a lui**.
	Non	**loro*	dirò mai	**loro**	tutto	**loro.*
	Non	**gli**	dirò mai	**gli*	tutto	**gli.*
	no	to.him/to.them	I.will.say never		everything	
b.	✓ Non		dirò mai		tutto	**a lui e a lei**
	* Non		dirò mai	**loro** e loro	tutto.	
	* Non	**gli** e le	dirò mai		tutto	
	no	to.him and to.her	I.will.say never		everything	

(54) SLOVAK

a.	✓	**Jemu**	*to*	bude pomáhat'
	✓	**Ono**	*mu to*	bude pomáhat'.
	*	***Mu***	*to*	bude pomáhat'
		to.him / it / to.him	it will help	
b.	✓	**Jemu** a Milanovi	*to*	bude pomáhat'.
	*	**Ono** a to druhe	*mu*	budu pomáhat'.
	*	***Mu*** a jej	*to*	bude pomáhat'.
		to.him and toM. / it and the other / to.him and to.her		it will help

(55) ITALIAN VS. TRENTINO

a. ✓ **Lui** mangia della zuppa e — beve del vino (Italian)

✓ **Egli** mangia della zuppa e — beve del vino (Italian)
he eats of.the soup and drinks of.the wine

* **La** canta e — bala (Trentino)
she sings and dances

b. ✓ **Lui** e la ragazza del bar sono gli unici ad apprezzare tutto questo. (Italian)
He and the girl of.the bar are the only to appreciate all this

* **Egli** e il cavalier Zampieri sono gli unici ad aver apprezzato quel nobile gesto. (Italian)
he and the *cavalier* Z. are the only to have appreciated this noble action

* **La** e la Maria è vegnude algeri. (Trentino)
she and the M. are come yesterday

(56) FRENCH

a.	✓	**Lui** aime les choux mais — ne les mange que cuits?	b.	✓	**Lui** et son frère ont accepté ?
	✓	**Il** aime les choux mais — ne les mange que cuits?		*	**Il** et son frère ont accepté ?
	*	Aime-t-**il** les choux mais ne les mange que cuits?		*	Ont **il** et son frère accepté ?
		he likes the cauliflowers but not them eats other than cooked			he and his brother have agreed?

The tripartitions of pronominal systems are extremely regular across and within languages:
(i) out of all the possible combinations of strong and deficient (personal) pronouns inside a tripartite paradigm, only one is attested: two deficient and one strong. It is never the case that a tripartition stems from there being two strong and one deficient, etc. Similarly, it is never the case (to our knowledge) that there is more than three classes, with two types of strong and two deficients, etc.
(ii) out of all the possible relations between three pronouns, only one obtains, identical across all paradigms. It is not the case, as might be expected, that the two deficient pronouns are simply opposed to the strong series, as represented by $\{x_D, y_D\}$ *vs.* z_S. What systematically obtains is a hierarchy of the type $x_D < y_D < z_S$, where x_D is the pronoun in the third example of each paradigm, and z_S in the first. The second pronoun is systematically intermediate between the first, strong, pronoun and the third, sharing the properties

characterising deficiency with the third against the first (here exemplified by lack of coordination), but sharing some distributional properties with the first, against the third. In other words, what systematically obtains is a ranking in deficiency: the third pronoun is systematically "more deficient" than the second.
(iii) by themselves, the preceding regularities strongly indicate that the tripartition reflects the existence of three abstract classes of pronouns (rather than being due to the existence of two abstract classes - deficient and strong - plus a series of idiosyncratic and irregular properties among deficient pronouns). The strongest evidence to this effect is however the fact that each series has uniform properties across paradigms: in each case above, the pronouns contained in the second sentence (y_D :*es*, *loro*, *ono*, *egli*, *il*) share properties which distinguish them from the pronouns contained in the third sentence (x_D: *s*, *gli*, *mu*, *la*, *il*). The properties opposing the two classes of deficient pronouns, properties differentiating so to speak "severely" deficient pronouns (i.e. x_D above) from "mildly" deficient pronouns (i.e. y_D above), are briefly summed up below, but are discussed in details in Cardinaletti & Starke (1994a) for Germanic paradigms, and in Cardinaletti & Starke (1994b) for Romance languages (cf. also Cardinaletti (1993)).

Terminology. To distinguish the two types of deficient elements, we will borrow two terms often used as designations for deficient elements: clitic elements and weak elements. Although these terms are usually understood as interchangeable, they here acquire two distinct meanings: weak pronoun refers to the set of mildly deficient pronouns illustrated in the second line of each above example (y_D), while clitic pronoun is reserved to the severely deficient pronouns in the third (x_D).

3.2. {Clitic} vs. {weak; strong}: Severe Deficiency

3.2.1. In each of the above cases the clitic heads an X°-chain. In Olang Tirolese, the head status of the clitic *s* is evidenced by its impossibility in XP-positions such as V2-initial position, (52a). For the Italian objects in (53), the same point is most clearly illustrated by the fact that the clitic is "transported" by the verb over the realised subject in conditional inversions, *gli avesse Gianni parlato in anticipo, niente sarebbe successo* "to.him had John spoken in advance, nothing would have happened", i.e. "had John spoken to him in advance, ...". In Slovak, the second-position clitics strongly amalgamate with the verb when enclitic, and pattern together with "clitic" verbs, particles, etc. themselves clear heads. Finally, the Trentino *la* and French postverbal *il* in (55) and (56) are standardly analysed as heads, cf. among others Brandi & Cordin (1981, 1989), Rizzi (1986b), Poletto (1993) for Northern Italian and Kayne (1983), Rizzi (1986b) for French.

On the other hand, weak pronouns uniformly occupy positions which seem to be those of maximal projections:

- the V2-initial position in Olang-Tirolese, (52a), where only full phrases can appear;

- the specifier position of an intermediate functional projection in Italian
  (cf. (53a) in which *loro* is both (i) not picked up by the verb (contrary to clitics), and thus not adjacent to the verb, and (ii) in complementary distribution with an object floated quantifier (a maximal projection containing a trace) (cf. ??*Dirà loro tutto Gianni* "will.tell to.them all Gianni"));
- the sentence-initial position in Slovak (*ono* being the only Slovak deficient pronoun to be able to appear there), a position which is only available to topicalised and subject XPs (except for the special case of verb-inversion).
  Embedded contexts make this point even more clearly: in strings of the type ...*C°* α *clit*..., the element α must be either itself a clitic (clustering with the subsequent *clit*) or *one and only one* XP. Since the sequence ... *C° ono clit* ... is possible, while *... *C° XP ono clit* ... is impossible, *ono* can only be an XP.

- the shared subject of a predicate coordination in Italian and in formal French, (55a)-(56a), a position available to XPs but not to heads. [32]

---

32 The recognition of a class of weak pronouns distinct from clitic pronouns, but also deficient, allows:
(i) a principled approach to the traditional mystery of French object "enclitics" in imperatives: the first and second person pronouns are intermediate between usual French clitic pronouns (both are deficient, i.e. non-coordinable, etc.) and strong pronouns (the enclitics share their morphological form with the latter). In the present approach, such "enclitics" are really weak pronouns (the paradigm being *me* 'me' (clitic), *moi* 'me' (weak), *moi* 'me' (strong)), much like English or German, which have homophonous weak and strong object pronouns (*him-him*, *ihn-ihn*, cf. Cardinaletti & Starke (1994a)). The relevant difference between "proclisis"

3.2.2. The two series differ w.r.t doubling: doubling is always *clitic*-doubling, in the sense that doubling must always involve (at least) one clitic, no combination of weak and strong pronoun is possible. This is neatly illustrated with the Italian dative paradigm, in which the {*gli*; *loro*}, and the {*gli*; *a loro*} pairs are possible doubling pairs, but where {*loro*; *a loro*} is impossible:

(57) a.✓ Glielo'ho dato loro.
him it I.have given to.them
b.✓ Gliel'ho dato a loro / ai bambini.
him-it I.have given to them/to the children
c. * L'ho dato loro a loro / ai bambini.

A similar constraint holds of the Slovak *ono*, which is found doubled by the clitic *to*, as in the above example (54a), but is never doubled by a full phrase. Northern Italian dialects also (trivially) exemplify this: a subject strong pronoun occurs with a clitic as a doubler (cf. *Ela la canta* "she she sings"). [33]

3.2.3. In all cases above, a cooccurrence of clitics leads to the formation of a "clitic-cluster" with characteristic morpho-phonemic processes applying (e.g. in Italian, the vowel /i/ of a clitic is obligatorily lowered to [e] inside a clitic-cluster: *mi da un libro* -> *me lo da* "[he] to.me gives a book" -> "[he] to.me it gives"). On the other hand, no such process is attestesd in a combination of weak pronouns.

3.2.4. The cooccurrence of several pronouns leads to a sharper contrast with one combination: an accusative first or second person clitic can never cooccur with a dative third person clitic. The sharp ungrammaticality of such examples (****Il me lui présente* "he me to.him presents") is constant across Romance and Slavic languages, but also in many different language groups (cf. Laenzlinger (1993), Bonet (1994)). No ungrammaticality obtains when one of the two pronouns is a deficient weak pronoun, thus the following minimal pairs (both for proclitic and enclitic pronouns):

(58) a.** Gianni **mi gli** ha presentato / ... di presentar**migli**.
b.✓ Gianni **mi** ha presentato **loro** / ... di presentar**mi loro**.
c.✓ Gianni **mi** ha presentato **a loro** / ... di presentar**mi a loro**.
Gianni me to.him has presented to.them / ... to present me to.him/ (to) them

3.2.5. The fact that the two deficient series of pronouns individuated by distributional properties consistently pattern asymmetrically (together with the systematic regularity of the tripartitions), is a clear evidence for the presence of an underlying pattern. The fact that clitics are uniformly best analysed as heads, while weak pronouns are uniformly best analysed as maximal projections, provide a simple distinction between the two series. Further, all other morpho-syntactic asymmetries above may be restated in X-bar terms: a doubled pronoun cannot be a maximal projection, only heads form clusters, and only heads are subject to the accusative-dative constraint, whatever the source of the latter is.

3.2.6. From now on, the terms *clitic* and *weak* pronouns will be used in this *strict technical sense*: clitic elements are deficient (underlying) phrases which are heads at surface structure, and weak elements are deficient (underlying) phrases occurring as maximal projections at surface structure:

---

and "enclisis" must therefore be that imperatives, for some reason to be determined, render the clitic form impossible, and therefore the choice principle (22) forces the next stronger form, weak pronouns (see also Laenzlinger (forthcoming) for a treatment of these facts in terms of the clitic/ weak distinction);
(ii) a principled approach to the less-noted fact that Italian deficient pronouns split into those which must be adjacent to the verb, and those which are not (nominative *egli*, *essi*, dative *loro*, etc.), the former being clitics and the latter weak.
For more details on both these points, cf. Cardinaletti & Starke (1994b).

[33] The doubling patterns could be taken as evidence for the fact that declarative deficient subject pronouns are clitics and not weak in French. Doubling of the type *Jean il mange* "John he eats", if a consistent analysis of doubling was put forward, would indicate that *il* is a clitic in that case, contradicting the claim in the text. Without paradox, it seems to us that this is a correct conclusion: the register / dialect of French which admits doubling with flat intonation also requires repetition in coordination, while the register / dialect which allows shared deficient pronouns in coordinations does not allow doubling with flat intonation. Cf. Cardinaletti & Starke (1994b) for more discussion.

(59)	strong pronouns:	strong, full phrases	(*jemu* (Slovak), *lui* (Italian), ...)
	weak pronouns:	deficient, full phrases	(*ono* (Slovak), *es* (Olang-Tirolese), ...)
	clitic pronouns:	deficient, heads	(*mu* (Slovak), *lo* (Italian), ...)

As a historical note, let us note that although the terminology of "clitic" and "weak" is taken from the tradition, the present syntactic tripartition of pronouns has, to our knowledge, never been proposed before. Earlier uses of the term *weak* are either synonymous to "clitic", or mean "Germanic counterparts to Romance clitics". Two proposals are closer to the present one, but both are suggested for and applicable to a constrained set of phenomena, neither proposes a syntactic tripartion and neither presents a global system covering all types of pronouns (cf. §4-§7): the PF-clitic system (Kayne (1983)), with two syntactic classes, clitic and strong, and a PF-class, a notion by definition limited to those (non-clitic) deficient pronouns which must be adjacent to their predicate (such as French *il* but not English *it*); and the N*-system (Holmberg (1991)) with two classes, strong and N* pronouns, the latter being an entity ambiguous between heads, and maximal projections (cf. Cardinaletti & Starke (1994a) for more discussion). [34]

3.2.7. Since both deficient series must occur at S-structure in some functional projection of their predicate, it follows from the X-bar distinction between them that clitics occur in a functional head, while deficient elements occur in some specFP.

It has been abundantly illustrated that subject deficient elements such as Italian *egli* "he", French *il* "he" are restricted to a preverbal subject position: they can only occur in specAgrsP at surface structure. A similar situation obtains for objects. The *loro* paradigm (53a) transparently shows that weak datives obligatorily occur in a high position, above the standard position of their strong counterpart.

That weak pronouns are limited to a derived position also transparently holds of weak direct objects. This is clear for instance in the English particle verb construction (Johnson (1991)). [35]

(60) a.	He took	**it**	in	***it**.	because of the rain.
b.	He took	**John**	in	**John**	because of the rain.

Anticipating on non-pronominal weak elements, the same is visible in French with weak quantifiers thanks to the absence of past participle movement (Cinque 1994), and in Italian with weak adverb placement with respect to the weak demonstrative *ciò* (cf. also §9): [36]

(61) a.	Il a	**tout**	vu	***tout**.
b.	Il a	***l'ours**	vu	**l'ours**.
	he has	all/the bear	seen	all/ the bear

---

[34] Facts which do not fit neatly into the traditional bipartition have in fact often been noted, and "local patching" have sometimes been proposed. Three additional cases are: Cardinaletti (1991), whose discussion of the properties of the aprepositional dative *loro* prefigured much of the present work without formally distinguishing the three classes, Halpern & Fontana (1992), with their notion of X-max clitics, which are also maximal projections, but which cover essentially those deficient pronouns which appear towards the front of the clause; i.e. roughly Germanic and Slavic deficient pronouns, some of which are clitic, and some of which weak, in our terms, and Koopman (1993), discussing the complex Welsh pronominal system. Again, in all these cases the proposed system is similar in spirit to the present proposal but widely different both empirically, and theoretically.

[35] The formal identity between the particle construction paradigm (*Mary took **him** in ***him/HIM***) and Romance paradigms such as *Marie **la** voit ***elle/ELLE*** or *Maria **la** vede ***lei/LEI*** 'Mary sees her' (the strong pronoun is impossible unless the deficient is ruled out by non-prominent referent focalisation) now shows the path to the solution to the puzzle observed in fn. 9: the amount of focus needed in French is much superior to that needed in Italian. Logically, this could be either because French transparently shows the effect of the choice preference, and an independent factor softens the effect in Italian, or that Italian is transparent w.r.t. the choice effect and an independent factor worstens French. Since the amount of stress needed in English to allow a post-particle *him* seems to pattern with the Italian case, and not with French, the second path is more plausible (all the more so given that a similar conclusion holds of German with post-adverbial pronouns, ... *daß Hans **ihn** gestern ***ihn/IHN** gesehen hat* 'that John yesterday seen has').

The additional effect observed in French may be due to the fact that French uses cleft sentences as the unmarked contrast-marking construction, whereas Italian focussed objects may freely stay postverbally.

[36] *Ha studiato la storia poco* is acceptable if the adverb is stressed (or coordinated or c-modified, etc.), cf. §9.

(62) a. Ha studiato **ciò** poco ***ciò.**
b. Ha studiato ***la storia** poco **la storia.**
he.has studied this/ the history little

Since none of the weak elements interferes with A'-chains, and all surface in the position where an AgrP would be postulated, subject and object weak elements may be subsumed under one general condition: [37]

(63) Weak pronouns must occur, *at S-structure*, in a case-assigning specAgrnP

or if case is limited to specifier-head configurations more simply:

(64) Weak pronouns must occur in a case-position *at S-structure*

(Rizzi (1986a) as reformulated by Chomsky (1992) arrives at the same conclusion (modulo the notion of weak elements) on the basis of one weak pronoun: the null personal pronoun *pro*, cf. §3.4)

3.2.8. An additional prosodic asymmetry seems to separate the two types of deficient pronouns: while both types of pronouns may receive phrasal and contrastive accent (cf. §2.4.1. and §2.8.5. above), weak pronouns but not clitic pronouns may have (lexical) word-stress. All of subject *egli*, dative *loro*, subject *ono*, V2 initial *es*, etc. are not obligatorily destressed morphemes, but may bear usual word-accent. On the other hand, the clitic-morphemes *lo*, *mu*, *s*, etc. are consistently destressed. In somewhat metaphorical terms: while both series can acquire accentuation, only one of the two has it from the start.

Unfortunately, the category of weak pronouns having been little studied, if at all, no extensive investigation is available on their prosodic properties. As a consequence, two types of interpretations are *a priori* plausible, the former focalizing on the lexical form of the pronoun, the latter on class-membership:
(i) the clitic / weak contrast is irrelevant, what matters is the monosyllabic / bisyllabic distinction. All monosyllabic deficient pronouns lack word-stress and restructure prosodically, while neither of those properties holds of bisyllabic deficient pronouns. Under this interpretation the only relevant prosodic asymmetry is that between deficient and non-deficient pronouns: deficient, but not strong, monosyllabic pronouns lack word-stress and restructure.
(ii) the monosyllabic / bisyllabic distinction is irrelevant, what matters is the clitic / weak contrast. All weak pronouns can bear word-stress, while no clitic-pronoun does so, i.e. clitic pronouns always restructure, while weak pronouns optionally do so.

What is at stake is the restructuring capacity of bisyllabic weak pronouns, such as Italian *loro*, *egli*, Slovak *ono*, on the one hand, and word-stress properties of monosyllabic weak pronouns, such as German *es*, French *il*, or English *him*, on the other.

In both cases, available indications point towards the second interpretation: Nespor & Vogel (1986) note that the bysillabic *loro* may optionally restructure with a preceding verb, invalidating the claim that only monosyllabic elements restructure. German V2-initial monosyllabic deficient subject *es* may occur both as a reduced *'s* and as a full prosodic word with its own accent, invalidating the claim that all monosyllabic deficient pronouns prosodically restructure.

This is most clearly indicated by the distribution of the German glottal stop which is only found before the initial vowel of a prosodic word. The glottal stop may be found either only in front of the sentence-initial deficient pronoun *es*, or both before the sentence-initial *es* and before the verb. In the latter case, *es* forms a prosodic word, and thus bears its own word-accent: [38]

(65) a. [?]Es ist schön.
b. [?]Es [?]ist schön.
it is nice

---

[37] By specAgrnp, we leave open the question of specAgr$_{acc}$P vs. the specAgr$_{dat}$P. Higher Agr projections should be assumed for languages such as German and West Flemish, displaying deficient object pronouns in positions higher than negation (cf. Haegeman (1994)).

[38] French subject pronouns are apparently the strongest example of weak element which are systematically stressless (but cf. fn. 33). However, in a preliminary phonetic experiment, one author (Starke) found a harmonic break between a weak subject pronoun and the verb, which is usually taken to indicate a prosodic boundary (Vater (p.c.)). To the extent that this is a genuine phenomenon, the full generality of the above prosodic observation is supported.

Both these facts indicate the relevance of the clitic/ weak distinction for prosody (as opposed to the monosyllabic/ bisyllabic opposition). The most plausible format for the generalisation concerning the prosodic asymmetries thus seems to be: [39]

(66) a. deficient, but not strong, pronouns may restructure (liaison, reduction, prosodic domains), §2.6
b. weak, but not clitic, pronouns bear lexical word-accent

3.3. Relative Properties: Morphology and Choice

All properties separating deficient from strong pronouns uniformly hold of both clitic and weak pronouns. This is trivial for monadic properties, (41) (with the above proviso about prosody), but more interesting for relational properties, (40), which both extend identically to the third class of pronouns, transparently showing the ranking between the classes.

3.3.1. The morphological asymmetries between the three classes of pronouns give an explicit illustration of the relation between the three series, a representative sample of which is:

(67)	*clitic*		*weak*		*strong*	
a.	s	<	es			(Olang Tirolese)
	il	=	il			(French)
b.	ho			<	jeho	(Slovak)
c.			loro	<	a loro	(Italian)
			il	<	lui	(French)
			sie	=	sie	(German)

So that:

(68) clitic ≤ weak ≤ strong

The two deficient series are not simply opposed to the strong: weak elements enjoy an intermediate status.

3.3.2. Whenever the two forms are in principle possible, a deficient form takes precedence over a strong form, §2.3. This is true of both weak pronouns and clitics: descriptively, a strong form is impossible if a reduced form is at disposal. As soon as the reduced form is impossible (for independent reasons, here ostension introducing a non-prominent discourse referent and c-modification), the strong form is possible again.

(69) clitic < strong
a. ✓ Je **le** vois.
b. * Je vois **lui**
c. ✓ Je vois ☞**lui**.
I him see him

(70) weak < strong
a. ✓ **Il** me voit.
b. * **Lui** me voit.
c. ✓ **Lui** aussi me voit.
he (also) me sees

Whenever a clitic and a weak form compete, as in Olang-Tirolese, it is the clitic that takes precedence. It is only when the clitic is a priori disqualified, as in (71c), that the weak form may surface.

(71) clitic < weak
a. ✓ ... daß **z** toire isch
b. * ... daß **es** toire isch
... that it expensive is

---

[39] A much more fine-grained analysis would be needed: the three discussed properties of restructuring sometimes seem to be dissociated while restructuring is sometimes obligatory, with no clear correlation with classes, number of syllables, etc.

c. ✓ **Es** isch toire
it is expensive

The complete precedence pattern thus mirrors the morphological pattern: weak pronouns are again intermediate between clitic and strong forms.

(72) clitic < weak < strong

3.4. Null Pronominals

To the extent that *pro* is pronominal, it is a deficient pronoun. [40]

3.4.1. It has the semantics of a deficient pronoun, not that of a full (strong) pronoun. It can be expletive, (73a), impersonal, (73b), can have non-human referents, (73c), but cannot occur with ostension to denote a non-prominent discourse referent, (73d) (while nothing rules this out in principle):

(73) a. ✓ ***pro*** piove molto qui. (*lui (strong))
[it] rains a lot here
b. ✓ ***pro*** mi hanno venduto un libro danneggiato. (*loro (strong))
[they] me have sold a book rotten
c. ✓ ***pro*** è molto costoso. (*lui (strong, non-human))
[it] is very expensive
d. * ☞***pro*** è veramente bello. (✓lui (strong))
[it] is very nice

3.4.2. Its distribution is that of a deficient pronoun, not that of a full argument: Rizzi (1986a), as rephrased by Chomsky (1992), concludes that *pro* can only occur in a case-marked specAgrP, exactly mirroring the distribution of weak elements (§3.2.7). This conclusion is thus supported by two distinct studies, based on two independent sets of facts (on the other hand, it also entails that the restrictions on *pro* are due to its being weak, and not to its being null).

3.4.3. Given the choice between a strong pronoun and a *pro* counterpart, *pro* is always chosen:

(74) a. Gianni ha telefonato quando ***pro*** è arrivato a casa.
b. * Gianni ha telefonato quando **lui** è arrivato a casa.
Gianni has called when he is arrived at home

This is sometimes referred to as the "Avoid Pronoun Principle" (cf. Chomsky 1981), which is a special case of a much broader preference for deficient elements over their strong counterparts, §7.

3.5. Generalisations [41]

(75)

	morphology	choice	distribution		interpretation	prosody		X-bar
	+reduced		in FP at SS	*coord, ...	no range	reduction rules	no word-stress	X°
clitic	1	1	+	+	+	+	+	+
weak	2	2	+	+	+	+	–	–
strong	3	3	–	–	–	–	–	–

[40] Modern Greek seems to be an example of a language with tripartion including *pro*. Joseph (1993) writes "Greek provides an example of a language with a three-way distinction in pronominal realizations", referring to the strong (nominative) *aftos*, the deficient *tos*, and *pro*. From preliminary tests, *tos* qualifies as a clitic, thus reproducing the clitic (*tos*), weak (*pro*), strong (*aftos*) paradigm.

[41] Descriptively, there is a progression from most deficient to totally free element: affix -> clitic -> weak -> strong. In this work, we are concerned in solidly grounding and finding the primitive of the distinction (i) between strong and non-strong (deficient) elements, and to distinguish clitic deficient element from weak deficient elements. The distinction affix/ non-affix is irrelevant to these points and the fact that many properties that do distinguish deficient/strong elements (such as coordination, morphological reduction, etc.) would put affixes together with deficient elements is therefore irrelevant to the extent that there exist *some* clitic elements (in the technical sense), which are uncontroversially not affixes (which we take to be the case).

PART II. ... AND HOW TO ACCOUNT FOR IT.

PRELIMINARY: The A Priori Desired Result.

Given (75), the initial question:

- What is γ, the underlying (universal) trigger of (1) which provokes a wide array of distributional, semantic, prosodic and morphological asymmetries between two forms of one and the same pronoun?

can be meaningfully addressed. Since deficient elements divide into two coherent classes, two triggers are needed: one which causes weak deficiency, γ' (weak pronouns), one which causes severe deficiency, γ'' (clitic pronouns). [42]

Logically, these two triggers could be unrelated, or widely distinct. But facts indicate the contrary. Deficient characteristics (DC) of weak pronouns are a *proper subset* of the deficient characteristics of clitic pronouns (i.e. all characteristics that differentiate weak from strong elements are also shared by clitics):

(76) $DC(strong) \subset DC(weak) \subset DC(clitic)$ (trivially $DC(strong)=\varnothing$)

What is needed in order to *explain* this state of affairs is that the trigger which causes mild deficiency, γ' be *shared* by clitic and weak pronouns. The second trigger, γ'', is an exclusive property of clitics and adds itself to γ' to cause severe deficiency. Two unrelated triggers could only accidently produce the pattern (76).

Pattern (76) repeats itself with the two relational properties linking the three classes:

(77) a. clitic ≤ weak ≤ strong (morphology)
b. clitic < weak < strong (choice)

Again, their general format (x<y<z) is explained only if x contains the same trigger as y, plus its own additional trigger. Two distinct triggers leave as a mystery both the nature of this format and its recurrence across the three generalisations ((76), (77a), (77b)).

The generalisations (77a-b) require that both γ' and γ'' trigger the same property. By (77a) both γ' and γ'' trigger the property of being morphologically reduced with respect to an element which does not possess the trigger. By (77b) both γ' and γ'' trigger the property of being preferred over an element which does not possess the trigger. The similarity of effects of γ' and γ'' would be most elegantly explained if the latter are two formally identical triggers.

A priori then, the format of the solution to the puzzle of deficiency should be (i) clitic pronouns are deficient in two respects, γ' and γ'', while weak pronouns are only deficient in one of these two respects, γ', (ii) the two aspects of deficiency, γ' and γ'', are two (formally identical) instances of a more general underlying phenomenon, γ, the unique cause of (75) and the scope of this paper.

---

42 Two points made in the introduction also come out clearly from (75): (i) a trigger which explains only a subset of the asymmetries is inadequate, (ii) given both the range of properties involved (distributional, morphological, semantic and prosodic) and the fact that none of these properties systematically correlates with the class-distinction, it is unlikely that the primitive of the explanation be a purely prosodic property (which would make it impossible to address semantic properties), or a purely semantic property (which could not explain prosodic asymmetries). Syntax is the only component plausibly linked to all relevant types of asymmetries, and thus capable of addressing all facts.

DERIVATION: MORPHOLOGY

3.5.1. The most direct manifestations of γ are the two relational properties: contrary to all other characteristics, they are uniformly valid across all three abstract classes, across widely different languages, and, to anticipate, across grammatical categories (cf. §9). Further, given the hypothesis that inflectional morphology is relevant to syntax (a conception recently popularised by the work of Baker (1988) and Pollock (1989)) the morphological asymmetry is by far the most telling of the two.

3.5.2. We contend in fact that the simple observation that some deficient pronouns are morphologically a proper subset of the corresponding strong pronouns (and that the reverse never obtains), is all that is needed to explain everything concerning the three classes of pronouns, and this without changing anything to grammatical theory.

3.5.3. How are the following morphological relations betwen pronouns of distinct classes formally represented?

(78)	strong:	je-**ho**	je-**mu**	a **loro**
	deficient:	**ho**	**mu**	**loro**
		him, Slovak	to.him, Slovak	to.them, Italian

Minimally, the deficient element must be taken to contain less morphemes than its strong counterpart. Under the hypothesis that morphemes are heads of discrete syntactic projections, it follows that the number of syntactic heads *realised* by the strong form is bigger than that realised by the deficient element. (The Italian pair, in which it is not controversial that the dummy preposition is syntactically represented, is particularly clear in that respect).

3.5.4. This simple, and unavoidable, conclusion provides the explanation of the systematic morphological reduction of deficient pronouns, (79a). A more deficient pronoun is morphologically lighter than stronger pronouns *because* it contains less (underlying) morphemes, (79a-b), and it contains less morphemes *because* it realises less syntactic heads (79b-c).

(79) a. clitic ≤ weak ≤ strong
b. morph(clitic) < morph(weak) < morph(strong)
c. struct (clitic) < struct (weak) < struct (strong)

The existence of opaque morphology is the only reason that this relationship is not always visible at the surface, as it is in the preceding cases (cf. fn. 4). Unless similar morphological pairs are to receive distinct explanations, the conclusion reached on the basis of (78) must extend to pairs such as <lo ; lui> in Italian or <me ; moi> in French, and (79b) is literaly entailed.

3.5.5. Why is it that "the more a pronoun is deficient, the less syntactic heads it *realises*"? It cannot be a simple matter of spelling out fewer heads, if the systematic nature of the asymmetry is to be explained. It must rather be that the syntactic representation of deficient pronouns contains less elements to be realised: the more a pronoun is deficient, the less features / projections it contains. The syntactic structure of deficient pronouns is itself deficient, (79c). [43]

3.5.6. Generalising, this reasoning yields that weak pronouns realise less structure than their strong counterpart, and similarly, clitics are structurally impoverished w.r.t. their weak counterpart.

In other words, taken seriously, simple morphological observations virtually entail "that what makes a clitic pronoun be a clitic" is that the latter's syntactic representation is impoverished w.r.t. that of weak and strong pronouns (and similarly for the weak vs. strong distinction).

---

[43] This syntactic impoverishment may be due either to (a) some syntactic nodes of the reduced pronoun being (always) radically empty, or (b) the syntactic structure of the deficient pronoun containing less projections than that of the strong pronoun. Both implementations explain the syntactic asymmetry, and the choice between the two involves delicate questions about the nature of syntactic structure (must all projections always be projected?, what does it mean to be a radically empty projection?, etc.). As far as we can see, nothing below hinges upon the choice between the two implementations. The more radical version is however adopted in the text for simplicity of exposition: the more a pronoun is deficient, the less it has syntactic structure.

3.7. The unique and purely abstract primitive, γ, underlying all asymmetries linked to deficiency, across lexical categories is identified:

(80) *Structural Deficiency*
γ = lacking a set of functional heads

Structural deficiency is (potentially) capable of deriving all relevant asymmetries: it is uncontroversial that variation in syntactic structure triggers variation in morphology, prosody, semantics and distribution.

Structural deficiency is also the right notion to explain the fact that the deficient properties of weak elements are a subset of those of clitic elements, since these properties are to be imputed to the set of heads which is lacking in both clitic and weak elements (cf. the desideratum of the above "preliminary"). Finally, structural deficiency straighforwardly explains the general format of the relation between the three classes (x<y<z), since each class literally *is* a (syntactic) subset of the other, with the general relation:

(81) clitic = weak - γ'' = strong - γ' - γ''

3.8. What follows, is a discussion of how γ triggers the three remaining aspects of deficiency: (i) what is the structure missing in all deficient elements and how does it trigger the set of properties distinguishing strong forms on the one hand from weak and clitic forms on the other, §5 ?, (ii) what is the structure missing in severely deficient elements and how does it trigger the set of properties distinguishing strong and weak forms from clitics, §6 ?, and (iii) how does syntactic reduction trigger the choice preference, §7 ?

4. Derivation: Mild Deficiency

A large number of properties of the set γ' of syntactic heads lacking in both clitic and weak pronouns is already known, given the preceding reasoning and the discussion in §2-3:

(82) a. in transparent morphology, γ' is overtly realised as the morpheme(s) missing on the weak form, but appearing on the strong form (i.e. γ' = *Morph(strong) - Morph(weak)* , cf. (79), (81))
b. the absence of γ' forces the pronoun to occur in a functional projection at S-structure
c. the absence of γ' renders coordination and c-modification impossible
d. the absence of γ' correlates with the absence of a range-specification in the pronoun
e. the presence of γ' forces a +human interpretation
f. the absence of γ' legitimates prosodic restructuring and phonological reduction rules.

4.1. The Missing Morpheme

These properties now unambiguously identify γ', the surface morpheme which realises the syntactic structure present inside strong elements but missing in their deficient counterparts.

4.1.1. The vast majority of known <weak; strong> pairs are homophonous: this is the case in English <*him*; *him*>, German <*sie*, *sie*>, French <*elle*, *elle*>, etc. One pair with transparent morphology has however been discussed above: the Italian dative *(a) loro*. In this case, the above discussion entails (i) that the strong element *a loro* is literally constructed out of the weak pronoun *loro* plus the morpheme *a*, so that (ii) γ ' = *a*.

(83) a. Non regalerei mai **loro** tutto ***loro.**
b. Non regalerei mai ***a loro** tutto **a loro.**
no I.would.give never (to) them everything

This (surprising) conclusion is directly supported by two sets of facts:

4.1.2. The "dummy marker" *a* has exactly the right distributional property (82c): its presence/absence correlates with possibility of coordination and c-modification. The *loro* complement to *a* in (85) has properties similar to that of the weak pronoun in (84). Only the whole projection, containing *a*, can be coordinated and modified; the same is true for new referents under contrastive stress:

(84) a. * Ho parlato [loro e loro].
b. * Ho parlato solo [loro].
c. * Ho parlato [LORO], non [loro].

(85) a. * Ho parlato **a** [loro e loro].
b. * Ho parlato **a** solo [loro].
c. * Ho parlato **a** [LORO], non [loro].

(86) a. ✓ Ho parlato [**a** loro] e [**a** loro].
b. ✓ Ho parlato solo [**a** loro].
c. ✓ Ho parlato [**a** LORO], non [**a** loro].
I.have spoken *a* only *a* {them and them; only them; THEM not them}

This is not an idiosyncratic property of *loro* or of pronouns in general. It is always true that the complement of *a* behaves as a weak element: the same paradigm is reproduced with strong nominal complements.

(87) a. * Ho parlato **a** [tuo fratello] e [quel sindaco].
b. * Ho parlato **a** solo [tuo fratello].
c. * Ho parlato **a** [TUO FRATELLO], non [quel sindaco].

(88) a. ✓ Ho parlato [**a** tuo fratello] e [**a** quel sindaco].
b. ✓ Ho parlato solo [**a** tuo fratello].
c. ✓ Ho parlato [**a** TUO FRATELLO], non [**a** quel sindaco].
I.have spoken (only) to your brother (and/not to that mayor)

More generally, the complement of dummy markers mirrors weak elements: it is a maximal projection which may not occur by itself in θ- and A'-positions, coordination, c-modification and introduce new referents (by contrastive stress).

4.1.3. Second, dummy markers like *a* have exactly the right semantic property, (82e): dummy markers such as the Spanish *a* and the Rumanian *pe* force a [+human] interpretation. This is replicated in Central and Southern Italian dialects, with the dummy marker *a* which appears on left-dislocated accusatives, (89), in exact parallel to the asymmetry found with dative *loro*, (90). [44]

(89) a. ✓ **A** quella bambina piccola, la metto in primo banco
a this small girl, her I.put in first row

b. * **A** quella tavola rossa, la metto vicino alla finestra
a that table red, her I.put near to.the window

(90) a. ✓ Ho parlato **a** loro.
I.have spoken to them

b. * Ho aggiunto i pezzi che mi hai consigliato **a** loro.
I.have added the pieces that to.me you.have recommended to them

4.1.4. That dummy markers like *a* realise the missing piece of deficient pronouns is strikingly confirmed by Central and Southern Italian dialects in which the above morphological similarity is widely generalised: a dummy marker appears on all strong objects, whether nominal or pronominal, but on *no* deficient objects. In the dialect spoken in the town of Senigallia, for instance, the dummy marker is spelled-out as *ma* and may appear on both dative and accusative objects in the base position (examples from Sellani (1988)).

(91) a. ✓ tutt l' ser arconta **ma**$_{DAT}$ i fiulini (p. 9)
all the evenings she. tells-tales *ma* the children

b. ✓ e po' s' sent urlà **ma**$_{ACC}$ i venditori (p. 39)
and then SI hears shout *ma* the venders
"and then one hears the venders shout"

The very same dummy marker appears on strong but is missing in deficient pronouns:

---

[44] The dummy markers also have the correct morphophonological properties: that of being light morphemes. Comparison of morphological pairs across languages shows the morphological difference between weak and strong elements to be systematically tenuous though present.

(92) a. ✓ Ho vist **ma**lu
b. ✓ **l'** ho vist
him I.have seen *ma*-him

This is the clearest possible evidence to the effect that the presence / absence of the dummy marker is correlated to deficiency. [45]

4.2. Missing Structure

Granted that the dummy markers realise the piece of structure missing in both weak and clitic pronouns w.r.t. normal strong pronouns, what is this piece of structure?

4.2.1. Given all the above argument, the missing structure must be some projection *inside* the nominal phrase, i.e. a functional projection associated to the noun (as first argued by Vergnaud (1974)):
(i) since γ' is a (set of) functional projection ***belonging to*** a strong element but lacking in weak and clitic elements, it follows that γ' is a member of the "extended projection" of the strong element. [46]
(ii) in complex prepositions, such as *instead of*, dummy markers typically appear as the final element, and are syntactically independent of the first preposition. To capture the rigid ordering, constituency, etc., the only reasonable approach is to assume that the lexical prepositions are PPs taking a full nominal projection as complement, part of which is the dummy marker (Starke (1993a)).
(iii) The correlation between the +human interpretation of the head noun and the presence of the dummy marker *a* in Spanish or *pe* in Rumanian can only be (naturally) implemented if these dummy markers are functional projections of the noun. It is a minimal assumption that the functional projections associated to the lexical head L° contain features of L°. If this is the case, nothing need be said except that accusative markers are specified for +human feature, and are thus compatible only with nouns with human referents. On the other hand, if the dummy markers were not functional heads associated to the noun, they would need to *select for* a +human complement, a type of grammatical selection never attested otherwise (trivial cases of θ role assignment are irrelevant).

4.2.2. Since dummy markers like *a* always appear topmost (leftmost in the SVO languages discussed here) in nominal phrases, not only do they realise some functional projection of the noun, but they realise a high, or the highest, functional projection (where the "n" subscript on XP and YP indicates that they are functional projections of the noun):

(93) *strong*: ... parlare $[_{XP_n}$ a $[_{YP_n}$ loro
weak: ... parlare $[_{YP_n}$ loro
... to.speak [a [them

4.2.3. In naming the high nominal functional head realised by dummy markers, we follow Starke (1993a) and call it "complementiser", i.e. that which makes something become "complement of". The original rationale for this is the extensive syntactic similarity between the dummy markers appearing in "complex prepositions" (e.g. *instead of*) and the complementiser appearing in "complex complementisers" (e.g. *avant que* 'before that'). Several other reasons however point to the same direction:
(i) The numerous analyses exploring the path known as the DP-hypothesis arrive on the one hand at the conclusion that the D-node contains two distinct sets of features: φ-features (Brame (1981), Abney (1987), Giusti (1993) etc.), and referential features (Vergnaud & Zubizarreta (1990), (1992), Longobardi (1991) a.o.). On the other hand, it is widely concluded that a serious study of adjective placement and of prenominal modifiers (quantifiers, demonstratives, etc.) requires a large number of functional heads associated to the nominal head (Ritter (1990), Cinque (1993) a.o.).

Putting these two trends together with the conclusion that dummy markers realise a high nominal functional head naturally leads to a "split-DP hypothesis": the two sets of features attributed to D° are realised in two distinct functional projections: one containing φ-features, Y°, and spelled out as such, and the other containing referential features, X°, and spelled out as a dummy marker, if at all.

---

[45] Since the <lo ; malu> pair is a <clitic; strong> pair, it only shows that the dummy marker is correlated to some degree of deficiency, weak or severe. It is only in conjunction with the <loro ; a loro> pair that this argument bears precisely on mild deficiency.

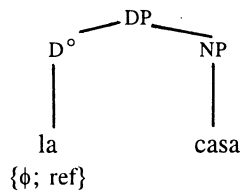
[46] The notion of Extended Projection is from Grimshaw (1991). It is used here in a loose sense, to refer to the unit formed by the lexical head and all the associated functional projection dominating it, where "associated to the lexical head" means "containing copies of features contained in the lexical head".

Now the parallelism between the topmost functional projections associated to the verb, and those associated to the noun, (94), is too striking not to be captured. In both cases, the highest realised layer contains a dummy morpheme (e.g. *that*, *of*), in both cases this dummy morpheme is (paradoxically) realising a head associated to abstract referential information of the whole phrase (i.e. range, one of the properties which distinguishes deficient from strong elements), and further, in each case, the next morpheme down contains agreement-type information:

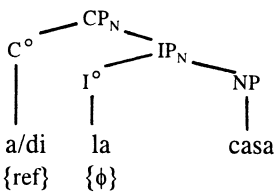
(94) a. $[_{CP_v}$ that {±wh} $[_{IP_v}$ {ϕ} [... $[_{VP}$]]]
b. $[_{XP_n}$ of / a {±range, ±human } $[_{YP_n}$ {ϕ} [... $[_{NP}$]]]

The most straighforward way to capture this parallelism is to assume that (94a-b) realise twice the same abstract structure: CP - IP - LexP. The traditional (95a) is split into (95b): [47]

(95) **a.**



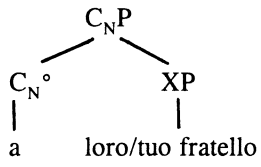
b.

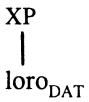


(ii) *a* and *di*, the realisations of the "to be identified" high functional layer of nominal phrases, are standardly taken to realise the topmost functional projection of infinitival phrases in Romance, i.e. C° (Kayne (1984), Rizzi (1982)). The proposed analysis renders this a natural fact: these markers always realise C°.

4.2.4. A functional preposition such as *a* in the above examples is thus interpreted as a nominal complementiser, which closes off the extended projection of the noun, exactly like the complementizer closes off the extended projection of the verb. Any strong element will contain such a complementiser-like preposition, whether realised or not (the identity of X below is irrelevant here, cf. §6).

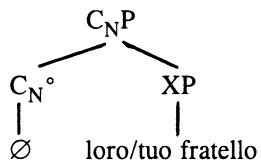
(96) a. strong b. weak





Strong elements appearing without a lexically realised preposition, for instance nominative and accusative DPs in Italian, are attributed the structure (97), which differs minimally from that of dative DPs, (96a): [48]

(97) strong(e.g. nominative/ accusative)



---

[47] Here and in subsequent representations, IP is used as a cover term for the (large) series of functional projections argued for in the above references. This proposal also implies that the definite article is not expression of the highest functional category, but of a lower functional head of the IP system. For the implications concerning clitics, often considered homophonous with determiners, cf. fn. 65.

That the syntactic representation of noun-phrases is similar, or identical, to that of verbal clauses is a hypothesis which has generated much recent work: Abney (1987), Szabolsci (1989), Siloni (1990), Cinque (1993), among others.

[48] The appearance of a functional preposition on accusative arguments (as in Spanish and Rumanian, cf. §5.1.3) also supports the above hypothesis that dummy prepositions are always associated to nominal extended projections, covertly or overtly.

Logically, the absence of the CP layer in deficient elements, must be the trigger of the remaining syntactic, semantic and prosodic asymmetries between deficient and strong elements: since absence of some XP (i.e. CP) is that which triggers the morphological reduction and the morphological reduction is correlated to all other asymmetries, all other asymmetries must be derived from the absence of XP (i.e. CP) in order to capture the correlation.

4.3. Syntax

The simple answer to distributional asymmetries between deficient and strong elements is that they are caused by the absence of CP in deficient but not in strong pronouns: because they lack CP, deficient pronouns must occur in some functional projection, cannot be coordinated, and cannot be modified.

(82) b. the absence of γ' forces the pronoun to occur in a functional projection at S-structure
c. the absence of γ' renders coordination and c-modification impossible

To go one step further, and explain why the absence of CP triggers these asymmetries, these asymmetries must be traced down either to the sheer absence of CP or to that of the content of C°. [49]

4.3.1. As earlier, morphology is an indicator of the solution: the morphological realisation of C°, the dummy marker, is commonly designated as a "(mere) case-marker". The distribution of noun-phrases with dummy-markers in one language largely corresponds to the distribution of case-marked noun phrases in other languages. [50]

We take this to indicate that the functional head which hosts the "reduplication" of the case feature of N° is C° (cf. fn. 46 for the assumed theory of functional heads). The distinction between, say, Italian and Slovak is that this case feature K is morphologically realised (if at all) on C° in Italian but on N° in Slovak (glossing over other differences, such as the relative richness of the distinctions morphologically expressed by K in the two languages).

Since deficient elements lack C°, they do not contain the (functional) case-feature. More precisely, *they cannot contain (functional) case-features*, the recipient of these features being absent.

Assuming, vaguely for the time being, that every noun-phrase must be associated to a functional case-feature (as opposed to the one on N°), it follows that deficient, but not strong, elements must undergo some process allowing them to be associated to the functional case-feature. In this context, the natural (and usual) interpretation of "x is associated with α" is that either x contains α or x is in a local structural configuration with an element containing α. If, as is often assumed, Agr° is necessary for case-assignment, deficient elements now need to occur in a local structural configuration with Agr°. Furthermore, weak pronouns have no space to represent the case feature internally (the locus of case is absent) and thus cannot "acquire" the functional case feature. If a weak pronoun is further displaced, the displacement destroys the local configuration with Agr° and the deficient pronoun lacks case again. The local relation between the weak pronoun and Agr° must be maintained as long as the pronoun needs case. [51]

4.3.2. Why do deficient pronouns need functional case? The central hypothesis of the present research is that the structural reduction observed in clitic and weak pronouns w.r.t. usual noun-phrases, is a *deficiency*, not a mere difference. As a deficiency, it must be compensated. A first tentative formulation of this is that (cf. §8 for discussion):

---

[49] This section is concerned only with mild deficiency, i.e. that what is common between weak and clitic pronouns, so (82b) requires some care. While it is evident that both weak and clitic pronouns must occur in some special derived position, the nature of this position seems substantially different in the two cases (X° vs XP). Accordingly, only that which is clearly weak deficiency will be addressed here, i.e. the placement of weak pronouns in a derived XP position, reserving discussion of placement of clitics for the next section, concerned with the derivation of severe deficiency.

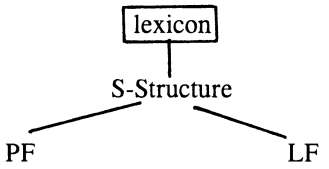
[50] As in §4 above, morphology is taken as an indicator of the underlying processes, not as the actual trigger of the surface asymmetries. This is not to be confused with so called morpho-syntactic accounts, which take morphology to be the "causal" factor.

[51] Proviso: as implemented here, the case-account is immune to a (strong) objection: that there exist deficient elements for which case is irrelevant (such as weak adverbs, §9). In the present approach these elements lack the features corresponding to their highest functional projection, C°, and these trigger deficiency. The nature of the feature in C° such that it generalises over adverbs, nouns, etc. remains an open question though.

(98) features missing in a deficient structure must be recoverable at all levels of representation

For deficient pronouns to be properly interpreted, the lack of functional case must thus be compensated. This can only be achieved through the establishment of an appropriate structural relation with Agr°. Given a model of grammar of the type, [52]

(99)



(98) entails that the relevant configuration must hold at S-structure (assuming there to be no displacement at PF).

Within a traditional X-bar model, "local configuration with X°" may mean one of two things: spec-head agreement with X°, or incorporation into X°. Weak pronouns being XPs, they establish a local relation with Agr° by appearing in specAgrP. It then follows from the above discussion that weak pronouns occur in a case specAgrP at S-structure, the generalisation to be derived. [53]

4.3.3. The ban on modification of deficient element (**I saw only it*) is to be traced down to the sheer absence of CP: these modifiers always modify a full clause, nominal or verbal, and never a subpart of the clause (*[only that it is so cold down there] bothers me* versus **[that only it is so cold down there] bothers me*). Their impossibility with deficient pronouns is a trivial consequence of structural deficiency. [54]

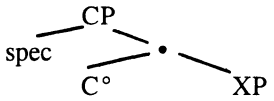
4.3.4. The ban on coordination of deficient pronouns could be treated similarly, given a theory of coordination of the type proposed by Wilder (1994): only CPs and DPs (i.e. only CPs, in our terms) can be coordinated.

In an approach to coordination in which any level of structure can be coordinated, one case is more delicate: that in which the conjunction (containing a deficient pronoun) occupies specAgrP (other cases are irrelevant: the pronoun is not in specAgrP and is thus ruled out exactly as dislocated or clefted pronouns). Nevertheless, the answer is straighforward enough: being embedded inside a coordination, the deficient pronoun is not in an adequate local configuration with Agr°, and is thus not associated to case, and, as a consequence, uninterpretable.

4.4. Semantics

4.4.1. The "referential" features usually attributed to the highest functional projection of noun phrases are referential indexes.

(100)



{index = ---}

---

[52] It is irrelevant whether "S-Structure" is taken to denote an actual level of representation, Surface Structure, or a point in a derivation to which spell out applies, Spellout-Structure, and similarly for the lexicon as insertion point vs. deep structure (cf. Chomksy (1981) vs. Kroch (1989), Chomsky (1992) for recent discussions).

[53] It is apparently strange that case is realised in C° but assigned by Agr°: given the strong similarity between nominal and verbal element, the locus of case should be uniform. This is however a false problem: the case features are always in C°, both in nominal and verbal extended projections (case is attested on verbal clauses across languages). Agr° on the other hand does not contain any case feature, but there is rather a "rule" akin to redundancy rules, which interprets all XP in specAgrP as associated to case.

[54] Whether c-modifiers are adjoined to CP, or they are in some higher position c-commanding the CP is not directly relevant. The second hypothesis is however favored by cases such as *seulement autour de la maison* versus **autour seulement de la maison* "only around of the house".

The exclusion of other modifiers, which occur neither with strong nor with deficient pronouns, (17a), must now be understood as a property of the L° lexical head of pronouns, the features of which must project onto functional categories that do not admit specifiers.

Deficient but not strong pronouns lack the highest functional projection, CP, and thus lack referential index. Also:

(82) d. the absence of γ' correlates with the absence of a range-specification in the pronoun
e. the presence of γ' forces a +human interpretation

Again, the simplest solution seems adequate: nothing need be assumed beyond (100) to explain (82d-e). If "referential index" is given its full meaning, that of associating a linguistic element with a non-linguistic entity, (82d) follows: having an index implies having a range. [55]

4.4.2. The exclusion of strong pronouns from expletives, impersonals, non-referential datives is straightforward: strong pronouns always have a CP and therefore contain an index and a range. Their having a range, is incompatible with occurrence in these constructions. On the other hand, because they have a referential index and a range, strong pronouns have no trouble denoting, even without being associated to an antecedent. [56]

Deficient pronouns on the other hand, have no CP, and thus no index. Lacking index, nothing forces them to be referential: they may occur as expletives, impersonals, etc. But since they lack a referential index, they can be interpreted as referential only if they are associated to a (non-deficient) antecedent, through coreference. As a consequence, deficient pronouns can only be referential if they are "old information", or "specific". They are uninterpretable in and by themselves. [57]

4.4.3. The fact that dummy markers differ w.r.t. the ±human characteristic, §5.1.3, could be taken to reflect their lexical specifications. Thus *a* in (101b) is lexically specified [+human] and only compatible with a [+human] noun (since the features of the noun must be identical to those in the functional heads). On the other hand, *of* in (101a) is lexically specified [±human] and thus occurs with both types of nouns.

(101)	a.	$[_{CP_n}$	of {±human}	$[_{IP_n}$ the	... $[_{NP}$ ✓car / ✓postman]]] ✓{-human} / ✓{+human}	(English)
	b.	$[_{CP_n}$	a {+human}	$[_{IP_n}$ (e)l	... $[_{NP}$ *coche / ✓ cartero]] *{-human} / ✓{+human}	(Spanish, accusative)

The requirement that strong pronouns refer to a +human entity could be exactly identical to (101b): the zero C° contained in strong pronouns, on a par with the Rumanian *pe*, the Spanish and Central-Southern Italian accusative *a*, is lexically specified [+human].

---

55 If there were reasons to keep the notions of range and index unrelated, only elegance would suffer: range would need to be postulated as an additional referential property of C° and some slight complication would be needed in the wording of §5.4.7. Similar remarks apply for human reference below.

56 We assume a theory of syntactic structure in which heads are nothing but the features "in" them and the presence of a node entails the presence of the features that constitute it (cf. fn.43). A C° without an index is not a possible entity.

57 This entails that coreference may be as in [ib]. Not only can two elements corefer by *refering* independently of each other to one and the same entity, [ia], but they can also do so if only one of the two refers, and the second is associated to the first, [ib] (i.e. the referent of the second is a function of the referent of the first), cf. also Fiengo & May (1994) for similar views. From now on, the word "coreference" will be used only for [ib].

[i] a.Coreference as (special) reference b. Coreference as a function of the antecedent

Deficient elements, such as the English *it*, are limited to the indirect path, [ib]. They never refer. (If coreference of the type [ib] did not exist, as is sometimes claimed, the "specificity" of deficient elements would be unformulable in any natural way which does not lose the correlation between (i) the asymmetry w.r.t. specificity and (ii) the asymmetry w.r.t. the capacity to be expletive).

(102) $[_{CP_n}$ ∅ $[_{IP_n}$ strong ... $[_{NP}$ ∅]]] (universal?)
{+human} * {-human} / ✓{+human}

Finally, the fact that deficient pronouns are free to refer to non-human entities trivially follows from (102): having no C°, they contain no +human specification and are thus in principle free to corefer with any (prominent) antecedent. [58]

(103) $[_{IP_n}$ weak ... $[_{NP}$ ∅]]] (universal?)
✓ {-human} / ✓{+human}

4.4.4. A still simpler, and more elegant account may however be closer to truth. It need not be postulated that C° contains two distinct (sets of) features (index/range and human). A range in C° is nothing but a set of features directly linked to interpretation. Since ±human is a feature in C° directly relevant to interpretation, it is best seen as *part* of the features which constitute a range, and not as a separate entity.

Now with respect to range, strong pronouns are in a contradictory situation: since they have a C°, the latter cannot be empty, and they must therefore contain some range-specification. On the other hand, they are associated to a dummy noun which does not provide any range-specification. To resolve the contradiction, a default range is inserted: +human.

There is substantial evidence that +human is indeed a default feature in natural language. Cf. fn. 59, and, among other, apparent "deadjectival nouns", such as *gli alti* ("the high", Italian), *the rich*. These can only be +human while the correponding adjectives are compatible with both human and non-human nouns (i.e. *the rich* may mean *the rich men* but not *the rich examples*). If such constructions include a null noun, the constraint on referents reduces to the fact that +human is a default range-feature.

As a result, only the accusative *a* in Spanish and Central-Southern Italian and *pe* in Rumanian need be lexically specified w.r.t. the ±human feature (i.e. +human only), all other complementisers are simply unspecified for this trait, and their behaviour follows from independent principles. [59]

4.4.5. An account in terms of default range is furthermore empirically superior to one in terms of lexical specification. The dummy marker appearing in the dative has distinct behaviour with nouns and pronouns: it is compatible with both human and non-human nouns, but only with +human pronouns:

(104)a.✓	Ho tolto una nota	al	{manoscritto; bambino}	
	I.have removed a footnote/mark	to.the	{manuscript; pupil}	
b.	Ho tolto una nota	a	loro	(= *manoscritti; ✓bambini)
	I.have removed a mark	to	them	

A lexical account would need to stipulate two distinct dative dummy markers, with no explanation of why one specified ±human occurs with nouns and not pronouns.

On the other hand, the facts follow straighforwardly if no lexical specification is involved in dative *a* (or English *of*): these markers are simply underspecified for the human feature, and take it from the head noun, (104a). With pronouns, the head noun does not provide any range-specification, and the default range is inserted in C°, +human, (104b).

---

[58] That the dummy nominal head is compatible with both values is attested by the zero noun of deficient pronouns and by overt realisations of the dummy noun, such as the English *one* (i.e. *the one I saw* may refer to both human and non-human entities). Nothing forces this though, cf. English, in which weak *it* is restricted to -human, and the weak version of *him/her* are restricted to +human. (cf. also fn. 15).

[59] Impersonals are particularly interesting: their interpretation is always arbitrary, i.e. associated to a default set of features, cross-linguistically including [+human], a range in the present assumptions. The difference between arbitrary and expletive subjects, one having both range and θ-role and the other neither, suggests a simple account of arbitrary reading: in the interpretive component, bearing a θ-role necessarily implies having a range. If this is the case, then deficient pronouns have three possibilites: (i) deficient pronouns may be θ-less and rangeless (expletives, discourse particles (non-referential datives)), (ii) deficient pronouns may bear a θ-role and acquire range through coreference, cf. fn. 57 above, (iii) deficient pronouns may have a θ-role but no range in syntax, in which case a default range is inserted at the (post-syntactic) semantic interface: +human.

In other words, if having an index entails having a range, not having an index does not entail not having a range.

4.4.6. All the facts linked to the lack of range of deficient pronouns ("specificity", expletives, impersonals and *non-referential datives), as well as the animacy asymmetry may thus reduce to the single fact that the highest* functional head associated to nouns contains a referential index, (100).

Whatever the fate of such a strong reduction, it is an empirical fact that this projection is linked to humanness, and it is barely controversial that it encodes referential features such as range. The link between the lack of CP and the wide number of apparently disparate surface semantic asymmetries, regarding contrastive stress (i.e. prominent discourse referent), ostension (i.e. prominent discourse referent), expletives, impersonals, non-referential datives, and specificity (i.e. prominent discourse referent), is thus established without special assumptions, through the notion of index (and range) in C°.

4.4.7. Putting the account of syntactic displacement (cf. §5.3.1-2) and semantic "non-referentiality" together now causes an unwelcome clash (although each is coherent in isolation): C° now contains two types of features: index (which entails range), and functional case. But the lack of index provokes different effects from the lack of case: case must be recovered (thus provoking displacement) while index may stay absent altogether (as in the case of e.g. expletives).

But this problem stems from the redundant assumption that C° contains *both* case and index. All and only strong elements have an index (cf. §5.4.2) and it is also true that all and only strong elements have a functional case feature (§5.3.1). This reduncancy vanishes if index is not a features *besides* K in C°, but rather *index is the interpretation of K.*

This now solves the apparent contradiction: it is still the case that all and only strong elements have range (since all and only strong element have an index, in turn a consequence of the fact that all and only strong elements have a functional case feature) thus deriving all the semantic asymmetries. Deficient element can but are not constraint to, corefer with an antecedent (cf. fn. 57), thereby seemingly inheriting index and range. It now follows that all deficient elements must recover case, while not all deficient element inherit index/range.

In short, both the obligatory (overt) displacement of deficient element and the whole range of semantic asymmetries follows from the presence/absence of one single feature of C°: functional case. [60]

4.5. Prosody

Although not much is known about the positive interactions between prosody and syntax, it is a standard assumption that prosody is sensitive to "major syntactic constituents" and that CP is such a constituent.

(82) f. the absence of γ' legitimates prosodic restructuring and phonological reduction rules.

Tentatively, it may be assumed that in absence of CP, a deficient pronoun does not qualify as a "major constituent": it does not constitute an independent (above word-level) prosodic domain, and it is subject to phonological rules / processes characteristic of non-major categories, such as reduction rules or liaison. None of this is true of strong pronouns which do constitute a major constituent, CP.

4.6. Summary

The general morphological asymmetry between deficient and full pronouns, together with the assumption that morphemes correspond to syntactic heads, leads to the conclusion that deficient pronouns correspond to less syntactic structure than full pronouns, a conclusion evident in transparent morphology (where the deficient form is a proper subset of the strong form).

From this it follows that "the missing piece" can be identified by a systematic morphological comparison between strong pronouns and (mildly) deficient pronouns, comparison which points to dummy prepositional markers. In turn, this entails that such markers are a functional projection of the noun, a conclusion supported by several independent studies. The similarity between this highest functional projection of the noun, and that found in the verb, is then so obvious that they are best viewed as two instances of one abstract category, here called complementiser. Finally, this topmost functional category is standardly taken to contain referential features (i.e. the referential index), while the dummy morphemes are typically associated to

---

60 Cf. Bittner & Hale (1994) for a recent discussion of KP, independently arriving at the same structural conclusion: noun phrases have a topmost functional projection which contains case (but not phi-features). In a different structural proposal, Giusti (1993) also argues against referential features in the topmost nominal projection and substitutes them by case information.

case, due to extensive distributional similarities. From these two features, the integrality of the asymmetries between deficient and full forms follows:

(105)a. Because strong elements have a CP but deficient elements lack it:
- deficient elements are morphologically "lighter" (i.e. have less heads to realise)
- deficient elements cannot be modified by modifiers of CP (c-modifiers)
- deficient elements are not "major constituents", and are thus subject to reduction rules, etc.

b. Because strong elements have functional case-features but deficient elements lack it:
- deficient elements are necessarily in a case-assigning specAgrP at S-structure (to recover case).

c. Because strong, but not deficient, elements have an index (from their functional case feature):
- strong elements cannot be expletive, impersonal, discourse-particles
- deficient elements cannot refer, they must associate to an antecedent prominent in the discourse (i.e. discourse-internal coreference: "specificity", "old information", etc., but also impossibility in most contrastive contexts, or with ostension)
- strong elements with dummy lexical heads (strong pronouns) are assigned a default range +human

5. Derivation: severe Deficiency

The properties of γ'', missing in clitics but present in both weak and strong pronouns, are:

(106)a. in transparent morphology, γ'' is overtly realised as the morpheme missing on clitics but realised on both weak and strong pronouns (i.e. γ'' = Morph(weak) – Morph(clitic)).
b. lack of γ'' causes X°-chain formation
c. lack of γ'' entails absence of word-stress

5.1. The Missing Morpheme

5.1.1. The single transparent <clitic; weak> minimal pair illustrated above, the <s; es> pair of Olang-Tirolese is not very informative. The Slovak <ho; je-ho> and <mu; je-mu> pairs discussed in §4.3 however provide valuable clues. First, the reduction is clearly not phonological, and second, the *je-* morpheme missing on the clitic is not plausibly assimilated to a dummy case-marker. As a consequence, the *je-* morpheme is not the realisation of γ' (C°) but rather of γ''.

5.1.2. A similar reasoning applies to a number of other cases, either closely related to Slovak, as in the Serbo-Croatian pair <ga, njega> "him", or more distant, the Spanish <los; el-los> "them", etc. (and maybe the Greek <tos; af-tos> "he", Joseph (1993)). To a certain extent, German provides the same type of clue: the distinction between forms such as the (ambiguous) weak form *ih-n* "him" and the corresponding (clitic) form *n*, found in dialects (cf. Abraham (1991)), results from the loss of a dummy morpheme *ih-*.

5.1.3. In each case what disappears is a semantically empty "dummy" morpheme which has no other apparent role than that of rendering an impoverished form stronger. To encode this into the terminology, we will refer to this dummy as a "syntactic support". [61]

5.2. The Missing Structure

5.2.1. The support morpheme attested in Slavic languages has the curious property of appearing as a support morpheme also outside of the realm of pronouns. The morpheme *je-* of Slovak pairs <ho; je-ho> and <mu; je-mu> also distinguishes, in one case, the clitic auxiliary from the strong copula: whereas the third person

---

[61] As noted by Davide Ricca (p.c.) an explanation in terms of phonological reduction is further not very plausible given the fact that the progressive phonological reduction of diacronic change systematically involve truncation of the ending of words, not of the initial part, while exactly the opposite holds of pairs of pronouns: it is systematically the initial phonemes/ morpheme which is deleted (while in many cases deficient pronouns still historically derive from full forms).

This strongly suggests that the diacronic process involved with deficient/ strong pairs is not an instance of the general reduction process but rather of the working of structural deficiency, maybe via the impetus of the choice principle (§2.3 , §7).

clitic is a null morpheme, the strong third person copula is realised as *je* (Toman (1981) discusses the difference between these verbal forms for quasi-identical Czech data).

This parallelism between nominal and verbal contexts is strongly illustrated in Serbo-Croatian: on a par with nominal pairs <ga; nje-ga> or <mu; nje-mu>, verbal pairs are systematically distinguished by a support morpheme (Browne (1974)):

(107)	*clitic*	*strong*		*clitic*	*strong*	
	sam	je-sam	I.am	smo	je-smo	we.are
	si	je-si	you.are	ste	je-ste	you.are
	je	je-st	he.is	su	je-su	they.are

5.2.2. Given the parallelism between the support in the verbal domain and in the nominal domain, and the hypothesis that nominal CPs parallel verbal CPs, conclusions from the study of one system may carry over to the other. This permits an understanding of the mysterious nominal support on the basis of its better studied verbal counterpart.

In Serbo-Croatian, the verbal support *je-* is part of a tripartite system of support prefixes: a reduced form such as *sam* "I.am" has as strong counterparts both *jesam*, in which *je-* is either emphatic or a simple dummy marker, and *nisam*, in which the support morpheme is interpreted as negative. The complete paradigm is illustrated in the following examples which show both the opposition between reduced and full forms ((non) string-initial) and the interpretive possibilities of the full forms.

(108)a. *	sam	ga pio	("clitic")
b.✓	je-sam	ga pio	(emphatic, dummy)
c.✓	ni-sam	ga pio	(negative)
	yes/no-am	it drank ("I.have(n't) drunk it")	

5.2.3. Exactly the same system is found in Basque: the bare form of the auxiliary is clitic-like (in the same sense as Serbo-Croatian), and the complex forms are non-clitic entering the same tripartite semantic system.[62]

(109)a. *	da	etorri	("clitic")
b.✓	ba-da	etorri	(emphatic, dummy)
c.✓	ez-da	etorri	(negative)
	yes/no-has	arrived	

5.2.4. The virtual identity of the Basque and Serbo-Croatian verbal support paradigms not only imposes a unified analysis in terms of deep properties of natural language (the two languages being largely unrelated), but also provides an analysis of the phenomenon: Laka (1990) argues at length that the Basque support prefixes correspond to a functional category between C° and I°, which she calls Σ° and which contains both polarity features (assertion / negation) and focus features.

5.2.5. Extending Laka's analysis not only to Serbo-Croatian but to all occurrences of support morphemes, the nominal support morpheme realises a nominal Σ° in a structure of the type: [63]

(110) C_LP Σ_LP I_LP LP (with L = any lexical category)

More generally, Σ° may be taken to be the locus of prosody-related features of L°. [64]

---

[62] The English *do* corresponds to *je-* in Serbo-Croatian and *ba-* in Basque in being a support morpheme (prefix) which is either dummy or emphatic. The necessity of a unified analysis of *ba-* and *do* is argued by Laka (1990).

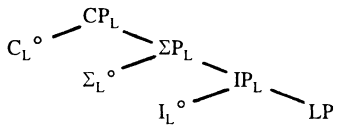
[63] What is here called Σ° has recently received widely different names: PolarityP, FocusP, Agr1P, AgrcP, WackP, etc. all denoting essentially the same entity.

[64] One feature may suffice: in the usual case, its negative value corresponds to negative interpretation, while its positive value is default (and non-realised) and corresponds to positive readings. Finally, a realised default value provokes emphatic (contrastive) reading. In Basque and Serbo-Croatian the default value may be independently needed for pure grammatical constraints and the emphatic reading is provoked where the support is realised without it being independently forced.

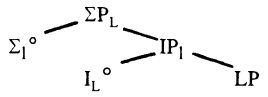
The fact that Σ° apparently contains both polarity features and focus features reflects a general (surprising) fact about language: non-lexical accentuation is largely related to affirmation / negation (i.e. to

5.2.6. Such a phrase-structure provides a pristine model of Structural Deficiency: just as weak element lack the superior layer of strong elements, CP, clitic elements lack the superior layer of weak elements, ΣP. Weak elements are "peeled" strong elements, and clitics are "peeled" weak elements. The resulting structure of the three classes is: [65]

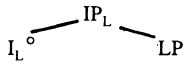
(111)a. Strong Pronouns



b. Weak Pronouns



c. Clitic Pronouns



5.3. Derivation: Syntax

5.3.1. To recover the features missing due to the lack of Σ°, a clitic pronoun must associate with prosodic features, a consequence of (98) above. There being no syntactic head which assigns such features structurally, clitic pronouns must surface in a local relation to a c-commanding Σ° itself.

With respect to syntactic placement, a clitic pronoun is thus faced with an apparent contradiction: to compensate the absence of functional case-features, it must occur in a specAgrP at S-structure and to compensate the lack of Σ° it must simultaneously occur in a local relation with Σ°.

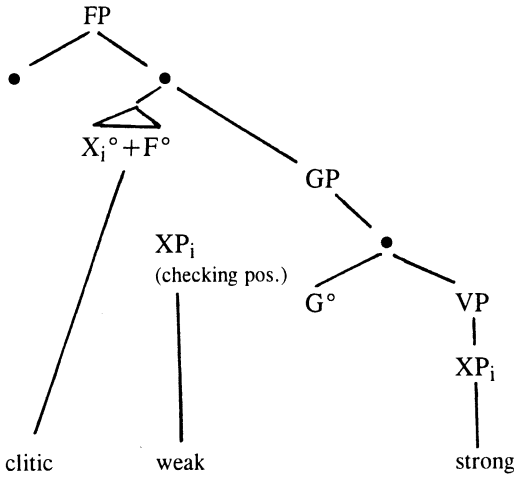
5.3.2. The only solution to this dilemma is to exploit both types of possible local configuration with an X°: specifier-head agreement and incorporation, through a derivation of the type:

---

emphatic and contrastive readings). It is an intriguing hypothesis that this link is a reflex of the fact that both polarity and accentuation features are realisation of one and the same set of features (i.e. one is derivative upon the other) realised in Σ°.

[65] Although the structure in (111c) could express the claim often made on the basis of Romance languages that (3rd person) clitics ARE determiners (compare (111c) with (95b)), it does not necessarily do so. This claim is in fact undermined by the observation that some languages manifest one paradigm but not the other: Slavic languages have clitics but not determiners, Brazilian Portuguese has determiners but no corresponding clitics. It would not be an unwelcome result that clitics realise more heads than determiners, given the non perfect homophony between the two paradigms (e.g. in Italian: *il/lo* (det.) vs. *lo* (3rd sing. clitic), *i* (det.) vs. *li* (3rd pl. clitic)).

(112)



Only through such a derivation is the clitic associated to both Agr° and X° at S-structure. [66]

5.3.3. That (112) is an accurate representation of clitic-placement has been argued many times, on grounds totally independent from the present concerns (mainly linked to agreement found on past-participles), cf. among others Sportiche (1989).

5.3.4. Incorporation opens an unexpected possibility: what is F° above? Obviously a head associated to the adequate (prosodic) features missing in clitic pronouns (but not in weak and strong pronouns). But there are two such heads: Σ°, by hypothesis, and (the head containing) L°, by definition, since a lexical head contains all features of its associated functional projections. It thus follows from the above system that the X° chain of the clitic has its head either in the head hosting V° or in Σ°.

This is precisely what is needed: typologically, clitic pronouns pattern in two groups: they appear either around the second position of the clause (C2 clitics), i.e. in Σ°, or on the verb (ad-verbal clitics), i.e. in the functional head hosting V° (cf. Starke (1993b) for discussion of C2 clitics along these lines). [67]

5.4. Derivation: Prosody

The fact that weak and strong pronouns, but not clitic pronouns, have lexical word-accent (possibly later erased in weak elements through prosodic restructuring) derives from the hypothesis that all prosody related features of L° are realised in Σ°. Clitic pronouns have no prosody related features and are syntactically associated to them only *pro-forma* through an adequate configuration. Clitic-pronouns thus end-up in the prosodic domain of an adjacent non-clitic element. Again, this is a sketch of a path which seems plausible, rather than a full fledged proposal.

6. Derivation: Choice of Pronouns

All but one of the asymmetries between clitic, weak and strong pronouns are now reduced to a unique underlying primitive, lack of the highest functional projection. The remaining fact to explain is the choice

---

66 A surprising result of this analysis is to *derive* the fact that there are three pronominal classes, and not two or four. This follows from the fact that there are only two possible (distinct) types of chains, XP and X° chains together with the strict locality condition on recoverability. In other words, the fact that there are two distinct types of chains entails that two elements at most may be recovered, and therefore that there can be only three classes of pronouns w.r.t. deficiency: non-deficient, deficients with one element to recover, and deficients with two elements to recover.

67 None of these possibilities are open to weak pronouns: weak pronouns are not forced to incorporate, the simplest option of spec-to-spec displacement being open to them. But displacement into specVP is impossible, this being the position of the external argument, and displacement into specXP where X° contains the verb, i.e. $[_{specXP}$ *weak* $[_{x°}$ *[v] x]* ... is plausibly not local enough to establish the correct configuration with V° (cf. Cardinaletti & Starke (1994a) for discussion of this last point).

preference. Given the existence of the three classes, every pronoun is potentially realised in three distinct ways; the profusion so created is regulated by a very strong, cross-linguistic generalisation:

(113) clitic < weak < strong

This generalisation captures asymmetries such as (115a-b) or (116a-b): the most deficient form must be chosen if it can be chosen (which is possible in (115b) but not in (116b) because deficient pronouns can refer to prominent but not non-prominent discourse referents, §2.4, §5.4). Similarly for (114b) vs. (115a): since there are no deficient nouns in French, *Jean* is the most deficient form possible in (114) and therefore allowed. In (115a) on the other hand, there exists a licit more deficient form, and the strong form is therefore not licit.

(114)a. A: Moi je vois Jean et Marie
b. B: Moi je vois Jean
I, I see Jean (and Marie)

(115)a. * Pierre voit lui_{odr} <$_{odr}$ = "old" (i.e. prominent) discourse referent>
b. ✓ Pierre le_{odr} voit

(116)a. ✓ Pierre voit lui_{ndr} <$_{ndr}$ = "new" (i.e. non-prominent) discourse referent>
b. * Pierre le_{ndr} voit
Pierre him sees him

6.1. Minimise Structure

Given the respective syntactic structures of the three classes, the generalisation (113) reduces to the statement that a "smaller structure" is obligatorily chosen, if possible:

(117) *Economy of Representations*
Minimise Structure

Only if the smaller structure is *independently* ruled out, is the bigger alternative possible. [68]

6.2. Null Pronouns

6.2.1. It is often held that a special filter, called "Avoid (lexical) Pronoun" in Chomsky (1981), applies to force the choice of null pronouns over their overt counterparts (in unmarked situations, i.e. where the referent is prominent in the discourse, as in the non-focussed coindexed embedded subjects):

(118) a. ✓ Gianni$_i$ partirà quando **pro**$_i$ avrà finito il lavoro. (Italian)
b. * Gianni$_i$ partirà quando **lui**$_i$ avrà finito il lavoro.
c. * Gianni$_i$ partirà quando **pro**$_i$, non il suo capo, avrà finito il lavoro.
d. ✓ Gianni$_i$ partirà quando **lui**$_i$, non il suo capo, avrà finito il lavoro.
Gianni will.leave when he (, not the his boss,) will.have finished the work

This filter reduces to Economy of Representations: the choice of *pro* over *lui* is a special case of Minimise Structure, *pro* being a weak pronoun and *lui* a strong pronoun.

Further, where the two proposals make different predictions, those of Minimise Structure, or Economy of Representations, are systematically favored over those of Avoid Pronoun. In languages which have two weak pronouns, one overt and one null, Avoid Pronoun requires that the null be chosen over the realised, where possible, exactly as in (118), while Minimise Structure leaves the choice free. The latter is correct: [69]

(119) a. ✓ Gianni$_i$ partirà quando **pro**$_i$ avrà finito il lavoro. (Italian)
b. ✓ Gianni$_i$ partirà quando **egli**$_i$ avrà finito il lavoro. (Italian)
Gianni will.leave when he will.have finished the work

---

68 Of course this may be translated as, or be derivative upon, "Minimise Features", cf. fn. 43. This holds of all subsequent discussion. See also Picallo (1994) for an "Avoid Features" view of Avoid Pronoun.

69 Although stylistic differences are involved, it is not the case that the two pronouns belong to disjoint registers, which would make the point irrelevant: at the stylistic level in which *egli* is possible, pro-drop is also allowed.

Conversely, if a language has two realised forms for one pronoun, one being deficient and the other strong, Avoid Pronoun, if anything, predicts a free choice, whereas Minimise Structure correctly requires the use of the deficient over the strong. I.e., only Minimise Structure captures the uniformity of the French (120), Italian (118), and Olang-Tirolese (121) paradigms:

(120) a.✓ Jean$_i$ pense qu'**il**$_i$ est intelligent (French)
b. * Jean$_i$ pense que **lui**$_i$ est intelligent
c. * Jean$_i$ pense qu'**il**$_i$ est intelligent, pas son chef
d. ✓ Jean$_i$ pense que **lui**$_i$ est intelligent, pas son chef
John thinks that he is intelligent (, not his boss)

(121) a.✓ Hans$_i$ denkt, daß **a**$_i$ intelligent isch (Olang Tirolese)
b. * Hans$_i$ denkt, daß **e:r**$_i$ intelligent isch
c. * Hans$_i$ denkt, daß **la a**$_i$ intelligent isch
d. ✓ Hans$_i$ denkt, daß **la e:r**$_i$ intelligent isch
Hans thinks that (only) he is intelligent

Similarly, Avoid Pronoun cannot explain the choice among objects pronouns across Romance, Slavic or Germanic: a realised object clitic or weak pronoun is chosen over an (equally realised) object strong pronoun, (115). Finally, Economy of Representations explains the noun-pronoun asymmetry, (114b) vs. (115a), while an approach in terms of Avoid (lexical) Pronoun would have nothing to say about this case.

In sum, Minimise structure, but not the Avoid Pronoun Principle, captures the parallelism between the behaviour of subjects and objects, between null-subject languages and non-null-subject languages (and across categories, §9) while explaining the noun-pronoun asymmetry. Such a coverage gives retroactively strong credence to the primitive upon which it is based: the tripartition between clitic, weak and strong elements and Structural Deficiency.

6.2.2. Most other filters / principles which have been stated in terms of null vs realised pronouns are to be restated in terms of the interplay between Minimise Structure and the three classes of clitic, weak and strong pronouns. Montalbetti's Overt Pronoun Constraint is a good illustration of this. Montalbetti (1984) observes that *if mild focalisation is ignored*, the following paradigm holds (the paradigm is slightly adapted and transposed from Spanish to Italian):

(122)a.✓ Gianni ha ammesso che ***pro*** ha bevuto tutto il vino.
b.✓ Nessuno ha ammesso che ***pro*** ha bevuto tutto il vino.
c.✓ Nessuno ha ammesso che ***pro*** ha detto che ***pro*** ha bevuto tutto il vino.
(123)a.✓ Gianni ha ammesso che **lui** ha bevuto tutto il vino.
b. * Nessuno ha ammesso che **lui** ha bevuto tutto il vino.
c.✓ Nessuno ha ammesso che ***pro*** ha detto che **lui** ha bevuto tutto il vino.
Nobody has admitted (that he has said that) he has drunk all the wine

and concludes that an overt pronoun cannot be directly Q(uantifier)-bound: an intermediate *pro* is necessary. Unfortunately, when focus is taken into account, as it must be given the preceding paradigms, judgments become very delicate: every example in (123) varies from inacceptable to fully acceptable, depending on the amount of focus on *lui*. To obviate this difficulty, it suffices to observe pairs of overt/ null examples, which do not involve focus. This is the case of coordination for instance (to be compared with (122)):

(124) a.✓ Gianni ha ammesso che **lui** e i suoi amici hanno bevuto tutto il vino.
b. ? Nessuno ha ammesso che **lui** e i suoi amici hanno bevuto tutto il vino.
c.✓ Nessuno ha ammesso che ***pro*** ha detto che **lui** e i suoi amici hanno bevuto tutto il vino.
Nobody has admitted (that he has said that) he and his friends have drunk all the wine

The same contrast obtains, although significantly weaker (judgments are somewhat unclear). Now the same argument as above holds: non-pro-drop languages which have two distinct realised pronouns, have the same (weak) contrast. The French paradigms (125a-b) reproduce the Italian (122)-(124), and the same holds of Germanic dialects with two forms for nominative pronouns, as St-Galler Swiss German for instance (M. Schoenenberger, p.c.) :

(125)a.✓	Jean a admis qu'	**il**	a fini la bouteille.
✓	Personne a admis qu'	**il**	a fini la bouteille.
✓	Personne a admis qu'**il** a dit qu	**il**	a fini la bouteille.
b.✓	Jean a admis que	**lui** et ses amis	ont fini la bouteille.
?	Personne a admis que	**lui** et ses amis	ont fini la bouteille.
✓	Personne a admis qu'**il** a dit que	**lui** et ses amis	ont fini la bouteille.

Nobody has admitted (that he has said that) he (and his friends) has/have finished the bottle

Since one and the same paradigm obtains in pro-drop (here Italian, but the same holds of Slovak) and non-pro-drop (here French and St-Galler German) contexts, the Overt Pronoun Constraint should not refer to the overt/non-overt distinction, but to the strong/deficient distinction, i.e. it should become the Strong Pronoun Constraint.

6.3. Up to Crash

6.3.1. Stronger (i.e. bigger) pronouns are possible only where smaller ones are impossible. But what renders the smaller one impossible? Intuitively, a bigger pronoun is possible only if generating a smaller pronoun in its place yields an impossible derivation, i.e. "crashes". *Minimise Structure* thus means "minimise up to crash".

Generating a deficient pronoun instead of a strong pronoun does not result in an acceptable derivation in coordination (because the deficient pronoun would not be in a local relation to Agr° at S-structure), with c-modification, with dislocation (considering the latter to be base-generation, Cinque (1990)), etc., i.e. exactly in the cases in which strong pronouns are allowed.

6.3.2. Adding a clause to Economy of Representations, (117), in order to incorporate its "up to crash" nature ("minimise structure, unless it leads to ungrammaticality") would not be an optimal solution: all other "economy" principles (Pesetsky (1989), Chomsky (1991, 1992)) would repeat the same clause.

Take chain-formation: a "longer (bigger) chain" is impossible where a "shorter one" is possible, and the longer is possible only if the shorter would lead to a crash (Rizzi (1990), Chomsky (1992)). The "up to crash" property is inherent to all principles of the "economy" type, among which Minimise Structure (typically resulting in informal modal renderings such as "if you *can* x, you *must* x", or "do the smallest *possible* x").

Ideally, the "up to crash" nature of all these principles should be stated only once, and not repeated in each principle.

6.3.3. Since all "economy" principles are of the "minimise α" format, where α = structure, chains, links, overt movement, etc.", an elegant solution is thus to postulate a unique general principle covering all economy-type constraints:

(126) Minimise α

of which Economy of Representations, Derivation , etc. are only special cases. This general principle must now be understood as incorporating the "up to crash" clause (*minimise α up to crash*). Of course, to limit the scope of α is then non-trivial. [70]

6.4. Level of Application

The preceding discussion presupposes that Economy of Derivations applies at the point where the pronouns are generated, i.e. at tails of chains. It need not be stipulated however that (core) lexical insertion is the place where Economy of Derivations applies. Optimally, Minimise α applies everywhere. It just so happens that Minimise α as Minimise Structure can apply at lexical insertion, while other instantiations of Minimise α

---

[70] By definition, all principles of the economy type, including Minimise Structure (and its precedessor Avoid Pronoun), are transderivational, albeit in a limited way (sometimes so limited as to be trivial): by definition, these principles allow a derivation *only if the "next more economical" is not possible*. In more intuitive terms: "how do you know whether to further minimise alpha, or stop there?", only by knowing that further minimalisation will trigger ungrammaticality. This is inherently transderivational.

Phrasing derivation in filtering terms (i.e. allowing derivations to crash) does not resolve this problem: the outcome of some derivations (crash or not) is determined by reference to the outcome of another derivation. (We thank Riny Huybregts for concentrating our attention on this issue).

cannot (such as Minimise chains), for purely independent reasons (i.e. there are not chains at core lexical insertion).

This also answers another curious observation: apparently Minimise α *qua* structure is contradictory with Minimise α *qua* chains: minimising structure leads to bigger (overt) chains. From this it could be inferred that there is a ranking among economy principles, to ensure that one takes precedence over the other in case of contradictory outcome. But no such thing is necessary in the present case: due to independent reasons, Minimise α *qua* structure applies prior to Minimise α *qua* chains.

7. Refinements

7.1. Results and Problems

7.1.1. Each of the asymmetries between clitic, weak and strong pronouns noted in (75) has now been derived, and this through a theory which meets the *a priory* standards set in the *preambule* to the second part.

The unique primitive is structural deficiency: lacking (the features of) the highest functional projection. Structural Deficiency is either directly responsible of the asymmetries (as with morphology, range, prosody, modification), or indirectly through its interplay with two conditions: Economy of Representation (deriving the choice-generalisation) and Recoverability (deriving displacement of a minimised element).

7.1.2. Although this may seem like a minimal account of the rich set of asymmetries (75) several background assumptions are built into the system. For instance:

(i) recoverability applies to deficient pronouns. Intuitive as it may be, this is an odd fact as things stand: nothing requires recoverability to apply at all. Why cannot a weak or clitic pronoun be base generated with nothing to recover since nothing is *lost* in the first place?

(ii) deficient pronouns must recover features that are lacking *with respect to a corresponding strong pronoun*. Again, as it stands, the Economy of Representations does not express this primacy of strong pronouns (just as *minimise* α does not encode any putative primacy of non-minimal forms). Why are not weak pronouns, or some arbitrary dummy noun, the reference point, for instance? It surely would be a logical possibility. [71]

(iii) nominal (extended) projections have the same structure (and labels) as verbal (extended) projections.

These assumptions need not be primitives. A particular implementation of Minimise Structure derives them all, significantly simplifying the overall profile of the theory (Section 8.3.).

7.1.3. Furthermore, as things stand at least one major misprediction occurs: if the possibility of deficient pronouns always blocks the use of strong pronouns (*Minimise Structure*), how can strong pronouns be used to refer to new discourse elements?

7.2. Indexes and Reference Sets

Superficially, the latter problem is trivial: deficient pronouns cannot bear an index, occasions where an index is needed will thus force a strong pronoun to be present, the correct result:

(127)a.✓ Jean a vu **lui**$_{ndr}$.
b. Jean **l'**a vu **t** ✓$_{odr}$ / *$_{ndr}$.
Jean him has seen him

The trouble is to give a coherent interpretation to "deficient pronouns cannot bear an index".

One possibility would be that indexes occupy C° and thus require the occurrence of a strong pronoun in order to appear (§5.4.2). This was rejected in §5.4.7 in view of the asymmetry between case and index: lack of case in C° always triggers Recoverability effects (i.e. displacement), so that lack of putative-index in C°

---

[71] The primacy of strong pronouns is also shown by acquisition data. In Cardinaletti & Starke's (1994c) interpretation, asymmetries linked to acquisition of principle B of binding theory (Jakubowicz (1984), Wexler & Chien (1985)) reduce to the fact that, confronted to a pair of <deficient; strong> ambiguous pronouns, a child always resolves ambiguity by postulating only one form, the strong one.

should have the same effect. This is however not true (cf. expletives, impersonals, non-referential datives). Unless the recoverability requirement can be made selectively blind to the index, this asymmetry indicates that the index is not in lost in the first place.

It was thus concluded that the index is absent from syntax altogether: all falls into place if indexes are the interpretation of the "functional case-features" by post-LF interpretive systems. To explain (127) it is now necessary that a post-syntactic process (index assignment) influences an intra-syntactic process (non-application of Minimise Structure in (127a)). If we follow Chomsky (1994), as we did in assuming the inexistence of indexes in syntax, this is a typical situation: post-PF filters on possible words for instance force larger displacement than would have otherwise been required.

But the external filter responsible for (127) is particularly delicate: contrary to the usual case in which such filters systematically rule out one possibility, allowing the other, both (127a) and (127b) have a well formed interpretation. What seems to be the case, is that the nature of the index contributes to the definition of the reference-set (the set determining which derivations are to be compared): compared LF-representations not only should share their building blocks (Deep Structure, Enumeration) but their interpretation must also be identical (the interpretation of two pronouns cannot be identical if one has a new index, "read off" from functional case-feature, while the other inherits it from an antecedent). Thus given "object = ndr" in (127), the strong pronoun is the minimal possible element, while the reverse obtains if the object is "odr".

Ultimately, Minimise α should integrate this requirement and become: Minimise α, up to crash, *given a particular choice of interpretation.*

7.3. Implementing Economy of Representations and Recoverability

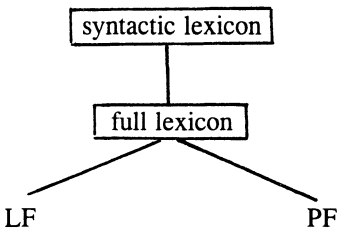
7.3.1. The puzzles of recoverability raised in §8.1.2. (why does it apply at all, why is there a primacy of strong pronouns) are answered at once by a particular implementation of Minimise Structure: only strong elements are ever generated in base.

7.3.2. As a consequence, *minimise* α, where α = structure (or features), can only be *erase* α: if only full (strong) structures are generable, deficient elements can only be obtained through deletion. [72]

Strong elements define which features there are to be recovered *because* deficient elements are obtained by erasing part of strong elements; recoverability applies to Economy of Representations *because* the latter involves actual deletion.

7.3.3. Deficient pronouns being a result of a syntactic process, it is a necessary consequence of this model that the morphophonological form of (these) lexical items is accessed only after (some) syntactic derivation. Access to the (morphophonological) lexicon must therefore take place after (some) syntactic derivation (cf. also den Besten (1976), Otero (1976), Halle & Marantz (1993), Jackendoff (1994)). In turn, this implies the existence of a presyntactic lexicon, providing the necessary features for syntactic derivations:

(128)



---

[72] An erasing-implementation was first proposed to us by Dominique Sportiche (p.c.). The particular use we have made of it is however not to be blamed upon him.

Such an implementation opposes the one in which structures are simply built in parallel and then compared w.r.t. economy. This conclusion that "starting from the most uneconomical and stripping down" is to be preferred over "parallel generation and choice" is not automatically transposable to other instances of minimise α: in the case of minimise structure, such a course of events is forced by the properties of the syntactic lexicon (cf. below). Nothing seems to force this in the case of chains, to the opposite.

Finally, such an implementation presupposes the (standard) view that "there cannot be holes in structure", i.e. *erase* α can only erase the *highest* functional projection, and not some intermediate layer.

7.3.4. Post-lexical insertion is independently needed: only through such an account can the existence of functional fused forms be explained. Forms such as *du* or *des* (of.the) in French and many other cases are portemanteau morphemes realising two distinct functional heads.

First, such forms cannot be base-generated since doing so would require generating features in the wrong head, i.e. phi-features under C° or referential and case features under I°. But this is highly unsatisfactory, and incoherent with the very idea of distinct functional projections, since the latter are distinct only to the extent that they encode distinct features. [73]

Second, such a base generation account is also empirically inadequate: the occurrence of some fused forms, such as French *du*, is dependent upon syntactic configuration. If a high nominal modifier occurs, the two heads remain distinct ... *de tout le monde*... "of all the people" versus ... *du beau monde* ... "of.the nice people". The choice between the morphophonological *de le* and *du* cannot be performed prior to syntax, or forms such as *... *de le monde*... (and maybe *... *du tout monde*...) would not be filtered out.

7.3.5. What are the properties of such a "split lexicon" ?

First, given that functional heads are nothing but reduplications of features contained in lexical heads, it would be redundant to postulate them in the syntactic lexicon. All information they contain is already contained in the lexical element, and the syntactic tree may be constructed on the sole basis of the latter.

Second, it is not the case that the full lexicon contains morphophonological information and that the syntactic lexicon is simply its syntactic counterpart: several entries exist in the full but not in the syntactic lexicon. Minimally, deficient pronouns and fused forms exist only in the full lexicon. More generally, if the first point above is correct, all function words exist only in the full lexicon.

The (pre)syntactic lexicon is thus a type of "abstract" or "core" lexicon, containing only grammatical features for a subset of lexemes, sometimes designed as a "functional lexicon" (an adequate though slightly misleading term in the present approach). The type of derivation intended here is that, first, (abstract) lexical items are selected from the syntactic lexicon, (cf. fn. 74 on abstractness), the features of these are then projected onto a set of functional projections, each reduplicating features of the lexical element, in a one-step operation (somewhat in the spirit of (a restricted view of) *Elementary Trees* of Frank & Kroch (1993), cf. also Kroch & Joshi (1985)). Those full phrases (extended projections) then combine to form the (deep structure) syntactic representation. Only after (some) syntactic derivation is the full lexicon accessed.

7.3.6. The fact that nominal phrases (C_nP) and verbal clauses (C_vP) are associated to identical functional projections, as well as the fact that only strong pronouns may be generated in the base, now comes down to a restriction on the syntactic lexicon. Given that functional heads reduplicate features of lexical heads, the identity of strong pronouns, noun-phrases and verbal clauses implies that all three realise an identical array/ set of (underspecified) features (this also holds of other categories, cf. §9). This may now be derived from a property of the pre-derivational lexicon: all entries of the syntactic lexicon realise one, and only one, array/ set of features, $\alpha_1 \ldots \alpha_n$,. These features then project onto what invariably becomes the {verbal, nominal, adjectival, etc.} complementiser phrases, agreement phrases, etc. (cf. also Starke (1993a), (to appear)). [74]

(129) entries of the syntactic lexicon all realise a fixed array of (underspecified) features, $\alpha_1 \ldots \alpha_n$,

---

[73] These arguments do not apply to cases in which one of the heads is lexical, since by definition all functional features also occur in the lexical head. The alternation between *does not sing* and *sings* is thus immune to the above discussion (unless *do* is taken to realise both phi and tense features simultaneously, in which case the same problem arises again).

[74] A posteriori, it is unsurprising that such a strong, universal condition holds of the core lexicon. Many of its properties are largely universal: classes such as ergative/transitive, noun/pronoun, psych verbs/ perception verbs etc., i.e. most syntactically relevant classes (which must therefore feed syntax), are attested in widely different and unrelated languages, and are thus optimally attributed to a fix core of language.

Both this observation and the text-conclusion that only one format is available for all entries of the syntactic lexicon suggest that the latter is a highly constrained, strongly UG-driven lexicon. An intriguing possibility is that this lexicon contains entries only for types (i.e. classes) of lexical elements (perception verbs, ergative adjectives, pronouns, etc.) but not for each individual lexical item. In such a case, entries of this lexicon are learned only to the limited extent of "parameter setting", i.e. fixing the value of features associated to word-classes. The fact that in some language verbs and nouns, say, have different surface orders w.r.t. their arguments (SOV vs. SNO) may thus reduces to the distinct feature-content of the abstract verb and noun in the syntactic lexicon.

The primacy of strong pronouns (§8.1.2), and the common format of verbal and nominal phrases (§5.2.3) are reflexes of this strong restriction on the entries of the syntactic lexicon.

7.3.7. If, given (128)-(129), *minimise* α is instantiated as *erase* α, what is the status of Recoverability? Supposing that the latter should not be integrated into *minimise* α but is an independent principle, a radical though natural reformulation of the Projection Principle yields exactly the correct result:

(130) *Projection Principle*
All information contained at level R must be present at level R+n.

7.3.8. The account proposed in §4-6 can now be significantly simplified, through the interplay of three independent constraints (i) the rigidity of the syntactic lexicon, (129), which is only capable of generating what corresponds to full CPs, (ii) the generalised economy principle, (126), understood as integrating an "up to crash" clause, and (iii) the projection principle, (130), forcing to recover erased features.

Full pronouns are always inserted in what corresponds to tails of chains. On the one hand the interplay of *minimise* α and of the Projection Principle determine which type of pronoun will be chosen: by *minimise* α pronouns are reduced as much as possible ("up to crash"), and by the Projection Principle erasing of unrecoverable feature leads to ungrammaticality. This entails that pronouns referring to non-prominent discourse entities, coordinated, dislocated, isolated pronouns, etc. will always be strong, and that in all other cases, the deficient counterpart will be obligatorily used. [75]

On the other hand, once the choice is made, if a deficient form is produced, the Projection Principle forces a peculiar derivation: recovery of the erased features implies that the deficient pronoun must be displaced at S-structure to establish a local relation with the relevant head, to respect the Projection Principle both at PF and LF. [76]

7.3.9. The initial question, i.e. what makes the class 1 / class 2 distinction (i.e. deficient vs. strong pronouns) so radically different from usual lexical classes (verbs/adjectives, transitives/ergatives, etc.) now receives a simple answer. While lexical classes arise from difference in feature composition of the lexical item (possibly in the core syntactic lexicon), class 1 / class 2 distinctions arise through syntactic processes (deletion). Class 2 (deficient elements) is a structural subset of class 1 (strong elements). This entails, among others, that the nature of lexical classes may slightly vary from language to language, but that the deficient/strong distinction is uniform, being due to an abstract grammatical process, *minimise* α , universal by hypothesis.

8. Beyond Pronouns

The preceding remarks focus on a narrow class of grammatical elements, personal pronouns. Ideally, this should not be so: the same tripartition (strong, weak, and clitic), revolving around the same set of asymmetries, obtains outside the realm of personal pronouns.

Unfortunately, while intensive work has been concentrated on pronominal clitics in the last two decades, scarcely any material is available on the differing properties of clitic, weak and strong adjectives, nouns, etc. Compensating for this asymmetry here would have been impossible without both exceeding space-limits of an article and provoking an unwelcome disbalance between description and analysis.

The approach developed above however naturally extends beyond the realm of personal pronouns to derive all major (known) aspects of non-pronominal deficient categories. In what follows, such an extention is illustrated through a few chosen pieces from adverb-syntax, a preliminary sample of a forthcoming more thorough treatment (cf. also (9), (61)).

8.1. An example: Description of Deficient Adverbs

8.1.1. A subset of Greek adverbs transparently illustrate the interplay between morphology, distribution and sensitivity to constructions such as coordination (cf. Rivero (1992), Alexiadou (1994), a.o.).

---

[75] Unless, of course, no such a form exists in the full lexicon of the language. It is worthwhile noting, however, that we have never encountered a language which does not have two series of pronouns, possibly a non-existent state of affairs.

[76] The asymmetry between case and index (§5.4.7) entails that case but not index is part of the fixed array of features constitutive of entries of the syntactic lexicon.

(131) a. To ***sigá** évrasa **✓sigá**.
b. To **✓sigo-** évrasa ***sigo**.
it slowly I.boiled slowly

The unaccented form *sigo* necessarily occurs in high derived position to the left of the verbal element in (131) while the full accented form *sigá* has no such restriction, and appears in the usual postverbal position. As with personal pronouns, the deficient form cannot be coordinated, or c-modified.

(132) a. ✓ To évrasa **sigá** ke **kalà**.
b. * To **sigo** ke **kalo-** évrasa.
it slowly and well-I.boiled

The triple correlation between morphology, necessary displacement and sensitivity to coordination and c-modification is exactly identical to that found amidst personal pronouns. [77]

8.1.2. Similar paradigms holds in Romance (cf. Starke (1994), see also Lonzi (1991)). The French *bien*, for instance, obligatorily appears before the past participle in compound tenses (all judgments hold for flat intonation):

(133) a. ✓ Il a **bien** essuyé la vaisselle.
b. * Il a essuyé la vaisselle **bien**.
he has well dryed the dishes well

unless it is coordinated, c-modified or focussed:

(134) a. ✓ Il a essuyé la vaisselle **bien** et rapidement.
b. ✓ Il a essuyé la vaisselle vraiment **bien**.
c. ✓ Il a essuyé la vaisselle **BIEN**, pas longuement.
he has dryed the dishes well and quickly / really well / well but not at-length.

This is a typical Economy of Representation paradigm: if the deficient (pre-participial) element can be used, it must, in formal identity to:

(135) a. ✓ Il **nous** a vu.
b. * Il a vu **nous**.
c. ✓ Il a vu **nous** et nos amis / que **nous** / **NOUS**, pas les autres.
he us seen us and our friends/ only us / US, not the others

To *explain* these paradigms, and to also capture the parallelism with personal pronouns, the postulation of a deficient *bien* in French is necessary. The same paradigm holds of other adverbs, both within French and across Romance.

8.1.3. But adverbs do not merely classify into deficient and strong. Both cross-linguistic and language internal differences require two types of deficient adverbs: maximal projections (weak adverbs) and heads (clitic adverbs).

Exactly as the clearest case of X° pronouns was provided by Italian I-to-C contexts, in which the pronoun is displaced over the subject along with the verb, resulting in a *[pron.-V]$_i$ Subject t_i ...* configuration, the clearest case of X° adverbs is given by Old Rumanian I-to-C constructions in which the adverb is displaced along with the verb over the subject, resulting in an *[adv-V]$_i$ Subject t_i ...* configuration (C. Dobrovie-Sorin (p.c.); for discussion of Rumanian clitic adverbs, cf. Motapanyane (1991), Dobrovie-Sorin (1992) a.o.). Greek examples such as (131b) provide another argument: given that *to* 'it' is analysed as an ad-verbal clitic incorporated into the (head hosting the) verb, the intervention of the adverb between the two entails the clitic status of the adverb (in compliance with the general observation that only clitics may break the adjacency between ad-verbal clitics and the verb).

On the other hand, the French *bien* neither attaches to the verb (contrary to Rumanian deficient adverbs), nor blocks verb-displacement (contrary to English negation). It is a weak adverb in a specifier position: [78]

---

[77] It is irrelevant if the morphological difference between the two forms may be phonologically defined: what matters, is that this difference strictly correlates with displacement, capacity to enter coordination, and other typical properties of deficient elements, as semantic deficiency (see below).

(136) ✓ Pierre cuisine$_i$ **bien** t_i
Pierre cooks well

The contrast between (133b) and (136) reduces to difference in French verb-placement in tensed and untensed verbs (Pollock (1989)). To give a more minimal pair, along with its underlying configuration: [79]

(137)a. Pierre [cuisine$_i$ ✓**bien** t_i
b. Pierre [a ✓**bien** cuisiné ***bien** / ✓**bien mais peu**
Pierre has well cooked well / well but little

8.1.4. Not only do adverbial asymmetries mimick pronominal contrasts in morphology, distribution, choice and sensitivity to coordination/ c-modification, but semantic and prosodic properties are also reproduced. Just as deficient but not strong pronouns may be rangeless (loosely "non-referential"), deficient but not strong adverbs may be "non-referential": in all above examples of deficient *bien*, it is ambiguous between a literal reading (the manner adverb corresponding to English *well*) and a discourse-particle meaning "certainly/ indeed, Pierre cooks" (cf. also Belletti (1990)). This is most clear in weak-climbing contexts:

(138) Il a **bien** du [parler t
he has well "must" speak = "he certainly/ indeed/ etc. has been obliged to speak"

This reading is never available with strong adverbs (i.e. post-participial *bien*, in coordination, etc.)

8.1.5. Finally, weak but not strong adverb trigger sandhi rules such as liaison. Whereas it is true of all adverbs that no liaison obtains before an adjunct (judgments hold of spoken Geneva French):

(139) a. * Je m'entends bien avec Marie.
I me hear well with Mary ("we get along well")
b. * Il aboie beaucoup après le dinner.
he barks lot after dinner
c. * Il avance lentement en ville.
he advances slowly in city
d. * On avance jamais avant une conférence.
one progresses never before a conference

a subset of adverbs, exactly those which qualify as weak, undergo liaison in front of past-participles:

(140) a.✓ Je me suis bien entendu avec Marie.
I me heared well with Mary ("we get along well")
b.✓?Il a beaucoup aboyé après le dinner.
he has a.lot barked after dinner
c.*? Il a lentement avancé.
he has slowly advanced

---

[78] That French adverbs belong to the class of weak elements and not clitic elements is further confirmed by their undergoing weak-climbing in modal constructions, a phenomenon found in French with weak elements such as *tout* 'everything', [ia], but never with clitics, [ib] (Kayne (1975)):
[i] a. Il a **tout** du [faire t ('he has everything must do' = he has been obliged to do everything)
Il a **bien** du [se comporter t ('he has well must to.behave' = he must have behaved well)
b. * Il **l**'a du [faire t ('he it has must do')
Finally, it would not be convincing to argue that French deficient adverbs are heads (clitics) but do not interfere with verb-displacement because they do not contain features related to the verb, which is thus free to skip them. The same effect is found with aspectual adverbs, the features of which are clearly linked to the verbal-features.

[79] This example further shows that motivation for deficient-adverb placement cannot be prosodic in a simple sense: from the preceding text-examples it could have been inferred that (deficient) adverbs can simply not be clause-final, maybe for prosodic reasons. This is however not the case: all versions of *bien* , for instance, may perfectly be clause-final, whenever verb-displacement (Pollock (1989)) has taken place, as in (136).

d.*? On a jamais avancé le travail avant une conférence.
one progresses never the work before a conference

The conditions on adverb-liaison may now be significantly simplified. Apart from usual locality conditions and segmental prerequisites, the list of adverbs which undergo liaison now reduces to a simple statement: weak but not strong adverbs undergo liaison. In the absence of such a distinction, a list of adverbs would have to be stipulated.

8.1.6. In sum, all of morphological, distributional, prosodic, choice, and construction-sensitive asymmetries are found with adverbs, exactly as with personal pronouns: [80]

(141) Deficient adverbs, contrary to strong adverbs
a. must occur in a derived position at S-structure
b. cannot enter coordination, c-modification
c. are preferably chosen
d. may be morphologically reduced w.r.t. the other series
e. may be "non-referential"
f. undergo prosodic restructuring

8.2. An example of Derivation: Deficient Adverbs

8.2.1. The derivation of the differing properties of the two classes of adverbs by and large mirrors that of the different classes of pronouns: lack of the highest functional layer, CP, triggers mild deficiency, and the additional absence of a second layer, SigmaP, triggers severe deficiency, i.e. clitichood.

8.2.2. First, that the high functional layers of adverbs are similar to that of nouns and pronouns, is suggested by the Senigallia dialect discussed above: not only does the dummy morpheme *ma* appear on strong noun-phrases and strong pronouns, (142), but also on strong adverbs, (143) (examples from Sellani (1988)): [81]

(142) a. ✓ tutt l' ser arconta **ma**$_{DAT}$ i fiulini (p. 9)
all the evenings she. tells-tales *ma* the children

b. ✓ e po' s' sent urlà **ma**$_{ACC}$ i venditori (p. 39)
and then SI hears shout *ma* the venders
"and then one hears the venders shout"

c. ✓ Ho vist **ma**lu
I.have seen ma.him

(143) a. ✓ so v.nuta **ma**quà (p. 9) (cf. the Italian counterpart: qua)
I.am came here

b. ✓ che s' magn.n anch'ogg **ma**lagiù (p. 9) (cf. the Italian counterpart: laggiù)
which SI eat also today there

c. ✓ **ma**lì dietra l' Cumun (p. 34) (cf. the Italian counterpart: lì)
there behind the town hall

8.2.3. Exactly as with personal pronouns, the full lexicon may contain both strong and reduced adverbs, but only full (strong) forms are generable by the syntactic lexicon. By *minimise* α, these full forms are then reduced as much as possible.

This already entails (i) the morphological asymmetry (deficient adverbs are reduced), (ii) the prosodic asymmetry (deficient adverbs do not count as major constituents), (iii) the choice asymmetry (having less structure, the more deficient version is preferred) (iv) the c-modification asymmetry (having no C° they cannot be modified by CP-modifiers).

Furthermore, since the highest layer (C°), when present, necessarily contains *some* feature (or it would not be projected), its absence entails the absence of some feature. To compensate for this absence, the

---

[80] As with pronouns, the simultaneous existence of semantic and prosodic correlates to the deficient/ strong distinction indicates that the primitive underlying the class-distinction is not restricted to one or the other component, i.e. not purely prosodic, nor purely semantic.

[81] A similar point holds of adjectives: in a number of languages a dummy morpheme appears on adjectival phrases which exactly parallels the dummy morpheme appearing on noun-phrases.

deficient adverb must occur in the checking position of this feature at S-structure (and similarly for clitic adverbs with the additional incorporation, exactly as with pronouns). [82]

Finally, since the syntactic lexicon is limited to one array of features, if referential features, α_i, occur in C° in one type of phrase, they occur in C° in all phrases. It follows that only deficient adverbs lack referential features and may act as "non-referential discourse-particles". [83]

Obviously, many details are left untouched here, and the existence of deficient adverbs would force a modification of several points of the preceding account, but no principled problem seems to arise. The theory developed on the basis of personal pronouns carries over *in toto* to deficient adverbs, dealing with all major asymmetries described above. By and large the same holds of tripartitions of adjectives, nouns, quantifiers, wh-words, etc., the topic of forthcoming work.

9. Summary and Conclusion

9.1. Prerequisites

9.1.1. The central thrust of the present proposal is that an adequate theory of clitic pronouns, i.e. oppositions between clitic and strong pronouns, such as:

(144)a. Gianni **la** vede con piacere.
b. Gianni vede **lei** con piacere.
Gianni her sees her with pleasure

should be a theory of a considerably enlarged paradigm.

---

[82] The existence of lexical clitics, as the above adverbs, but also nouns, adjectives, etc. provides a strong argument against an alternative approach to deficient-placement proposed in Sportiche (1992). Sportiche suggests that the only analysis of clitics open to the learner is that in which the clitic realises a functional head (his §1.1), and provokes displacement of a silent argument into its specifier, due (roughly) to a clitic-criterion (his §1.3) (but cf. his fn. 27).

Such an analysis entails that all clitics are functional, in contradiction with the existence of lexical clitics, if a meaningful generalisation englobes pronominal and lexical deficient clitics. A criterion-based analysis is further open to several less important objections, to which a deficiency analysis is immune, among which (it should be noted that irrespectively of these problems with clitics, Sportiche's approach can be made compatible with everything we have presented on weak pronouns):
(i) the distinction between ad-verbal clitics and C2 clitics is not readily explained by criteria: C2 clitics are apparently much higher than the ad-verbal clitics, although there is no trigger to clitic-displacement beyond the base-generation site, in a criteria-approach. The sole solution would seem to be that C2 clitics are NOT much higher, i.e. that both types of clitics are generated in the same place, and remain there, modulo verb-movement. But such an analysis would entail a range of severe difficulties w.r.t the respective clause-structures of Slavic and Romance, w.r.t. verb-movement, and w.r.t. the general parallelism between V2 and C2 structures (which would be mostly obliterated).
(ii) the reduced auxiliaries of Slavic, illustrated in §6.2 for Serbo-Croatian, distributionally and prosodically pattern with pronominal clitics. Accordingly, it is traditionally assumed that a strong generalisation relates the two. To the extent that this is correct, it is a counterexample to the criterion approach, there being no plausible XP counterpart to auxiliaries (in the present framework this is not a strong objection: since verbal clitics do not provoke deficiency of their whole CP, they cannot be clitics in the present sense of the term, but rather auxiliaries which move very high for independent reasons).
(iii) a criterion approach would be at pains to explain the regular precedence of clitics over weak and weak over strong, given that it considers clitic and weak as two fundamentally distinct elements. To explain the clitic-over-weak precedence, it would presumably require a principle akin to Avoid Pronoun, but the preference of weak over strong (as in the English particle-constructions or the German adverb-pronoun combinations, cf fn 35) is then left unexplained.
(iv) it was argued above, fn. 19, that explanation of the semantic contrasts between the deficient and strong pronouns requires the *absence* of a property, and not the *presence* of a special feature, a state of affairs contradictory with the criterion approach.

[83] This system makes the prediction that while weak elements of diverse categories (i.e. weak adverbs, weak pronouns) occur in distinct positions, having distinct (maybe θ) features to "check", all clitics occur in the same position within a given construction of a given language. The rare available observations confirm this prediction, but again, this remains an open issue until more data is gathered.

9.1.2. It should be a theory of tripartitions (not bipartitions) of clitic, weak and strong elements, tripartition into which pronominal systems consistently divide, across languages.

(145)a.	Non	**gli**	dirò mai	**gli*	tutto	**gli.*
b.	Non	**loro*	dirò mai	**loro**	tutto	**loro.*
c.	Non	**a lui*	dirò mai	**a lui*	tutto	**a lui.**
	no	to.him/to.them	I.will.say never		everything	

Cross-linguistically, each class shares the same properties, which oppose it to both other classes, with a regularity that indicates the presence of three abstract underlying classes, rather than idiosyncratic lexical accidents.

Most notably, both weak and strong elements cross-linguistically occupy XP positions at S-structure, contrary to clitics found only in X° positions, while, on the other hand, clitic and weak are both *deficient*, i.e. restricted with respect to a large set of constructions, among which coordination (neither of them is coordinable, while strong elements are).

These two properties further illustrate the intermediate status of weak elements (identical to strong w.r.t. x-bar, but like clitics w.r.t. coordination (deficiency)), resulting in a typical *clitic* < *weak* < *strong* relationship across the three classes. This is most strikingly illustrated by the fact that all properties differentiating weak elements from strong elements also differentiate clitic elements from strong elements. Deficient characteristics of weak are a *proper subset* of deficient characteristics of clitics, again *clitic* < *weak* < *strong*.

9.1.3. A theory of clitic pronouns should also handle morphological, as well as distributional, semantic, prosodic, and phonological contrasts. The rich net of asymmetries distinguishing the three classes, cuts across all these components: morphology (clitic ≤ weak ≤ strong), distribution (clitic and weak pronoun must be in a derived position, contrary to strong; clitics are heads at S-structure, contrary to weak and strong, etc.), semantics (clitic and weak lack range, strong always have one), prosody (clitic and weak restructure prosodically, contrary to strong; weak and strong may have word-accent, contrary to clitics) and phonology (liaison and contraction rules are restricted to clitic and weak elements).

Surprisingly, while these asymmetries seem to be absolute universal, none of the interpretive asymmetry is systematic: it is not the case that there is a strict covariation between being of one class, and having one type of semantic/ phonetic interpretation. The interpretational characteristics are asymmetric but overlapping: the three classes are purely abstract (both deficient and strong elements can for instance refer to human entities and to prominent discourse referents, although an asymmetry holds w.r.t. non-human entities and non-prominent referents).

9.1.4. Finally, a theory of clitic elements should be applicable across lexical classes: just as personal pronouns may be either clitic, weak or strong, all of adverbs, adjectives, quantifiers, wh-pronouns, nouns, etc. are found in all three format. Furthermore, the characteristics of clitic, weak and strong elements are largely *identical across categories*. A clitic pronoun differs from a strong pronoun *in the same way* as a clitic adverb differs from a strong adverb.

9.2. Summary

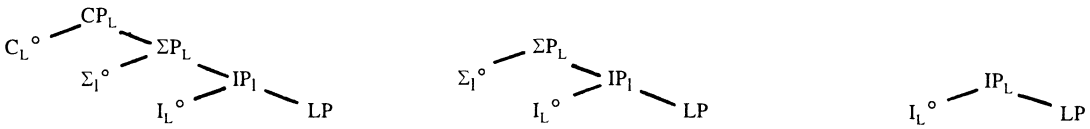
9.2.1. The morphological asymmetry between the three classes (clitic ≤ weak ≤ strong), together with the *Principle and Parameters* framework as it stands, indicates a simple analyis complying with all the above prerequisites. Since deficient elements are systematically morphologically reduced w.r.t. the strong elements, and since morphemes are syntactic terminals, deficient elements realise less syntactic structure than strong elements. This is particularly clear in transparent morphology, where one class is a morphological subset of the other.

9.2.2. Based on surface morphological forms (which are taken as indicators of the underlying trigger and not as actual triggers) the missing structure is systematically identified as a high functional morpheme: while strong pronouns are full nominal projections, weak pronouns lack the highest functional layer, and clitic pronouns further lack both of the two highest functional layers.

The uniformity of these layers across classes has then led to the hypothesis that there is one and only one format for all syntactic structure, across languages, constructions and lexical items. Deviation from this basic format, an extremely rare fact, leads to *deficiency*, triggering strong consequences for the deficient element.

Based on the nature of the interpretive properties involved, and on the most widely accepted labelling, that attributed to verbal (extended) projections, the labels adopted are (where IP is a cover term for a set of functional projections, and L refers to any lexical category):

(146)a. Strong Pronouns b. Weak Pronouns c. Clitic Pronouns



From this, most aspects of deficiency directly follow: morphological reduction is a direct reflex of lack of structure, impossibility of modification follows from the observation that those modifiers that can modify strong but not deficient elements only modify full CPs, the choice preference follows from the diverse structures in combination with a general principle *minimise* α , the prosodic asymmetries comply with the observation that "major constituents" (i.e. CP) are treated differently from non-major constituents w.r.t. prosodic processes. Other distributional and semantic asymmetries follow not from the sheer absence of structure, but from the absence of features contained in those structures: case features in C° (and consequently, referential information) and polarity and prosodic features in Σ°.

9.2.3. An attentive observation of the choice preference shows that strong elements are treated as *prior* to deficient elements: a deficient element must be chosen but only if it is associated to the same features as *those which would have been contained in the strong counterpart.*

This primacy, together with the whole general theory of deficiency, may be implemented through three general assumptions, two of which are hardly more than expressions of what is generally assumed:

(147)a. minimise α (Economy Principle)
b. information of level R must be present at R+n. (Projection Principle)
c. all entries of the syntactic lexicon are limited to one array of features, $\alpha_1 \ldots \alpha_n$

The identity of all (extended) projections now follows from (147c): features $\alpha_1 \ldots \alpha_n$ invariably project onto what becomes CP, IP, etc. It also follows that only strong pronouns are generable. The (generalised) economy principle then forces to reduce structure as much as possible, α being in this case structure (strictly speaking, *minimise* α is forced to operate as *erase* α in this case), thus deriving both the choice asymmetry and the primacy of strong elements.

Finally, the projection principle forces recovery of features erased by *minimise* α. This recovery is possible only through a local relation between the deficient element and an adequate head at S-structure (assuming there to be no displacement at PF), thus deriving the distributional asymmetries.

9.2.4. The relevant set of properties now all follow, independently from the nature of the lexical head (across verbs, nouns, adjectives, etc.), with the desired morphological, semantic, prosodic and syntactic consequences:

(148)a. From the sheer absence of highest projections in deficient elements (clitic and weak), it follows that:
- the more an element is deficient, the more it tends to be morphologically reduced
- deficient elements cannot be modified by modifiers of (elements selecting) CPs
- deficient elements are not "major constituents", a central notion in prosody
- the most deficient element possible is preferred (by *minimise* α)

b. From the absence of C-features in deficient elements it follows that:
- deficient elements never have their own range (and are thus always either expletive or coreferent)
- deficient elements must be displaced to recover missing (case-)features (Projection Principle)

c. From the absence of Σ-features in clitic elements it follows that:
- clitic elements do not have word-accent

- clitic-elements must be displaced to recover missing (prosodic) features. To not destroy the effect of the recovery of the C-feature, a X°-chain must be created. (Projection Principle)

9.3. Conclusion

The present investigation, we hope, illustrates the interplay of abstract theoretical constructs and empirical generalisations. The first part seeks to establish that what is to be explained by a theory of simple oppositions between clitic and strong pronouns, arriving at several new conclusions. Most prominently, that the relevant opposition is between three distinct classes: clitic, weak and strong; but also that these classes are separated by a regular range of semantic (referential) oppositions. The global picture then becomes uniform: clitic pronouns are deficient w.r.t weak pronouns which are in turn deficient w.r.t. strong pronouns, both distributionally, morphologically, semantically and prosodically.

This generalisation, (75), then indicates a simple abstract primitive: *structural deficiency*. Some pronouns are deficient in that they have a deficient syntactic structure. For the first time, to our knowledge, this opens a (tentative) road towards a unified derivation of the whole range of syntactic, morphological, semantic and prosodic effects involved, but also of the similar properties of pronominal, adverbial, adjectival, etc. clitic, weak and strong elements.

The postulation of diverse structures then entails a set of constraints which regulate the generation and derivation of syntactic structure, further constraining the general model of grammar upon which it is based. Precedence patterns among distinct classes of pronouns indicate that only full, i.e. strong, structures are generable. The existence of deficient structure must then be attributed to a reduction process in syntax, traced down to a general *Minimise Structure* principle, subsumed under a global economy principle *minimise* α (cf. Chomsky's (1992) economy guidelines). Finally, this entails a split lexicon, with post-syntactic access to morpho-phonological information (cf. Halle & Marantz (1993), Jackendoff (1994)).

The "theory of clitics" thus developed, is a general theory of arguments and adjuncts, and of their syntactic structure; thereby defining a novel set of central questions, which we hope to be productive avenues of research.

REFERENCES

Abney, S. (1987) "The English Noun-Phrase in its Sentential Aspects", MIT PhD dissertation.
Abraham, W. (1991) "Rektion und Abfolge pronominaler Satzglieder und ihrer klitischen Formen im Deutschen", *Language and Cognition*, 1.
Alexiadou, A. (1994) "Adverbial Licensing and Clause Structure", paper presented at FAS, Berlin.
Arnauld, A. & C. Lancelot (1846) *Grammaire Générale et Raisonnée de Port Royal*, Paris (reprint (1993) Slatkine, Geneva).
Baker, M.C. (1988) *Incorporation*, The University of Chicago Press.
Berendsen, E. (1986), *The Phonology of Cliticization*, Foris, Dordrecht.
den Besten, H. (1976) "Surface Lexicalisation and Trace Theory" in H. van Riemsdijk (ed.), *Green Ideas Blown Up*.
Bonet, E. (1994) "The Person-Case Constraint", paper presented at the Morphology-Syntax Connection Workshop, MIT.
Booj, G. (1993) "Connected Speech III: Cliticization", ms.
Brame, M. (1981) "The General Theory of Binding and Fusion", *Linguistic Analysis* 7.3:277-325
Brandi, L. & P. Cordin (1981) "Dialetti e italiano: un confronto sul parametro del soggetto nullo", *Rivista di grammatica generativa* 6, 33-87.
——————— (1989) "Two Italian dialects and the null subject parameter", in O. Jaeggli and K.J. Safir (eds.) *The Null Subject Parameter*, Kluwer, Dordrecht, 111-142.
Browne, W. (1974) "On the Problem of Enclitic Placement in Serbo-Croatian", in C. Brecht & C. Chvany (eds.), *Slavic Transformational Syntax*, Michigan Slavic Publications.
Brugger, G. (1990) Über obligatorische Elemente im "Restrictive Clause", Diplomarbeit, University of Vienna.
Cardinaletti, A. (1991) "On Pronoun Movement: The Italian dative *loro*", *Probus*, 3.2, 127-153.
———— (1992) "On Cliticisation in Germanic Languages", *Rivista di Grammatica Generativa*
———— (1993) "A Weak Pronoun in the Noun-Phrase: Italian possessive *loro*", paper presented at the Research Seminar, University of Potsdam.
———— **& M. Starke** (1994a) "Deficient Pronouns: A View from Germanic", paper presented at the Germanic Comparative Syntax Workshop, Harvard.
——————— (1994b)"Deficient Pronouns: A View from Romance", paper presented at the Linguistic Symposium on Romance Language, UCLA/USC.
——————— (1994c) "On Weak-Pronoun Learning", paper presented at *Workshop on the Acquisition of Scrambling & Cliticisation*, Bern.
Cartocci, A. (1990) 'Cose de ogi, cose de sempre', *Scenette comiche in vernacolo anconetano*, Ancona.
Chomsky, N. (1981) *Lectures on Government and Binding*, Foris, Dordrecht.
———— (1986) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.
———— (1991) "Some notes on Economy of Derivation and Representation", in R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge, Mass.
———— (1992) "A Minimalist Program for Linguistic Theory", in K. Hale & S. J. Keyser (eds.), *The View From Building 20*, MIT Press, Cambridge, Mass.
———— (1994) Class Lectures, MIT.
Cinque, G. (1988) "On *si* Constructions and the Theory of *arb*", *Linguistic Inquiry*, 19.4, 521-581.
———— (1990) *Types of A' Dependencies*, MIT Press, Cambridge, Mass.
———— (1993) "On the Evidence for Partial N-movement in the Romance DP", ms, Univ. of Venice.
———— (1994) "Sulla sintassi del participio passato nelle lingue romanze", paper presented at the Incontro di Grammatica Generativa, Padua.
Corver, N. and D. Delfitto (1993) "Feature Asymmetry and the Nature of Pronoun Movement", ms., Utrecht Univ.
Damourette & Pichon (1911/1957) *Des mots à la pensée*, Paris, D'Atrey.
Diesing, M. (1991) *Indefinites*, MIT Press, Cambridge, Mass.
Dobrovie-Sorin, C. (1992) *The Syntax of Romanian*, Mouton.
Enç, M. (1991) "The Semantics of Specificity", *Linguistic Inquiry*, 22.1
Fiengo, R. & R. May (1994) *Indices and Identity*, MIT Press.
Frank, R. & A. Kroch (1993) "Generalised Transformations in Successive Cyclic and Long Dependencies", paper presented at GLOW, Lund.
Giusti, G. (1993) *La sintassi dei determinanti*, Unipresss, Padova.
Grimshaw, J. (1991) "Extended Projection", ms., Brandeis University.
Guillaume, G. (1949/1973) *Leçons de linguistique*, Klincksieck, Paris

Harris, J. (1991) "The Exponence of Gender in Spanish", *Linguistic Inquiry*, 22.1., 27-62.
Haegeman, L. (1994) "Object Clitics in West Flemish", ms., University of Geneva.
Hall, T. (1992) *Syllable Structure and Syllable-related Processes in German*, Niemegen.
Halle, M. & A. Marantz (1993) "Distributed Morphology and the Pieces of Inflection", in K. Hale & S. J. Keyser (eds.) *The View From Building 20*, MIT Press.
Halpern, A. & J. Fontana (1992) "X° and X^{max} clitics", proceedings of WCCFL, 12.
Holmberg, A. (1986) "Word Order and Syntactic Features", PhD. Diss., Stockholm.
———— (1991) "The Distribution of Scandinavian Weak Pronouns" in Henk van Riemsdijk & Luigi Rizzi (eds) *Clitics in Romance and Germanic*, Eurotyp Working Papers 3, 33-76.
Jackendoff, R. (1994) "Lexical Insertion in a Post-Minimalist Theory of Grammar", ms., Brandeis University.
Jakubowicz, C. (1984) "On Markedness and Binding Principles", in C. Jones and P. Sells (eds), *Proceedings of NELS* , 154-182.
Jaeggli, O. (1982) *Topics in Romance Syntax*, Foris, Dordrecht.
Johnson, K. (1991) "Object positions", *Natural Language and Linguistic Theory*, 9.4, 577-636.
Joseph, B. (1993) "On the Absolute Nature of the Pro Drop Parameter", abstract of LSA paper, Ohio.
Kayne, R.S. (1975) *French Syntax*, MIT Press, Cambridge, Mass.
———— (1983) "Chains, categories external to S, and French complex inversion", *Natural Language and Linguistic Theory* 1.
———— (1984) *Connectedness and Binary Branching* Foris, Dordrecht.
———— (1991) "Romance Clitics, Verb Movement and PRO", *Linguistic Inquiry*, 22.4., 647-686.
Koopman, H. (1993), "The internal and external distribution of pronominal DPs", ms., UCLA.
Kroch, A. (1989) "Asymmetries in Long-Distance Extraction in a Tree Adjoining Grammar", in M. Baltin & A. Kroch (eds.) *Alternative Conceptions of Phrase Structure* The University of Chicago Press
———— **& A. Joshi** (1985) "The Linguistic Relevance of Tree Adjoining Grammar", UPenn Dept. of Computer & Information Sciences Technical Reports MS-CIS-85-16.
Laenzlinger, Ch. (1993) "A syntactic view of Romance pronominal sequences", to appear in *Probus*, 5.3.
———— (forthcoming) "Enclitic Clustering: the Case of French Positive Imperatives", ms. Univ. of Geneva.
Laka, I. (1990) Negation in Syntax: On the Nature of Functional Categories and Projections, PhD Diss, MIT
Lattewitz, K. (1994) "A minimalist view on verb raising", Paper presented at the Comparative Germanic Syntax Workshop, Harvard.
Longobardi, G. (1991) "N-movement in Syntax and in LF", University of Venice, Working Papers in Linguistics, 91.I.9.
Lonzi, L. (1991) "Which adverbs in specVP?", *Rivista di grammatica generativa*, 15, 141-160.
Montalbetti, (1984) "After Binding", PhD Diss, MIT.
Motapanyane, V. (1991) Theoretical implications of the complementation in Romanian, PhD Diss., University of Geneva.
Nespor, M. & I. Vogel (1986) *Prosodic Phonology*, Foris, Dordrecht.
Oberleiter, M. and I. Sfriso (1993) "Pronouns in the Olang Dialect", ms., University of Venice.
Otero, C. (1976) "The Dictionnary in Generative Grammar", ms., UCLA.
Perlmutter, D. & Oresnik (1973) "Language particular rules and explanation in syntax", in S. Anderson & P. Kiparsky (eds.), A Festschrift for Morris Halle, Holt, Rinehart, and Winston, New York.
Pesetsky, D. (1989) "Language-Particular Processes and the Earliness Principle", ms., MIT
Picallo, C. (1994) "Catalan Possessive Pronouns: The Avoid Pronoun Principle Revisited", *Natural Language and Linguistic Theory*, 12.2, 259-299.
Poletto, C. (1993) La sintassi del soggetto nei dialetti dell'Italia settentrionale, PhD Diss., University of Padua & Venice.
Pollock, J.-Y. (1989) "Verb-Movement, Universal Grammar, and the Structure of IP", *Linguistic Inquiry*, 20, 365-424.
Postal, P. (1969) "On the so-called "pronouns" in English", in D. Reibel and S. Schane (eds.), *Modern Studies in English*, Prentice-Hall, Englewood Cliffs NJ, 201-224.
Ritter, E. (1990) "Two Functional Categories in Noun Phrases: Evidence from Modern Hebrew", in *Syntax and Semantics* 26, Academic Press.
Rivero, M.L. (1992) "Adverb Incorporation and the Syntax of Adverbs in Modern Greek", *Linguistics and Philosophy* 15: 289-331.
Rizzi, L. (1982) *Issues in Italian Syntax*, Foris, Dordrecht.
———— (1986a) "Null objects in Italian and the theory of *pro*", *Linguistic Inquiry*, 17, 501-557.
———— (1986b) "On the status of subject clitics in Romance", in O. Jaeggli and C. Silva-Corvalán (eds.), *Studies in Romance Linguistics*, Foris, Dordrecht, 391-419.

———— (1990) *Relativised Minimality*, MIT Press, Cambridge, Mass.
Sandfeld, (1965), Syntaxe du Français contemporain, Paris
Schroten, J. (1992) "On Spanish definite determiners: Personal pronouns and definite articles", Recherches de Linguistique Romane et Française d'Utrecht, 9-24.
Sellani, R. (1988) 'Ade' nonna v'arconta S.nigaja' (100 sonetti in dialetto senigalliese), Tecnostampa Edizioni, Ostra Vetere.
Siloni, T. (1990) "On the parallelism between CP and DP: the case of Hebrew semi-relatives" in J. van Lit, R. Mulder & R. Sybesma (eds.) *Proceedings of LCJL* 1, 135-153.
Sportiche, D. (1989) "Movement , Agreement and Case", ms., UCLA.
———— (1992) "Clitic constructions", ms., UCLA.
Staaf, E. (1906) *Etude sur les pronoms abrégés en Ancien Espagnol*, Lundström, Uppsala.
Starke, M. (1993a) "Notes on prepositions and clause-structure", Pré-Mémoire, University of Geneva.
———— (1993b) En deuxième position en Europe Centrale, Mémoire de Licence, Univ. of Geneva.
———— (1994) "Two notes on Adverb-chains", ms. University of Geneva and MaxPlanck Berlin.
———— (to appear) "On the format for Small Clauses" in A. Cardinaletti & M.T. Guasti (eds.) *Small Clauses*, Syntax & Semantics, 27, Academic Press.
Szabolcsi, A. (1989) "Noun-phrases and clauses: is DP analogous to IP or CP?" to appear in J. Payne (ed.) *The Structure of Noun-Phrases*, Mouton, Amsterdam.
Tasmowski-De Ryck, L. & S.P. Verluyten (1982) "Linguistic control of pronouns", Journal of Semantics, 1.4., 323-346.
Toman, J. (1981) "Weak and Strong: Notes on *be* in Czech" *in* G. Brettschneider & C. Lehmann (eds.) *Wege zur Universalienforschung* Narr, Tübingen
———— (1993) "A note on clitics and prosody", in L. Hellan (ed.) *Clitics in Germanic and Slavic*, ESF Working Papers, 113-118.
Uriagereka, J. (1992) "Aspects of the syntax of clitic placement in Western Romance", ms., UMD.
Vergnaud, J.-R. (1974) "French Relative Clauses" PhD Thesis, MIT.
———— **& M.L. Zubizarreta** (1990) "On the Form and Interpretation of Definite Noun Phrases", ms. USC
———————— (1992) "The Definite Determiner and the Inalienable Constructions in French and English", *Linguistic Inquiry*, 23.4, 595-652.
Wexler, K. & Y.-C. Chien (1985) "The development of lexical anaphors and pronouns", *Papers and Reports on Child Language Development*, 24, 138-149.
Wilder, Ch. (1994) "Coordination, ATB and Ellipsis" GAGL, 291-329.
Yvon, H. (1949) "Le, la, les, articles ou pronoms?" *Le Français Moderne* , 251-264.
———— (1957) "Le, la, les, articles ou pronoms? (Supplément)" *Le Français Moderne* , 248-255.
Zwart, C.J.W. (1992) "Notes on Clitics in Dutch", in L. Hellan (ed.) *Clitics in Germanic and Slavic*, ESF Working Papers, 19-155.

SCRAMBLING AND INCORPORATION IN TURKISH*

Jaklin Kornfilt, Syracuse University

1. Turkish has a syntactic phenomenon which I shall refer to as "Case Drop"; this is to be understood as a purely descriptive term. Case Drop is found with structural Case[1] only (most conspicuously, with Accusative), and under certain semantic/pragmatic conditions whose nature shall not concern us here[2]. It will suffice for our purposes to say that non-specific NPs which are in a syntactic configuration where they are assigned structural Case are not marked with an overt Case morpheme and have to surface to the immediate left of the verb of their clause. Some examples of Case Drop follow:

"Nominative Drop":

(1) a. Çocuğ-u arı sok - tu
child-Acc. bee sting-Past
'Bees stung the child'

---

* I would like to thank Bahar Arsoy, Ayşe Can, and Mehmet Yanılmaz for sharing their native intuitions with me, to Gerald Greenberg for discussion of some of the issues addressed here, and to Peggy Speas for her comments on a previous draft of this paper. All errors and shortcomings are my responsibility.

[1]By "structural Case" I mean those Cases that are assigned independently from θ-roles and which are assigned after D-Structure; for Turkish, these are Nominative, Accusative, and Genitive. Nominative and Genitive are assigned by verbal versus nominal AGR, respectively; Accusative is assigned by V. In contrast, I assume that oblique Cases enter the level of D-Structure already attached to their NPs and are necessary for successful thematic linking of those NPs to the verbs that select for such Cases. The term "structural Case", the way it is used here, is independent from morphological realization; thus, an NP can bear overt or non-overt structural Case. As it happens, oblique Case is always overt in Turkish. Accusative and Genitive are overt, unless they undergo Case Drop. Nominative is never overt. However, I view the "silent" nature of the Turkish Nominative as a very superficial property and choose to group the Nominative as an overt Case for syntactic purposes. A parallel can be drawn to pro as an empty element which is, however, exempt from the ECP and behaves in many respects as an overt pronominal.

[2]For detailed descriptions of this phenomenon, the reader is referred to Dede, Tura, Erdal and Nilsson as recent sources. Most textbooks of Turkish grammar will mention this phenomenon.

"Genitive Drop":

(2) a. [Çocuğ-u arı sok - tuğ -un -u] duy -du -m
child-Acc. bee sting-Nomin.-3.sg.-Acc. hear-Past-1.sg.
'I heard that bees stung the child'

"Accusative Drop":

(3) a. Ahmet bütün gün pasta ye-di
whole day cake eat-Past
'Ahmet ate cake all day long'

The corresponding NPs with overt Case are not limited to the pre-verbal position; as subjects, their canonical position will be sentence-initial; as direct objects, while their unmarked position is pre-verbal, they can be scrambled freely:

(1) b. Arı çocuğ-u sok - tu
bee child-Acc. sting-Past
'The bee stung the child'

(2) b. [arı -nın çocuğ-u sok - tuğ -un -u] duy -du -m
bee -Gen. child-Acc. sting-Nomin.-3.sg.-Acc. hear-Past-1.sg.
'I heard that the bee stung the child'

(3) b. pasta-yı Ahmet dün akşam ye-di
cake-Acc. yesterday evening day eat-Past
'Ahmet ate the cake yesterday evening'

In this short paper, I shall offer examples showing that word-order rigidity arising from "Case Drop" can lead to word-order flexibility. More specifically, I shall suggest that "Case Drop" is due to noun incorporation into the verb, and that such incorporation of N-heads enables stranded subconstituents of NP to move out of NP, since the trace left by scrambling can be antecedent-governed, while such antecedent-government is impossible without incorporation, due to the nature of the NP as a barrier to government. Incorporation voids the NP of barrierhood, since it renders the head of the NP non-distinct from the complex verb after incorporation.

I now turn to some examples to illustrate the correlation between scrambling and incorporation.

2. While Turkish is known to be rather word-order free (to such an extent that even non-verb-final orders are possible), phrases like NPs and PPs cannot usually be broken up:[3]

(4) a. Dün sokak-ta [çok yaşlı bir adam-a] rasla-dı -m
yesterday street-Loc. very old a man-Dat. meet-Past-1.sg.
'Yesterday I met a very old man in the street'
b.*Dün sokak-ta [e_i bir adam-a] rasla-dı-m çok yaşlı$_i$ [4]

Suppose we have a theory of boundedness that rules out the ungrammatical instances of Scrambling, because nothing can adjoin to the NP (since it is an argument) before moving further, as indeed it cannot:

(5) *Dün sokak-ta [e_i bir adam-a] çok yaşlı$_i$ rasla-dı-m

---

[3]One exception to this generalization is a possessor in a possessive NP:

(i) a. [Ahmed-in karı -sın -ı] tanı -mı -yor -um
-Gen. wife-3.sg.-Acc. know-Neg.-Pres.Progr.-1.sg.
'I don't know Ahmet's wife'
b. [e_i karısını] tanımıyorum Ahmed-in$_i$

Note that a possessor and the head in a possessive NP are "linked", since the head is marked for agreement with the possessor; modifiers, quantifiers and *other material in an NP, on the other hand, do not agree with the head in any* way.
Let us say that the reason why possessors appear to be moveable out of their NP is that there is a resumptive pro in their original position, licensed and identified by the agreement marker on the phrasal head. In non-possessive NPs, scrambling of subconstituents is impossible, since a pro would not be licensed, and a non-pronominal empty category would be illicit, due to locality constraints.

[4]Such constructions are sometimes found in poetry, but not in prose.

3. I now turn to the type of example which is of crucial interest to this paper, namely a construction where incorporation and scrambling interact. In such examples certain phrasal parts can move out, if the head has been incorporated into the V, i.e. if the head lacks the overt structural Case marker that is expected; corresponding examples where the head carries an overt structural or oblique Case marker are ungrammatical:

(6) a. ?Bir daha [e_i bir terzi[5] bul-a -ma-m sen-in gibi$_i$
one time a taylor find-Abil.-Neg.-1.sg. you-Gen. like
'I won't ever be able to find a taylor like you again'

b. *Bir daha [e_i bir terzi-yi] bul-a-ma-m sen-in gibi$_i$[6]
-Acc.

(7) a. ?[e_i Bir haydut] gör-dü -m dev gibi
a robber see-Past-1.sg. giant like
'I saw a robber (big) like a giant'

b. *[e_i Bir haydut-tan] kaç-tı -m dev gibi
-Abl. flee-Past-1.sg.
'I fled from a robber (big) like a giant'

---

[5]It might be objected here that a string like bir terzi is a syntactic N' rather than a bare N. If so, it wouldn't be clear how incorporation could take place. However, I assume that only items like terzi, an N, incorporate, leaving the rest of the phrase behind. The question that now arises is whether we wouldn't be--wrongly--predicting that the string [bir t] can scramble freely. This is not a problem, however, for two reasons: firstly, bir certainly cannot scramble by itself, since only maximal projections seem to be able to scramble. Secondly, even if the string [bir t] were a maximal projection (rather than an X', as I am assuming here), it wouldn't be able to scramble, either, since there is independent evidence that empty-headed maximal projections cannot scramble. (The reason for this restriction might well be due to the ECP, as is suggested in Lamontagne and Travis (1987).)

[6]Note that (6)b. is ungrammatical only under the crucial reading where the scrambled PP originates within the NP object. The example is grammatical under the (for us) irrelevant reading: 'I cannot find a taylor like you did (i.e. in the manner in which you found a taylor).

Needless to say, the non-scrambled versions of all of these examples are grammatical. The generalization seems to be that when the head of the phrase cannot be incorporated because it is overtly marked for Case, the rest of the phrase cannot scramble out.

4. Time has come to address the question of how incorporation yields the two word order effects just illustrated, namely the fixed pre-verbal position of non-specific NPs without overt structural Case morphemes on the one hand, and the freedom of subconstituents of such NPs to scramble out of the larger phrase on the other. I address these two issues in turn.

I am assuming with Baker (1988) that incorporation is due to head movement. The moved (N-)head then forms a complex word with the verb. The verbal complex governs the whole NP (just as the simple verb did before the movement) canonically, which is to the left in Turkish, since the language is head-final; thus, the verb also governs the head position of the NP. The trace left behind by the moved N-head is thus properly governed, while such a trace in head position of an NP which is in (VP-external) canonical subject position (i.e. in Spec/IP) (or in any other VP-external "scrambled" position) would violate the ECP. The obligatory position of NPs without overt (structural) Case to the immediate left of the V is thus explained.

Note that I am assuming, along with Baker, that incorporation from canonical, IP-initial subject position is not possible. However, in a language like Turkish, subjects can incorporate nevertheless, since they can optionally originate in VP-internal position. I thus assume an account similar to what has been proposed by other syntacticians for Germanic languages like Dutch and German, where a VP-internal subject can receive Case in-situ (cf. den Besten 1984, Reuland 1990 and others) and which I advanced as a possible analysis for Turkish in Kornfilt 1984. AGR, which is the Case assigner to subjects, is part of the verb in these languages and is thus able to govern and, consequently, to Case-mark within the VP. If so, subjects as well as objects can have their head nouns incorporate into the verb under government.

Before turning to the second central issue, namely how incorporation can make scrambling possible, I would like to address the following question:

5. Why should Case Drop be limited to structural Case?

Let us assume that oblique Case enters syntax already pre-attached to its phonological host, so as to enable the NP to receive a θ-role. Structural Case, on the other hand, is assigned in the syntax, and the appropriate θ-roles can be assigned to (as yet) un-cased arguments.

Assume now further that NPs are actually embedded within Case Phrases (KPs), as complements of a K-head. Where such a K-head is phonologically filled with an overt Case marker (as it always has to be for oblique Case and may be so filled for structural Case), the N-head of the NP cannot move further up into the V[7]; even if it could, its trace would not be governed by the V: the overtly headed KP would act as a barrier to government. I am further assuming that functional heads like K and AGR cannot incorporate into the lexical category V. Thus, the NP-complement of a KP or of an AGR-P (cf. Kornfilt 1984, corresponding to DP in Abney 1987) is never stranded as a result.[8]

Where the K-position is empty, the N could move into that position, and further into the V; the traces left behind would both be properly governed.[9] (Note, incidentally, that if the assumptions made here are correct, we would have an

---

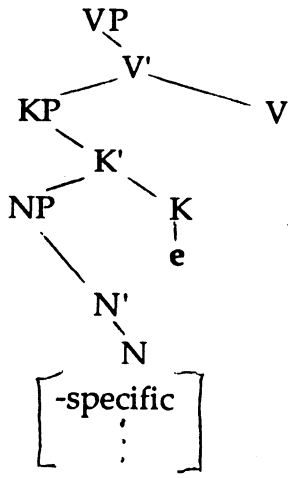
[7]I am thus assuming that the moving N, which might have undergone movement to a filled K, cannot move onwards to V, taking the K along. This is presumably due to the aforementioned restriction against functional categories incorporating into V. Neither can such an N excorporate from a filled K.

[8]If such categories were to incorporate, two problems would arise: Where an overt functional head incorporates, it would not be obvious how to then raise the N so as to form the overtly cased or overtly agreeing word; where the empty K incorporates, the stranded NP without overt Case would be predicted to be able to scramble freely, and this is obviously impossible, as we saw earlier in this paper.

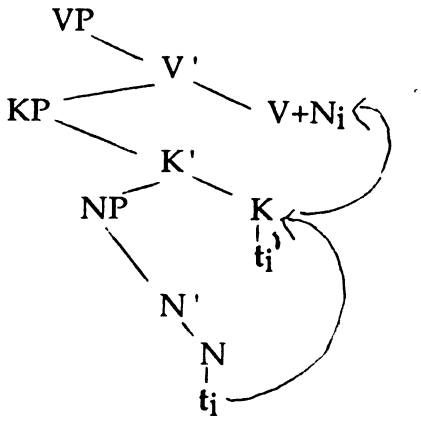
[9]The trace in the K-position would be properly governed by the V, or, actually, by the incorporated antecedent in the verbal complex; the trace in the N-position would be governed by the intermediate trace in K.

additional argument in favor of Baker's claim that incorporation is syntactic.) The following rough diagrams illustrate the structures before and after noun incorporation:

Before incorporation



After incorporation



While this account covers the facts of Case Drop[10], it is not immediately corroborated by word-order facts, since Turkish is head-final, and it is not obvious that noun incorporation indeed strands the remainder of the NP: The sequential order between the (putatively) stranded material and its head would remain the same after incorporation.

Furthermore, incorporation in Turkish (if this is indeed the nature of Case Drop in Turkish) does not involve any change in grammatical relations for the remainder of the NP--a phenomenon of particular interest for some of the languages studied by Baker.

In spite of these inconclusive points, I would like to claim that, through the interaction of scrambling and incorporation, Turkish does make an interesting

---

[10]Actually, I have not addressed the question of why Case Drop is restricted to--and, indeed, obligatory in--non-specific NPs. This is a very wide-ranging topic which cannot be made justice to in this short paper which concentrates on a different point. However, I would like to mention an idea which I have sketched out elsewhere (cf. Kornfilt forthcoming), which posits overt structural Case realization as a PR phenomenon. The phonological feature matrices in K need the presence of the feature [+specific] in N, as an assimilatory requirement of sorts. Where the feature [specific] has a negative value, it will be treated as absent altogether, and the phonological features will not be filled in.

contribution with respect to incorporation. Note that these facts have remained undiscussed in the literature on Turkish syntax so far. Furthermore, they are of a type expected by Baker's theory and yet not often found, or, if found, are "murky" (cf. Baker, p. 103).

6. I would like to suggest the following account to explain the observed correlation between incorporation and scrambling:

There is a weak Subjacency violation in all of these examples, and this explains why even the better examples are not perfect. If so, it is irrelevant for the (im)possibility of adjunction to NP whether there is a trace (left by incorporation) in the head position of the NP as in the a.-examples of (6) and (7) or whether there is a full N as in the b.-examples; such adjunction is always (weakly) ruled out.

In order to explain why the b.-examples are worse, I will take recourse to the notions of government and Case (as assigned under government).

I shall assume, along with Baker and Chomsky (1986), that a verb can govern its complement and the head of that complement, but not the rest of that complement phrase. In other words, the maximal projection node dominating the complement phrase acts as a barrier between the verb and the non-head part of the phrase. However, when the head of the phrase moves into the verb, a movement chain is established between the trace in head position and the antecedent within the complex verb. According to B., the head of the previous barrier is now <u>not distinct</u> from the complex V, due to this movement chain (since the complex V includes the antecedent of the trace in the head position of the former barrier); therefore, the phrasal projection is not a barrier any longer. The verb will now govern whatever material the incorporated head governed previously (cf. Baker, p. 64, definition 65 of the Government Transparency Corollary: "A lexical category which has an item incorporated into it governs everything which the incorporated item governed in its original structural position.").

Why should government by the verb be important for the trace left by scrambling? Note that this trace is governed by the head of the NP in the ungrammatical examples; hence, no ECP violation should ensue.

However, the scrambling trace needs to be antecedent-governed, since it is not θ-governed (cf. Chomsky (1986), p. 17). Where the NP out of which scrambling has taken place is overtly headed, the NP-projection will act as a barrier to antecedent-government. On the other hand, where incorporation has applied, the NP is not a barrier to government, as just explained. I assume that the scrambled antecedent has adjoined to VP; if so, the VP is not a barrier to antecedent-government, either (since it does not exclude the adjoined antecedent).

What is important here is not that the verb governs the scrambling trace, but rather that antecedent-government of that trace is not blocked--neither by the NP itself, nor by the verb as a closer governor. The latter is due to Minimality: According to Chomsky's (1986) definition of "narrow" Minimality (cf. p. 42, #91) the verb would not be a closer governor than the antecedent, since the VP doesn't immediately dominate the scrambling trace; Rizzi's (1989) notion of Relativized Minimality would have the same effect, since V would be a "Head-Governor" and not an "Antecedent-Governor". What is important, rather, is that the original barrierhood of the NP is voided, due to incorporation of its head (and the "non-distinctness", in Baker's terms, of the head of the NP and the complex V). As a result, antecedent-government of the scrambling trace becomes possible.

Note that the array of the scrambling and incorporation facts we have seen is just as predicted by Baker via his Government Transparency Corollary. However, Baker also states that while possessor stranding examples due to N-incorporation are found, they are more restricted than his theory would predict (p.103), and also that "complement raising" is apparently not found (p.104), while predicted by the theory; both types of "data available in the literature are unfortunately murky and unclear" (p.104). If the treatment of the Turkish stranding facts offered here is on the right track, the data are of the kind predicted and, while coming from an unexpected source, are neither murky nor unclear.

References:

Abney, S. P. (1987) The English Noun Phrase in its Sentential Aspect, unpublished Ph.D. dissertation, MIT, Cambridge, Massachusetts.

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

Besten, H. den (1984) "The Ergative Hypothesis and Free Word Order in Dutch and German", in J. Toman, ed., Studies in German Grammar, Foris Publications, Dordrecht, 23-64.

Chomsky, N. (1986) Barriers, The MIT Press, Cambridge, Massachusetts.

Dede, M. (1986) "Definiteness and Referentiality in Turkish Verbal Sentences," in D. I. Slobin and K. Zimmer, eds., Studies in Turkish Linguistics, John Benjamins, Amsterdam/Philadelphia, 147-163.

Erdal, M. (1981) "Turkish Participles and the Absence of Reference," in D. W. Young, ed., Studies Presented to H. J. Polotsky, Pirtle and Polson, Beacon Hill.

Kornfilt, J. (1984) Case Marking, Agreement, and Empty Categories in Turkish, unpublished Ph.D. dissertation, Harvard University, Cambridge, Massachusetts.

Kornfilt, J. (forthcoming) "Naked Partitive Phrases in Turkish", to be published in J. Hoeksema, ed., Studies on Partitives, Mouton-deGruyter, Berlin.

Lamontagne, G. and L. Travis (1987) "The Syntax of Adjacency", Proceedings of WCCFL, vol, 6, 173-186.

Nilsson, B. (1986) "Object Incorporation in Turkish,", in A. Aksu Koç and E. Erguvanlı Taylan, eds., Proceedings of the Turkish Linguistics Conference 1984, Boğaziçi University Publications No. 400, Istanbul, 113-128.

Reuland, E. (1990) "Head Movement and the Relation Between Morphology and Syntax", in G. Booij and J. van Merle, eds., Morphology Yearbook 3, 129-161.

Rizzi, L. (1990) Relativized Minimality, MIT Press, Cambridge, Massachusetts.

Tura, S. (1986) "Definiteness and Referentiality in Turkish Verbal Sentences," in D. I. Slobin and K. Zimmer, eds., Studies in Turkish Linguistics, John Benjamins, Amsterdam/Philadelphia, 165-194.

Ewald Lang (HUB & FAS, Berlin)

Basic Dimension Terms: a first look at universal features and typological variation*

0. Introduction

Though published some 25 years ago, Berlin & Kay's (1969) pioneering study on color terms (which the title of the present paper alludes to) has not found its counterpart in the field of spatial vocabulary up to now. This is indeed a surprising gap in the literature - in view of the observation that all languages seem to have a certain sample of lexical items to make reference to spatial dimensions such as height, length, width, depth, distance etc.[1], and taking into consideration that our spatial concepts more or less directly originate in human perceptual endowment, which provides the source of their supposed universality.

In the past two decades, there was a considerable amount of research work on space perception, shape recognition, visual discrimination etc. which attempted to prove or at least to support the claim that "the dimensions languages pick out are just those dimensions the human perceptual apparatus is tuned to pick out" (cf. Clark 1973, Clark & Clark 1977). Plausible as this guideline undoubtedly may be, things turned out to be more complicated. In the meantime we know that there is more to dimensional designation than perception-based categorizing of axes, planes, extensions etc. and simple projecting of top - bottom, front - rear, left - right sides from some observer-based body-schema (cf. Herskovits 1986) onto spatial objects.

In a series of studies (notably Bierwisch & Lang (eds.) 1989; Lang 1990 a,b; Lang, Carstensen & Simmons 1991), the grammar of Dimension Assignment (DA), i.e. the set of conditions according to which natural languages pick out and lexically encode spatial dimensions, has been shown to comprise at least the following components:

(I) **Position and gestalt properties.** DA is basically organized not by a single body-schema but by two interacting categorization grids called "Primary Perceptual Space" (PPS) and "Inherent Proportion Schema" (IPS), both being independently traceable to human perceptual endowment ;

(II) **Parameters.** DA makes use of a limited set of Dimension Assignment Parameters (DAP) which - emerging from the grids in (I) - go to make up primary candidates for being lexicalized and thus provide the basis to look for universals of how natural languages encode spatial dimensions;

(III) **Conceptual-Semantic Interfaces .** DA involves a set of devices that account for the differentiation between, as well as the interaction of, non-linguistic conceptual *world knowledge* and language-bound *word knowledge* as regards the way in which spatial objects are assigned primary and/or contextually induced dimensions in terms of linguistic expressions.

---

* Earlier versions of the paper were presented at the FAS Inaugural Conference " Linguistic Universals & Typological Variation", Berlin, March 17-19, 1994, and at the L.A.U.D. Workshop "Language & Space", Duisburg, March 22-24, 1994. Special thanks go to Paul Kay (Berkeley) for encouraging discussions on the methodology of such an investigation. For providing and/or checking the data which this study draws on I am indebted to Byong-Rae Ryu [Korean], Horst D. Gasde, Chen Xuan [Mandarin], Joanna Blaszczak [Polish], Svetlana Poljakova, Vladimir Klimonov [Russian], and Marcela Adamiková [Slovak]. For stimulating ideas I am grateful to Zubin & Choi (1984), Zubin & Svorou (1984).

[1] Bierwisch's seminal paper of 1967, which proposed a handy set of features, inspired much work in acquisition research but has not been challenged as semantic classic or revised until the appearance of Bierwisch & Lang (eds.) 1989.

After having gained some idea of what is going on in DA in general and having worked through the grammar of DA in German and English in some detail, which inter alia includes a PROLOG implementation (Lang, Carstensen & Simmons 1991) that served to prove and improve the DA model outlined in (I) - (III), we now may feel encouraged to tackle questions of universal features and typological variation in the realm of basic dimension terms.

The paper is divided into five sections. Section one presents a brief overview of the framework mentioned in (I) - (III) above; sect. two offers a couple of preliminary universals of lexicalization regarding dimension terms; sect. three illustrates typological variations within the scope of the universals outlined so far; sect. four examines - within a group of cognate languages that share the same set of lexical items for DA - a case of variation that is induced via interference by a language from outside the group; sect. five gives a brief outlook on problems to be tackled in the future.

1. Overview of the framework

In order to be brief, I will outline the analysis of DA as presented in Lang 1990 a,b; Lang, Carstensen & Simmons 1991 by means of a summarizing diagram - cf. Fig. 1 below. It indicates the way in which perceptual information (from vision, organ of equilibrium, upright walk etc.) is being conceptually categorized by two independent but interacting grids, PPS and IPS, each of which contributes in its own way to delimit and identify what is to be taken as a spatial dimension.

(1) The main tenet is that DA to spatial objects works by designating certain axis extensions of a given object as **spatial dimensions** by picking out some axis extension **d** of object **x** due to

(a) **d**'s coincidence with an axis of PPS (e.g., an axis **d** of object **x** is designated as "**x**'s height" if **d** coincides with the Vertical, or as "**x**'s depth" if **d** coincides with the Observer Axis etc.) or/and

(b) **d**'s showing some distinctive gestalt feature as defined by IPS (e.g., an axis **d** of object **x** is designated as "**x**'s length" if **d** is the Maximal axis of **x**, or as "**x**'s thickness" if **d** is the SUBstance Axis of **x** etc.).

Fig.1 below shows three levels on which spatial information involved in DA has to be represented.

The Perceptual Level is determined by sensory perception that emerges from how our senses instatiate physical parameters, in the case of spatial perception it is above all those parameters that in one way or other derive from the force of gravity.

The Conceptual Level is determined by what results from categorizing perceptual input by means of PPS and IPS in view of its relevance to human behaviour. In other words, perceptive distinctions are conceptualized only to the extent that they are needed for the naive physics which underlies our everday knowledge of space. The conceptual categorization of spatial objects by PPS and IPS provides us with an inventory of spatial features that are essential to the way in which DA works in natural languages. The role of this inventory of features is twofold:

(2) The spatial features that emerge from PPS and IPS, while keeping their conceptual content constant, occur in two representational formats:

(a) as entries of so-called Object Schemata (OS), that is, as elements which our conceptual representations of spatial objects are made of (= Conceptual Level in Fig. 1)

(b) as linguistically relevant Dimension Assignment Parameters (DAPs), that is, as semantic elements which the meanings of dimension expressions are made of (= Semantic Level in Fig. 1)

It is important to note that the DAPs form but a designated proper subset of the inventory of spatial features defined by PPS and IPS, that is, only a subset of the entries in OS also occur as DAP. This reflects the basic idea that DA rests on designating certain specified object extensions as spatial dimensions. The components shown in Fig 1. will be briefly commented upon in the sequel.

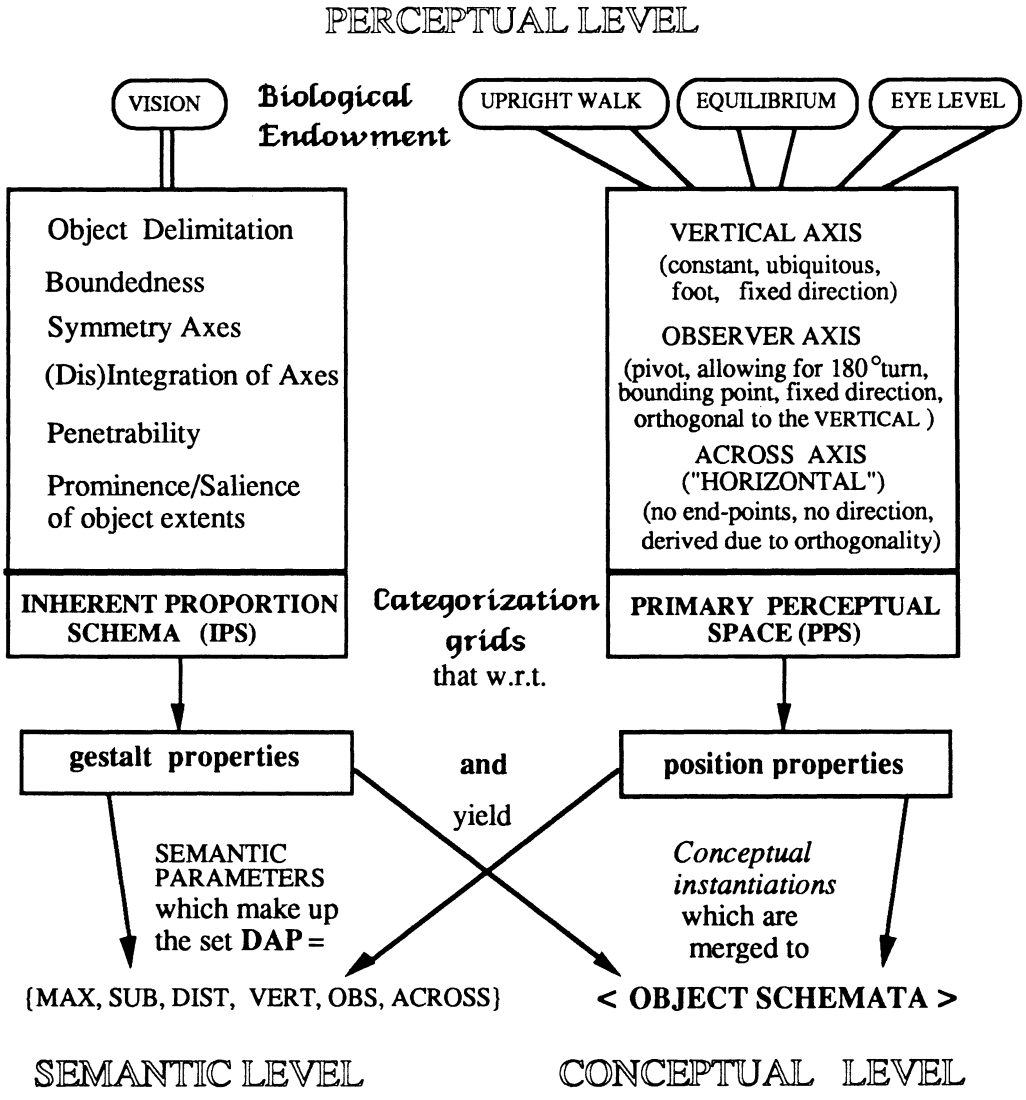


Fig. 1 Cognitive components involved in assigning spatial dimensions to objects

1.1 Primary Perceptual Space (PPS).

PPS consists of a system of axes that form our internal model of the external space. The PPS axes include the Vertical, the Observer axis, and the Horizontal or Across axis, each of which has its own properties relevant to the way in which we conceive of the spatial environment around us.

(3) It is by reference to PPS as a categorization grid that objects
  (a) are assigned a **position** and/or a **location** in space
  (b) are **localized** with respect to one another
  (c) are said to **move** (i.e. to change their position and/or location)

Taking reference to the Vertical as a case in point, we may illustrate (3) (a - c) with respect to some of the spatial relations between the objects shown in **Fig. 2** below by the sample given in (4).

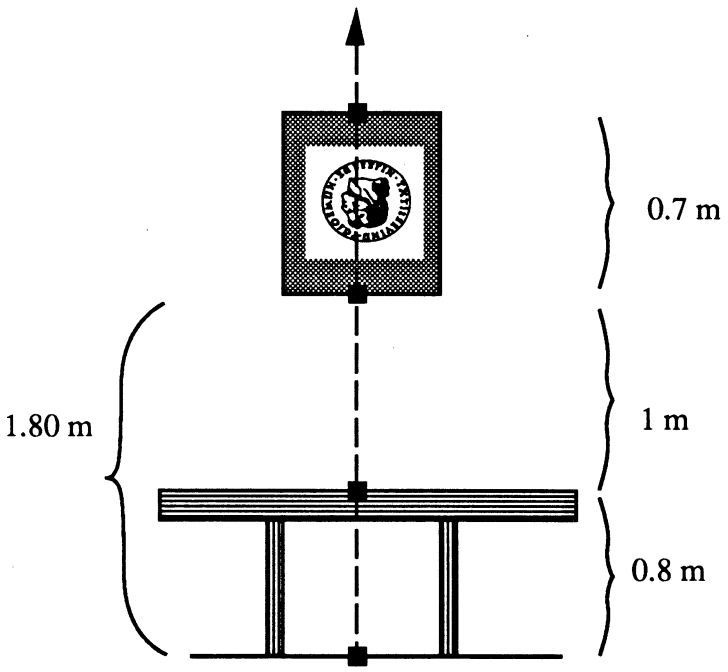


Fig. 2

(4) (a) The *table* is 0.8 m high / 0.8 m in height / is standing / is upright [position]
The *poster* is 0.7 m high /0.7 m in height / is hanging / is upside down
The *poster* is 1.80 m high, at a height of 1.80 [location]

(b) The *table* is under /below the poster [localization]
The *poster* is 1m above / 1m higher than the table

(c) The *poster* is aligned with the table / turned upside down [change of position]
The *poster* was lifted / raised to a place above the table [change of location]

First, all underlined items in (4) involve reference to the Vertical of the surrounding space. This obviously forms a constitutive part of their lexical meaning and is being represented by the parameter VERT on the Semantic Level (cf. Fig 1 and 1.3). Thus, VERT is present in the meaning of each of these items. However, the way VERT is packaged into the lexical meanings differs, depending inter alia on the syntactic category of the given item (for dimension adjectives - cf. Lang 1989; for position verbs - cf. Maienborn 1990a,b, 1993; Lang, Carstensen, Simmons 1991; for projective prepositions - cf. Lang 1991, 1993).

Second, in addition to that, reference to the Vertical is also an essential part of our conceptual knowledge of spatial objects such as "table" or "poster". The concept "table" (and possibly also the semantic representation of the word *table* [2]) includes the feature of having a canonical orientation. A table has an axis **d** which is designated for being aligned to the Vertical and which thereby defines a table's canonic height, i.e. its canonic top - bottom extension, and hence its normal position in the spatial environment. The concept "poster" (and possibly also the semantic representation of all picture nouns) includes the feature of having an inherent orientation. A poster (or picture etc.) has an axis **d** which (due to the object it shows) is assigned an inherent height, i.e. an inherent top - bottom extension, absolutely independent of the Vertical and of the poster's actual position.

When lexical items like *stand, lie, upright, upside down* etc. are predicated of spatial objects, they explicitely refer to such conceptually fixed verticality features. Based on this the examples in (4) get regular interpretations whereas things like **The ball /line/hole is standing upside down* are ruled out as (conceptually) deviant. Hence, having a fixed, canonic or inherent orientation or no orientation at all, is part of the representation of spatial objects on the Conceptual Level (cf. Fig.1). Technically, this is achieved by providing the Object Schemata of "hill", "table", "tree" "poster" etc. with an obligatory primary entry *vert* (which is only typographically distinct from the semantic parameter VERT) while preventing the Object Schemata of "ball", "line", "hole" etc. from being available for any verticality specification whatsoever.

Instead of telling the analogous story about the Observer Axis of the PPS, the semantic parameter OBS involved in the meaning of *deep, in front of, behind* etc. and in the canonical vs. inherent perspectivization of spatial objetcs on the Conceptual Level (e.g. "river", "cupboard" vs. "hole", "wound"), I refer to Lang, Carstensen, Simmons (1991).

I should note, however, that the third axis of the PPS, the so-called Horizontal or, as I prefer to say, the Across axis, is rather different in that it lacks the power of categorizing spatial objects into subclasses based on putative features of canonical or inherent horizontality or transversality. Due to their differing origin in human perceptual endowment, the three axes of the PPS have distinct properties (briefly noted in Fig.1, in more detail discussed in Lang 1989) and hence each of them plays a distinct role in determining our internal model of the external space. In short:

(5) The **Vertical**

> Originating from the effects of gravitation as perceived by the organ of equilibrium, the Vertical is an orientation cue which is **ubiquitous and constant**, that is, available everywhere and with the same effect at all times. The Vertical is physically and conceptually the **most salient** and also the **dominant** axis of PPS ; the other axes are defined in relation to the Vertical.

---

[2] The question of whether all or only a specified subset of the spatial features emerging from PPS and IPS should be considered to be part of the lexical meaning of *table, poster* etc. is one of the facets of the world knowledge vs. word knowledge debate within the two-level approach to semantics (cf. Lang 1994). For the purpose of this paper, I will adopt the latter view without further ado.

(6) The **Observer axis**

Originating in the visual organ, this axis is determined by the line of sight of a (potential or actual) **observer**. Due to this, the Observer axis is flexible in two respects:
(a) it is - unlike the gravitation based Vertical - not anchored in the physical space but rather induced by a movable and moving human interpreter of the physical space;
(b) it has an anatomically determined pivot allowing for a 180° turn in either of two planes.

In the **unmarked** case, given by the position of the eyes of an observer in upright posture, the Observer Axis is orthogonal (at **90°**) to the Vertical. In the **other relevant** configuration, the Vertical and the Observer axis lie at an angle of **180°** such that they run parallel but in diametrically opposed directions. The Observer axis is the source of depth perception.

(7) The **Across axis**

The third axis of PPS is not an axis we are endowed to identify by primary perceptual information, but is **derived from**, hence **dependent on**, the two others just to fill the gap determined by the properties of the latter. The Horizontal or Across axis is exclusively defined by its **orthogonality** to the Vertical and to the Observer Axis.

PPS inter alia provides us with the semantic parameters VERT and OBS (and possibly ACROSS) that are relevant to DA. As we will see in sections 3 and 4, the different status of these axes has direct consequences for delimiting the scope of universal features and typological variation in the lexicalization patterns that cross-linguistically underlie basic dimension terms.

1.2. Inherent Proportion Schema (IPS)

Spatial objects are furthermore categorized by their gestalt properties, i.e. by features drawing on whether or not an object has boundaries, symmetry axes, salient axes (e.g. a Maximal axis, which terms like *long, short, tall, along* etc. refer to), on whether an object axis is visually penetrable (distance axis) or not (substance axis), and on an object's dimensionality (1D, 2D, 3D). IPS provides us - among other things - with the semantic parameters MAX, SUB, DIST (and possibly ACROSS) that are relevant to DA.

1.3 Dimension Assignment Parameters (DAP)

The **interaction** of the two categorization grids PPS and IPS provides us with an inventory of semantic Dimension Assignment Parameters (DAP) and with an inventory of Object Schemata (OS) of conceptually admissible objects specifying the full range of dimension, position, and mobility properties of spatial objects. The latter inventory, which yields a complete catalogus mundi of possible spatial objects, cannot be repeated here (see Lang, Carstensen, Simmons 1991).

I will, however, enumerate the inventory of semantic DAP (with the exception of the holistically assigned parameter SIZE involved in e.g. *groß - klein* , see Lang 1989). Note that things like MAX, SUB etc. are not mere labels, but theoretical constructs having a clear-cut interpretation within the scope of the DA model outlined in the Introduction. Slightly simplified, the conditions encoded in each of these DAPs may be spelled out like this:

MAX identifies the most extended disintegrated axis of some object **x**, which in turn presupposes there to be exactly one such axis of **x** available (that is what makes *long - short* inapplicable to circles, squares, balls etc.).

SUB identifies either a non-maximal third axis (cf. *thick board, thin slice of bread*) or a non-maximal integrated axis, e.g the diameter of a circular cross section (cf. *thick pole*).

DIST identifies an object axis perceived as inside diameter of a hollow body. Though SUB and DIST identify the same type of axis in terms of geometry, they draw on mutually exclusive perceptual properties viz. permitting or preventing being looked through. Thus, SUB refers to axes determined by solid (parts of) objects, DIST to axes determined by hollow ones.

VERT identifies, if assigned via *high* or *tall* etc. to some spatial object **x**, exactly that disintegrat ed axis of **x** which coincides with the Vertical of PPS.

OBS identifies, if assigned e.g. via *deep* to some spatial object **x**, any disintegrated (non-minimal) axis of **x** which coincides with the Observer axis of PPS.

ACROSS designates a disintegrated object axis which is left unspecified by any of the other DAPs referring to maximality, substance, verticality, or alignment to the Observer axis.

Notice that ACROSS is a stop-gap with respect to both IPS and PPS. Within PPS, ACROSS covers horizontality in that it is assigned to an axis to which neither VERT nor OBS apply; within IPS, ACROSS supplements the parameters MAX and SUB in that it is assigned to an axis to which neither of these applies. In other words, ACROSS represents the overlap of the two categorization grids PPS and IPS, and due to this it provides a major source of ambiguity within and typological variation between languages - as will become clear in sections 3 and 4.

To sum up: it is this small set of semantic parameters DAP = {MAX, VERT, OBS, ACROSS, DIST, SUB} which controls the way in which natural languages assign spatial dimensions and positions to objects[3]. DAP is the stuff which the lexical meanings of dimension terms are made of. Taken as categorized semantic components, the elements of DAP are packaged into more complex representations that e.g. for dimensional adjectives meet the following general format (for details see Lang 1989, Bierwisch & Lang 1989):

(8) $\lambda c\ \lambda x\ [\text{QANT}\ [\text{DIM}\ x]\ =\ [v\ \pm\ c]]$

DA crucially involves gradation and comparison, thus QUANT is a semantic component for a scaling operation which assigns a scale value composed of **v** and **c** to some spatial object **x** with regard to a dimension **d**. The latter is represented by the placeholder DIM, a variable that is to be replaced by the elements of DAP. Though (8) offers a lot of aspects that invite further typological considerations, this paper will focus on the dimension component DIM in discussing the conditions on which the elements of DAP are lexicalized and arranged to lexical fields. Next step in doing this is to take a look at the internal structure of the set DAP = {MAX, VERT, OBS, ACROSS, DIST, SUB} w.r.t. the division of labor between, and the mutual compatibility of, the elements it contains.

---

[3] Note that the approach adopted by Herskovits (1986) and other Cognitive Semanticists is, roughly speaking, observer-centered and situation-based. The approach I am advocating is object-centered and axes-based. This difference in view, I hope, will stimulate the discussion.

1.4. Compatibility conditions.

By claiming that DA to spatial objects is the joint outgrowth of the two categorization grids IPS and PPS, the approach adopted here suitably accounts for the fact that there are various cases in which a given object axis is not identified by a single parameter but by a combination of parameters from both grids. Such combinations of DA parameters occur on the Semantic Level as well as on the Conceptual Level. Here are a few examples for each:

Semantically, the English adjective *tall* comprises a combination of MAX and VERT, though not as a (symmetric) conjunction. The fact that the antonym of *tall* is *short* (not *low*) suggests that MAX and VERT are combined in such a way that the axis referred to by *tall* is identified as the object's maximal axis which is furthermore specified as being aligned to the Vertical.

Conceptually, the Object Schemata for "tree" or "tower" contain as primary (i.e. as defining) entry the complex *max-vert* , which indicates (i) that the objects at issue have a canonical orientation regarding verticality, (ii) that this canonical verticality is bound to the objects' maximal axis.

Besides occurring as conceptually fixed, combinations of parameters can also result from contextual specification. So the primary OS for "pole" contains the entry *max* which suffices to assign things like *the pole is 3 m long* a regular interpretation. The interpretation of *the pole is 3m tall / high*, however, provides the OS for "pole" with the contextually induced verticality specification which also results in a complex entry *max-vert.* This is how a gestalt property (*max*) is turned into a position property (*max-vert*). The fact that specification the other way round (that is, turning position into gestalt properties) is excluded adds further evidence to the analysis of *tall* sketched above and, in general, shows that the relation between IPS and PPS is an asymmetric one.

Now, the claim is this: both the full range of possible objects in OS and the scope of admissible dimensional designations and positional variations of spatial objects (irrespective of being primary or contextually induced) are determined by a small set of compatibility conditions that specify which axial properties may combine. The details are given in Lang 1989, here I will only list the results. Given the interpretation of the elements of DAP in 1.3, we are left with 14 in three groups :

(9) single parameters:
MAX, VERT, OBS, ACROSS, DIST, SUB

(10) admissible combinations (based on compatible axial properties):
MAX-VERT, ACROSS-MAX, MAX-OBS, VERT-OBS ($\measuredangle$ 180°)

(11) inadmissible combinations (due to incompatible axial properties):
*MAX-SUB, *DIST-SUB, *OBS-SUB, *OBS-VERT ($\measuredangle$ 0°)

Note that a combination of parameters - as illustrated with *tall* above - is not a mere conjunction but a more structured complex made up of a basic parameter (left part) and a specificatory one (right part). The combination VERT-OBS in (10) is reserved for concave objects that are canon-

ically aligned to the Vertical and that are specified for a canonic ("river") or a contextually induced perspectivization ("pot") that refers to the same axis but in the opposite direction. The combinations *MAX-SUB,*DIST-SUB,*OBS-SUB in (11) are excluded due to perceptually incompatible axial properties - cf. 1.3 above. The combination OBS-VERT, where the Vertical and the Observer Axis run in the same direction (at 0°), is perceptually quite conceivable but, interestingly enough, does not constitute a conceptually relevant parameter. [4]

To sum up: the claims concerning admissible single and combined DA parameters in (9) - (11) sort out 10 out of 14 and thus lay the ground on which we might now look for universals.

2. Some semantic and lexical universals in the realm of DA

2.1 What are semantic universals supposed to be?

Let me start with a few assumptions on what specific theoretic contributions we should expect from semantic universals in comparison with universals in phonology, morphology, or syntax. Given the two-level approach to meaning worked out in the literature quoted above (Bierwisch, Lang (eds.)(1989) and much subsequent work), I take the status of semantic universals to be that given in (12) and their role in linguistic theory to be determined by the requirements posited in (13) - (16):

(12) Semantic universals are statements on how semantic primes and combinations of primes are encoded into lexically categorized, morpho-syntactically specifiable, hence compositionally suitable building blocks out of which phrasal and sentential structures are formed.

Put in terms of a modular view on grammar, (12) might also be reformulated to the effect that semantic universals are statements on the interface between lexicon-based grammatical structures and conceptual structures (knowledge stored in memory). The attribute ' lexicon-based' is due to the view that meaning in language is necessarily linked with lexical items, that is, with those linguistic units that in a specific way integrate phonological, morphological, syntactic, and semantic information and thus - by projecting this information onto combinatorial structure - constitute the basic elements for grammatical structure formation. In this paper, we will focus on the underlined part of (12) and examine it w.r.t. some requirements that are imposed on semantic universals.

(13) Semantic universals are expected to contribute to clarifying the difference between, as well as the interaction of, linguistically encoded meaning and extralinguistic context information;

(14) Semantic universals are to serve as basis on which (a) notions like ambiguity, polysemy, and unspecifiedness can be distinguished, (b) the various types of inferences can be explained;

---

[4] There are data which prove that this kind of doubly determined axis identification is not utilized semantically. As piece of evidence let me quote the following documented example from a TV report on Cape Canaveral:
(i) *The rocket rose into height and disappeared in the depth of space*
The fact that the visible path of the rocket covers one continuous segment (simultaneously determined by the Vertical and the Observer axis running equidirectionally) cannot be semantically accommodated in one dimension term. Instead, the semantic structure of dimension terms necessitates a way of designating the path of the rocket which construes the relevant projections as concatenated, the point of linkage being marked by a shift in the reference system (from PPS to the orbit).

(15) Semantic universals that state which (combination of) primes is lexicalized, that is, packaged into categorized lexical items and this way put into grammatical structure formation, should draw on independent explanations as far as possible.

With these preliminaries and the definitions of the elements of DAP in mind, we may now move on to formulating some tentative universals under the follwing heading:

2.2 What of DAP is lexicalized ?

Given the decisive role of the compatibility conditions discussed in 1.4, especially the assumptions posited in (9) - (11), there is a universal constraint which suggests itself as it follows from (11). We can put it in two equivalent versions:

(16) Only admissible combinations of elements from DAP are lexicalized (=U-1)

There are no lexical items covering simultaneous reference to axes that would be identifiable by MAX-SUB or DIST-SUB or OBS-SUB or OBS-VERT. (=U-1')

In view of the fact that the supposed combinations MAX-SUB, DIST-SUB, and OBS-SUB would embody perceptually incompatible cues (regarding OBS-VERT cf. FN.3), (16) is not a surprising result. It nevertheless is worth being mentioned as a it illustrates what the requirement to look for independent explanations as posited in (15) is assumed to mean. Furthermore, (16) has some direct implications for lexicalization. It predicts that parameters that draw on mutually incompatible axial properties are lexicalized separately, that is:

(17) If MAX, SUB, DIST, OBS are included in a lexical field of DA, the terms drawing on any subset of them are lexically distinct.

The next two tentative universals draw on the prominence of the axes determined by IPS and PPS and are formulated a positive hypotheses. The first one again in two versions:

(18) The most prominent axes of both IPS and PPS are lexicalized separately (=U-2)

In a lexical field of DA, there are at least distinct items for MAX and SUB as well as for VERT and OBS (=U-2')

Unlike (17), which is a corollary of (16), the statement in (18) is an independent one. It claims separate lexicalization for MAX (from IPS) and VERT and OBS (from PPS) and ACROSS (possibly from either) despite the fact that the axial properties to which these parameters refer to are pairwise compatible - as shown in the admissible combinations in (10). (18) is a plausible tribute to saliency as a source for lexicalization which, of course, has to be proved by massive empirical evidence yet. It is a claim about minimal distinctness in the lexical field of DA, and hence is not at variance with e.g. English having items covering VERT (*high - low*), MAX (*long - short*), and MAX-VERT (*tall*) separately, nor with the possibility that there may be languages that encode OBS (at $\measuredangle$ 90°to VERT) and OBS-VERT (= OBS at $\measuredangle$ 180°to VERT) into distinct lexical items.

The following is a tentative universal concerning the relation between conceptual salience and lexical primes (a lexical prime being a lexical item that is neither a compound nor a derivative):

(19) DA terms drawing on MAX, SUB, VERT, OBS are rendered by lexical primes. (= U-3)

Well, in all of the 15 or so languages I have examined so far (besides the major Germanic, Romance, and Slavonic languages my data base includes Turkish, Korean, Mandarin and Khmer), the field of DA belongs to the core lexicon, so the claim in (19) may be weak. It is meant as an attempt to correlate basic perceptual contrasts with preferences for lexical primeness. I hasten to add for all of (16) - (19) that, though I did not find a counterexample to any of them, all three need further confirmation by massive empirical data. So much for tentative universals of what is (minimally) lexicalized within the field of DA. The scope thus defined leaves room for variations in the internal structure of the lexical field of DA. This is what we will take up now.

3. Typological variation in the structure of the lexical field of DA: the basics

3.1 What are the basics ?

Taking stock of the inventory of single and admissible combined DAP with the help of Fig. 3. below, we observe that - disregarding polar antonyms - there are at least 10 candidates for being lexicalized as basic dimension terms. This number of potential DA terms is much larger than the number of basic dimension terms we observe in most known languages, which ranges between 5 to 8 (again disregarding antonyms). Hence we face the problem of what lexical items cover what subsets of the DAP inventory shown below.

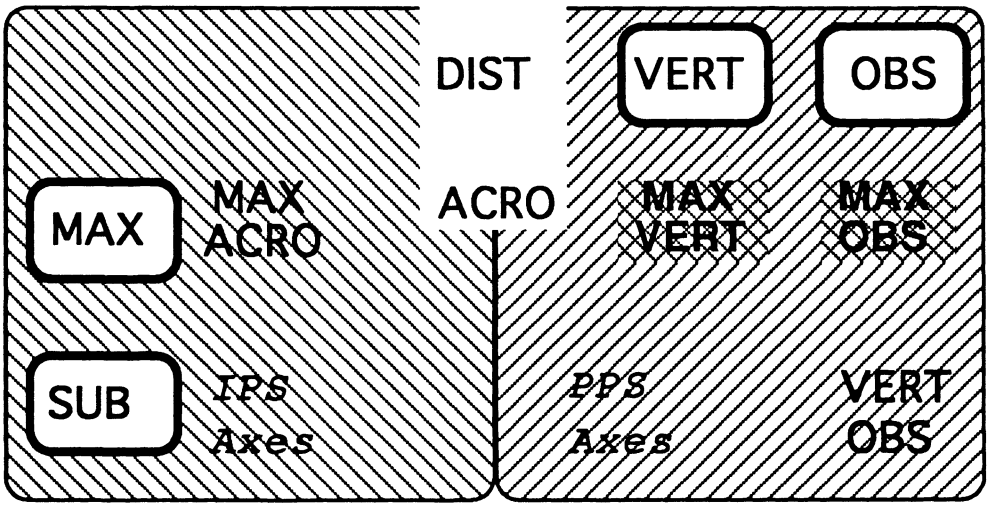


Fig. 3. IPS & PPS Sources of DAP.

As a consequence of this inevitable few-many mapping from lexical items to DAP, we should expect various possible partitions of the set of DAP as regards the scope of lexical coverage. Languages actually vary as to these partitions, and going in search of recurrent patterns and principles explaining them we enter the field of typology.

Taking for granted that, following (19), the two salient DAP of each grid (i.e. MAX, SUB, VERT, OBS - bold encircled in Fig. 3) are rendered by lexical primes, the scope of varying partitions is supposedly confined to the rest. If so, what determines the boundaries of a given partition?

Presumably, it is not the boundary separating IPS and PPS as we observe (indicated by the hatching) two areas which blur such a simple division:

(a) the combinations MAX-VERT and MAX-OBS embody parameters from both IPS and PPS. As there are lexical primes that cover such a combination (recall *tall*), the boundary between IPS and PPS hence does not necessarily serve as a demarcation line in the lexical field;

(b) the parameters ACROSS and DIST are somehow indeterminate between IPS and PPS (cf. 1.2 and 1.3 above), which does not make them clear-cut boundary posts either.

Given that dimension terms rest on conditons for identifying object axes, a cross-linguistic examination of the lexical field of DA suggests that the assignment of dimensions to objects inter alia follows two different **strategies** (20), (21) and an invariant principle (22) which in turn seem to determine the partition of the lexical field of DA .

(20) proportion-based strategy: the relative size of an object's axial extensions **a**, **b**, say **a** > **b**, determines the assignments of dimensions to objects.

(21) observer-based strategy: the condition whether or not **a** = OBS, **b** = ACROSS to OBS (or vice versa) determines the assignment of dimensions to objects.

Independent of (21) and (22), the fact that the Vertical is the domint axis of our spatial orientation (cf. (5) in 1.1) lays the ground for another invariantly observable principle which reads:

(22) the Vertical prevails: if a significant (i.e. non-minimal) object axis **a** coincides with the Vertical of PPS, then this coincidence will determine what dimension **a** is assigned.

To illustrate (20) - (22), imagine a writing desk in normal position sized **a** = 2m, **b** = 1m, **c** = 0.80m. Whatever the proportion of its axial extensions, (22) will pick out its canonic top-bottom extension, say **c,** and reserve it for being identified by a term covering VERT, e.g. *It is 0.80 m high. / in height.* Now, if the remaining extensions **a**, **b** are described by *It is 2m long / in length and 1 m wide / in width,* (20) applies. If the same extensions are described by *It is 2m wide / in width and 1 m deep / in depth,* (21) applies. Actually, English can make use of both strategies (which - as we will see below - is a typologically relevant feature) though not simultaneously in one and the same construction: **It is 2m wide / in width and 1 m wide / in width* is clearly out. The unacceptability of such expressions leads us to a constraint which - being a corrollary of some more general conditions on wellformed coordinate structures - can be narrowed down regarding DA to this:

(23) Uniqueness constraint: in an instance of identifying distinct axial extensions **a, b, c** of some object x, a dimension term t may apply only once.

With (20) - (23) in mind we are prepared to approach the typology of lexicalization patterns.

3.2 How languages partition the lexical field of DA - a typology

Taking (20) and (21) as key factors in determining possible partitions of the lexical field of DA and examining DA data from a sample of 15 languages (a selection of which will be exemplified below) allows us to reach the following conclusions. Typologically, languages differ as to whether or not that subset of the lexical field of DA which covers ACROSS

(24) (a) is clearly determined by the proportion-based strategy (e.g. Mandarin Chinese) or
(b) is clearly determined by the observer-based strategy (e.g. Korean) or
(c) is determined by a conflation of both strategies (German, English, French, Russian).

Note that (24)(a-b) allows for both strategies to apply for disjoint subsets of the same lexical field, the crucial point is whether or not both are conflated on the same lexical items. Moreover, the languages in (24)(c) can be ordered or scaled as to the relative share of either strategy they involve. Thus - as will be shown immediately - German is number 1 on the scale of observer-basedness, Russian is number 1 on the scale of proportion-basedness, English and French are in between.

In the following, I will illustrate (24) by data gathered with the help of two sets of elicitation tests with native informants. I will start with the conflated type (24)(c) as it allows to introduce the experimental design on the basis of more familiar languages and then move on to show in which way each of the other types (24)(a) and (24)(b) significantly differs from English or German.

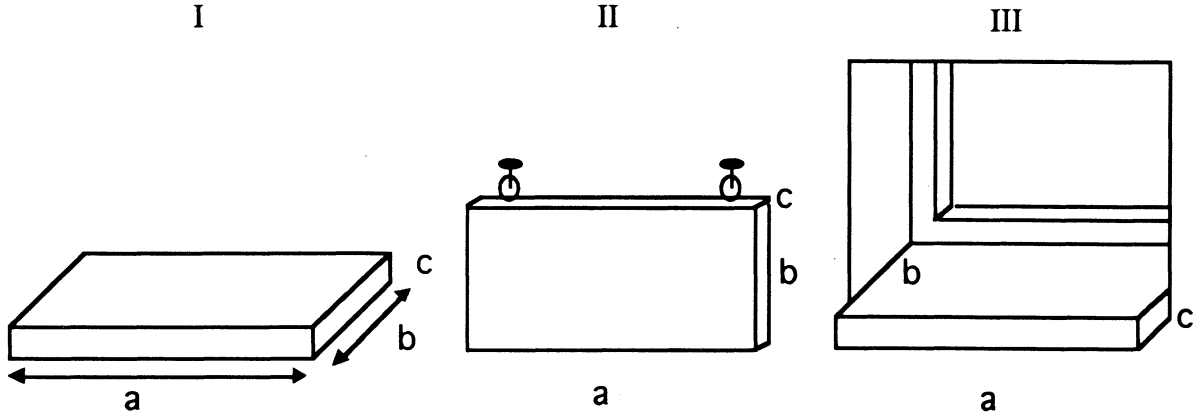
Note that the search for typologically relevant lexicalization patterns is made within the scope of possible variations that is delimited by the (tentative) universals (16) - 19) stated in 2.2. Hence, any typological assumption we are going to formulate is likewise subject to the requirements (13) - (15) posited in 2.1. So, as regards the lexical coverage of DA terms, the elicitation tests should be arranged in such a way that their foreseeable results are relevant (a) to clarifying the relation between encoded meaning and context information, (b) to deciding on ambiguity, polysemy etc., (c) to accounting for various types of inferences. I will comment on these aspects while presenting the tests. The elicitation tests reported on this paper are tasks of naming object extensions.[5]

3.3 ACROSSing the board in English and German

At first I will report on a tried and tested naming task the results of which have proved to be especially telling as regards the methodological requirements repeated above. Subjects were presented a picture showing a board with constant size in three different spatial settings (I - III) -

[5] The data bases underlying the analyses presented in Lang 1989, Lang 1990a,b were obtained with the help of a larger set of elicitation tests which include tasks like object guessing (*x is wide, deep, and high. What might x be* ?) and acting out tasks where subjects were given an object of fixed size, say, a book or a brick, and were asked to position the object according to its possible description by sentences like *The brick is 24cm long, 11cm high, and 7cm wide* . These tests are not discussed in the present paper as they were made only for German and English so far. It goes without saying that they deserve to be carried out for other languages — provided the approach advocated here should prove to be fruitful.

see. Fig. 4. below. Subjects were asked to name the object's axial extensions, which were labeled by the letters **a**, **b**, **c**. Subjects were given hints on plausible measures of the object, e.g. **a** = 1m, **b** = 30cm, **c** = 3cm. For German and English, which in this case do not differ, the results are those listed in (25) and (26) under the respective settings shown in **Fig. 4.**



	I	II	III
(25)	a = lang b = breit c = dick	a = breit b = hoch c = dick	a = breit b = tief c = dick
(26)	a = long b = wide c = thick	a = wide b = high c = thick	a = wide b = deep c = thick

Well, that extension **c** is constantly labeled *dick / thick* is not surprising in view of the claims about SUB (which is encoded by *dick / thick*) made in section 1 above: (9) states that SUB does not enter combinations, (17) claims that SUB is lexicalized distinctly from MAX, OBS, DIST, (19) claims SUB to be packaged into terms that occur as lexical primes.

With extensions **a** and **b**, however, we observe correlative changes of DA terms which obviously (i) centre on how the stop-gap parameter ACROSS is lexicalized (or, put the other way round, what is semantically covered by the terms *breit / wide*); (ii) depend on the spatial settings the board is in. Now, what do (i) and (ii) tell us about the lexical coverage of the terms *breit / wide* ?

As for (i), there are two points to be made. The first point is that - given the 10 admissible (combinations of) parameters in (9)-(10) and in Fig. 3 - *breit /wide* cover the single parameter ACROSS (25)(I) but also the combination MAX-ACROSS (25)(II - III). Looking for the conditions on which these assignments take place leads us immediately to the second point.

Recall that the parameter ACROSS serves as a stop-gap in PPS and/or IPS as it lacks independent perceptual support, which in turn implies that ACROSS comes into play only as dependent on other parameters which do draw on independent perceptual support. This dependency is what defines the respective place of the terms that cover ACROSS in the lexical field of DA. We will see that the typology of languages outlined in (24)(a - c) essentially rests on what strategy takes care of ACROSS and in what way, that is, the proportion-based or the observer-based in disjoint subsets of the lexical field or both conflated onto one subset, which leads us back to German and English.

Now, given that there is no independent defining spatial property according to which *breit / wide* are assigned to physical objects, an object extension **d** to which *breit / wide* are assigned is determined in relation to some other object extension **d'**, where **d'** is identifiable independently. Let's try then to interpret the data in (25/26) against the background of the claims made so far.

In setting I, the assignments **a** = *lang /long*, **b** = *breit /wide* follow the proportion-based strategy (20), that is, ACROSS is determined in relation to some other axis **d'** which is identified as the maximal (i.e. **d'** = **a** = MAX) by *lang / long*. The occurrence of the latter is, as (19) predicts, reliably indicative of maximality being the defining property of the given extension.

In setting II, the assignment **b** = *hoch /high* follows from (22) - "the Vertical prevails" - and as such yields an independently identified axis onto which ACROSS can be hooked (i.e. **d'** = **b** = VERT), thus **a** = *breit / wide.* But what about the underlying strategy? If **a** in setting II were assigned proportion-based, it would be labeled *lang / long*, which is, however, unacceptable in both languages if paired with **b** = *hoch / high* . Hence, the assignment **a** = *breit /wide* in setting II (where it is paired with **b** = *hoch / high*) is due to the observer-based strategy as well. This claim is supported by the fact that normally the axes of PPS are orthogonal to each other (OBS at $\measuredangle$ 90° to VERT and ACROSS at $\measuredangle$ 90° to OBS) as has been claimed in (6) and (7) above.

In setting III, the assignments **a** = *breit / wide*, **b** = *tief / deep* follow the observer-based strategy (21), that is, ACROSS is determined in relation to some other axis **d'** which is identified as being aligned to the Observer axis (i.e. **d'** = **b** = OBS) by *tief / deep*. The occurrence of the *tief / deep* is, as follows again from (19), a reliable indicator of observer-basedness.

Concerning the aspects of ACROSS noted in (i) above, we are now ready to state the relevant partition of the lexical field of DA in German and English. The complete patterns of DA terms, which also show differences between the two languages beyond (27), will be given in 4.3 below.

(27) If *breit /wide* do not cover MAX but are assigned relative to **d'** = MAX, they are assigned proportion-based, otherwise they are assigned oberver-based.

Next, in order to take up (ii), let's take a look at the part played by the spatial settings in (25/26). In (I) the board is conceived as a freely movable object, hence the dimensional terms assigned to it refer to its inherent gestalt properties which allows the proportion-based strategy to apply. Note that due to referring to the object as such, the assignments made in (I) will also hold for settings (II) and (III), though not vice versa. This fact is important for the inferences to be accounted for. In (II) the board can be conceived as having undergone an orientation towards the Vertical due to which it can be assigned position properties (to be hanging, to have a height etc.). This is what the dimension terms **b** = *hoch/high*, **a** = *breit / wide* in (II) refer to.

Finally, in (III) the board can be conceived as part of the window niche such that it is assigned position properties it inherits from the surrounding macro-object, i.e the depth and width of the niche are transferred to the board thus specifying its inherent properties. This is what explains the assignments **b** = *tief /deep* , **a** = *breit / wide* in setting (III).

Having discussed the test so far, it is time for a few remarks assessing its heuristic value. Methodologically, there are at least three aspects that make this test a powerful tool in DA research.

First, it shows a board of constant size in three settings that reflect the board's increasing integration into the spatial environment. In fact, the board as shown in Fig. 4 can be claimed to be distinctly conceptualized in the three settings[6], which, in turn, will reveal to what extent the DA terms used in each of the settings are context-dependent or, put the other way round, what the DA terms used in (II) and (III) induce as contextual specification.

Based on this, the test yields an interesting means to assess the equivalence of situational and linguistic contextual information. The conditions governing the assignment of *breit / wide* that were illustrated in Fig.4 (I - III) by means of the non-verbal contextual settings can just as well be illustrated by means of sentences like (28)(I - III) which, drawing on the Uniqueness constraint in (23), provide exactly the same contextual information on **d'** that is needed to assign *breit / wide* .

(28) Das Brett ist lang und breit genug, aber zu dünn — [**d'** = MAX, **b** = ACROSS
The board is long and wide enough but too thin — as in (25/26) (I)]

Das Brett ist breit und hoch genug, aber zu dünn — [**d'** = VERT, **a** = MAX-ACROSS
The board is wide and high enough but too thin — as in (25/26) (II)]

Das Brett ist breit und tief genug, aber zu dünn — [**d'** = OBS, **a** = MAX-ACROSS
The board is wide and deep enough but too thin — as in (25/26) (III)]

We thus get an instrument that allows to account for the impact of non-verbal context information to the extent that it can be captured by the mutual determination conjoined DA terms impose upon each other within a sentence. (By the way, this equivalence of situational and linguistic contextual information can only be accounted for by representations on the Conceptual Level (cf. Fig. 1), which provides a strong argument in favour of the two-level approach mentioned in section 1.)

Second, the test provides us with a clear-cut example by means of which we can check the inferential relations holding between gestalt and position properties of spatial objects. Note the valid inferences in (29). The sentences contain measure phrases in order to secure the constancy of the object extensions at issue.

(29) (a) The board is 1 m wide, 30 cm deep → The board is 1 m long, 30 cm wide
(b) The board is 1 m wide, 30 cm high → The board is 1 m long, 30 cm wide
(c) The board is 1 m long, 30 cm wide ↛ The board is 1 m wide, 30 cm high / deep

This proves that an object's inherent gestalt properties can be validly inferred from its contextually induced position properties but - as witnessed by (29)(c) - not the other way round. The general pattern underlying these inferences is that of de-specification (see Lang, Carstensen, Simmons 1991).

---

6 The fact that in many languages the board in (III) has a special name pointing to its spatial integratedness might be taken as an additional hint in that direction. So unlike the German compound *Fensterbrett* , which keeps *Brett* as head, the English compounds *window-sill, windowledge* have heads that are spatial meronyms; Russian *podokonnik* is derived from *pod* [under] *okno* [window].

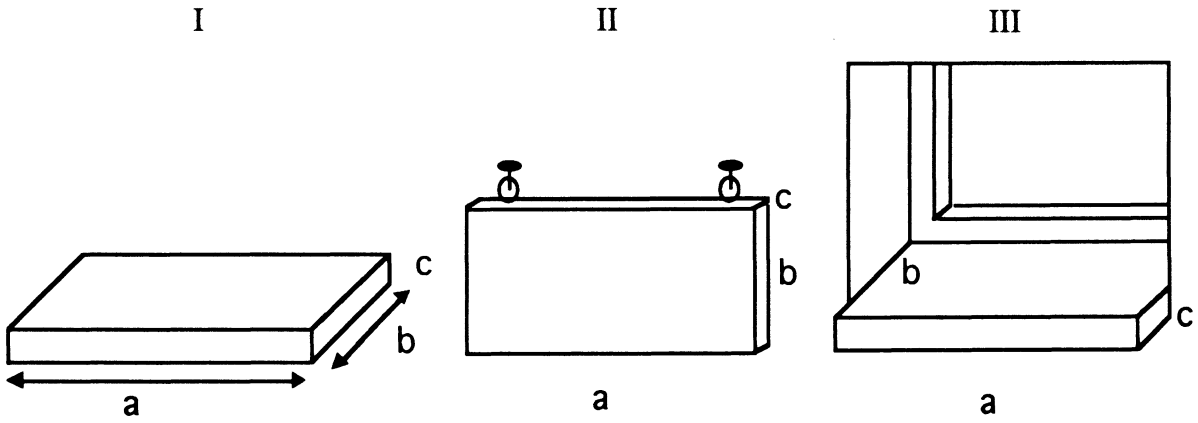
Third, if linked with the typological assumptions in (24), the test provides us with a useful diagnostic for ambiguity. Recall that German and English are of the mixed type regarding the division of labour between proportion-based and observer-based assignments as spelled out for *breit / wide* in (27). Given this we will not only expect, and empirically confirm, that isolated from the context a sentence like (30) is ambiguous as to what extension *breit / wide* refer to,

(30) Das Brett ist 50 cm breit / The board is 0.5 m wide / in width

we can now also exactly determine the range and the source of the different assignments the terms *breit / wide* are able to cover. Moreover, if the ambiguity of *breit / wide* in (30) is an outcome of the fact that German and English are of the mixed type, we would expect the translation of (30) into a language belonging to the "pure" type, say Chinese or Korean, should turn out to be non-ambiguous in this respect. In fact they are - as we shall see in the sequel.

3.4 ACROSSing the board in Chinese, Korean, and Russian

Presenting the naming task in Fig. 3 to speakers of Mandarin Chinese reveals that this language - as regards the lexical encoding of ACROSS - is exclusively proportion-based.[7]



(31)

I	II	III
a = cháng	a = cháng/ *kuān	a = cháng/ *kuān
b = kuān	b = gāo	b = kuān / *shēn
c = hòu	c = hòu	c = hòu

As distinct from (25/26), we observe here that extension **a** is constantly labeled by the term *cháng* that encodes exlusively MAX (as do *long / lang*), while **b** in setting I and III is reserved for *kuān* which is strictly confined to cover ACROSS in relation to **d'** = MAX, and hence is unavailable for **a** in (31) on principle (indicated by * *kuān*).

In setting II, the assignment **b** = *gāo* ['high'] follows from (22) - "the Vertical prevails", which does not interfere with the proportion-based assignment **a** = *cháng/ * kuān* in setting II. If the hanging board is revolved by 90°, the assignments are **a** = *gāo* ['high'], **b** = *kuān / * cháng*, which again confirms (22) as an independent principle.

---

[7] The data shown in (31) confirm observations on Chinese extension terms by Zubin & Choi (1984) and Li (1988), two stimulating papers from which I profited much, not the least because they adhere to a different framework.

The fact that **b** = *shēn* [' deep'] is excluded in any of the settings yields another piece of evidence that Mandarin does not have an observer-based strategy at its disposal. Of course, as predicted by (18) and (19), Mandarin does have a primary term for depth viz. *shēn* ['deep']. However, *shēn* does not at all interact with the conditions of encoding ACROSS but is selectionally restricted to concave or hollow objects. This selectional restriction on *shēn* is again indicative of proportion-basedness (see section 4.1 below.)

Finally, that extension **c** is constantly labeled by one and the same term is not surprising, we would expect this from the comments made on SUB with (25/26). But there is an interesting addition to be noted: unlike German or English, Mandarin Chinese makes subtle distinctions in the lexical items that cover the parameter SUB, so *hòu* is selectionally restricted to identifying a non-maximal disintegrated third axis of an object (i.e. to the first part of the definition of SUB given in 1.3). The issue of lexical granularity of DA terms will be taken up again in section 4 below.

In sum, all observations we can squeeze out of the data in (31) testify to the proportion-based strategy we assumed to be a typological feature of Mandarin. Accordingly, the non-ambiguity prediction regarding *kuān* is borne out: while *breit /wide* in (30) are ambiguous, *kuān* in (32) is not.

(32) Zhè kuài mùbǎn kuān wǔ-shí límǐ
this CL board wide 50 cm
This board is 50 cm wide/in width [i.e. in its secondary extension !]

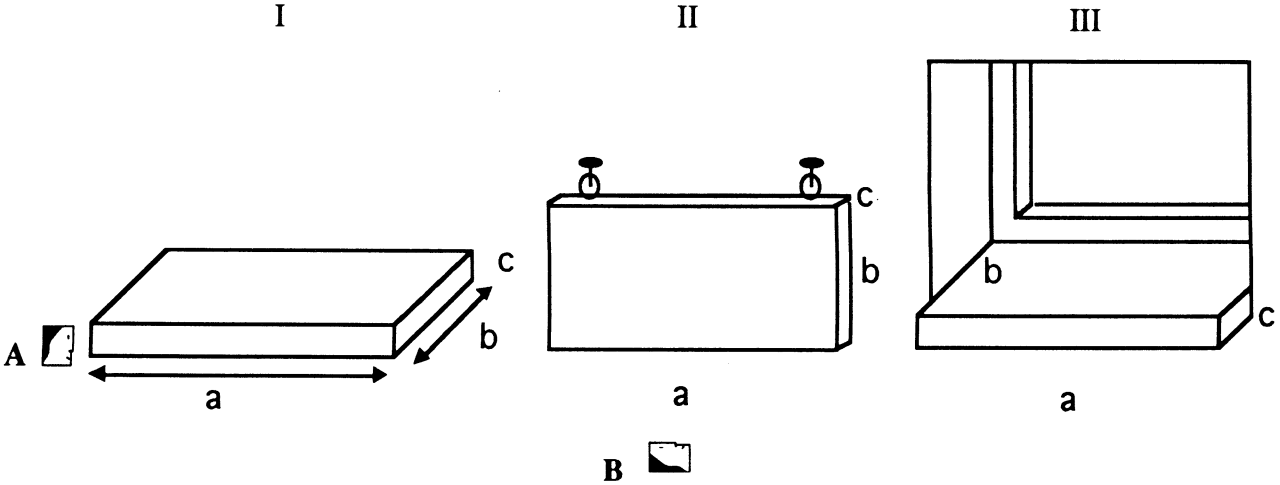
Likewise, the inferences (31) allows for are entirely different from the ones (25/26) do. In order to reveal the semantic consequences of proportion-basedness, I will repeat the English examples of (29) and contrast them with their non-existent equivalents in Mandarin (for brevity' s sake I insert European style measure phrases):

(29) (a) The board is 1 m wide, 30 cm deep → The board is 1 m long, 30 cm wide
(b) The board is 1 m wide, 30 cm high → The board is 1 m long, 30 cm wide
(c) The board is 1 m long, 30 cm wide ↛ The board is 1 m wide, 30 cm high / deep

(33) (a) Zhè kuài mùbǎn kuān 1m, *shēn 30 cm ↛ Zhè kuài mùbǎn cháng 1m, kuān 30 cm
(b) Zhè kuài mùbǎn *kuān 1m, gāo 30 cm ↛ Zhè kuài mùbǎn cháng 1m, kuān 30 cm
(c) Zhè kuài mùbǎn cháng 1m, gāo 30 cm → Zhè kuài mùbǎn cháng 1m, kuān 30 cm

As a matter of fact, the inference patterns in (29) and (33) turn out to be nearly complementary. First note that strictly proportion-based Mandarin cannot provide literal equivalents of the antecedent sentences in (29)(a, b). So it simply lacks the premises needed for this type of inference. Recall that this type of inference draws on the de-specification of oberver-based specifications — and this is what Mandarin does not provide. Second, while Mandarin does not have an analogue of (29)(c) either, it does allow for an inference like (33)(c), which draws on the de-specification of the verticality feature contextually induced according to (22).

Let's now turn to Korean, which works in the opposite way — at least to the extent that oberver-basedness is concerned. In fact, the lexical field of DA terms in Korean (which I quote here in their basic nominal form) also contains a subset of terms (*kili* ['long'], *phok* ['wide']) that resemble Chinese *cháng* , *kuān* in being proportion-based items for an object's maximal and secondary (dis-integrated) axis, respectively.[8] But Korean DA terms contain another subset which is clearly observer-based in that, following Zubin & Choi (1984: 337), the "spatial terms *kalo* and *selo* [...] pick out the edges of a surface which are across and in line with the observer's visual field, respect-ively, with no regard for the relative extension of theses edges." Rephrased in our terminology: these terms are correlatively and context-dependently assigned to non-minimal axes in such a way that *kalo* is reserved for covering ACROSS in relation to **d'** = OBS, *selo* for covering OBS in relation to **d"** = ACROSS. Hence, in order to put the test in Fig. 3 to work, we have to add hints on the respective positions taken by the observer - indicated by the faces and marked by A or B. The proportion-based versions are listed under C.



		I	II	III
(34)	(A)	a = selo b = kalo c = khuki	a = selo b = nophi/kalo c = khuki	a = selo b = kalo c = khuki
	(B)	a = kalo b = selo c = khuki	a = kalo b = selo/nophi c = khuki	a = kalo b = selo c = khuki
	(C)	a = kili b = phok c = khuki	a = kili b = nophi c = khuki	a = kili b = phok/kiphi c = khuki

Extension **c**, as we would expect, is constantly labeled by the term *khuki* ['thick'] covering SUB as defined in 1.3 and does not interfere with taking proportion-based or observer-based strategy. But what is worth noting is the complementary distribution of *kalo* - *selo* onto the extensions **a** and **b** within and between the settings (A) and (B). The additional option for **b** = *nophi* ['high'] in II for all of (A- C) follows independently from (22). The only structural difference of the data in (C) to their likewise proportion-based counterparts in Mandarin is the option for **b** = *kiphi* [' deep'].

---

[8] See also Zubin & Choi (1984). The Korean data in this paper are due to my informant Byong-Rae Ryu (Tübingen), the transcription system for Korean used in the examples below is based on Yale Romanization.

The observation that the two subsets of the lexical field that cover ACROSS are strictly disjoint as concerns source and distribution is furthermore confirmed by the following fact: any conjoined co-occurrence of elements from either subset, that is, any combination of *kalo - selo* and *kili - phok* in a coordinate construction, is ruled out.[9] This homogeneity constraint can be construed as a language-particular tightening of the general Uniqueness Constraint in (23).

Regarding predictable ambiguities, we are faced with two cases as we might have guessed. The proportion-based version of referring to the board's ACROSS axis in (34) is non-ambiguous

(35) Ku-nelphanci-nun phoki il meta ita.
DET.board.TOP width.SUBJ 1 m DECL
"This board is 1 m wide/in width" [i.e. in its secondary extension !]

whereas observer-based versions like (35), if presented without contextual cues for (A) or (B), is ambiguous (or unspecified) as regards reference to extension **a** or **b.**

(36) Ku-nelphanci-nun selo-ka/kalo-ka il meta ita.
DET.board.TOP observer-axis.SUBJ / across axis.SUBJ 1 m DECL
"This board is 1 m wide/in width/long/in length"

Finally, the inferential behaviour of the data in (34) can be easily extrapolated from what has been stated so far: if the absolute measures of the object's extensions are as required, there is room for valid inferences from oberver-based assigned axes to proportion-based ones, that is, from an object's contextually induced position properties to its inherent gestalt properties. Thus, (37) proves to be valid along the lines of de-specification illustrated by (29) in 3.3.

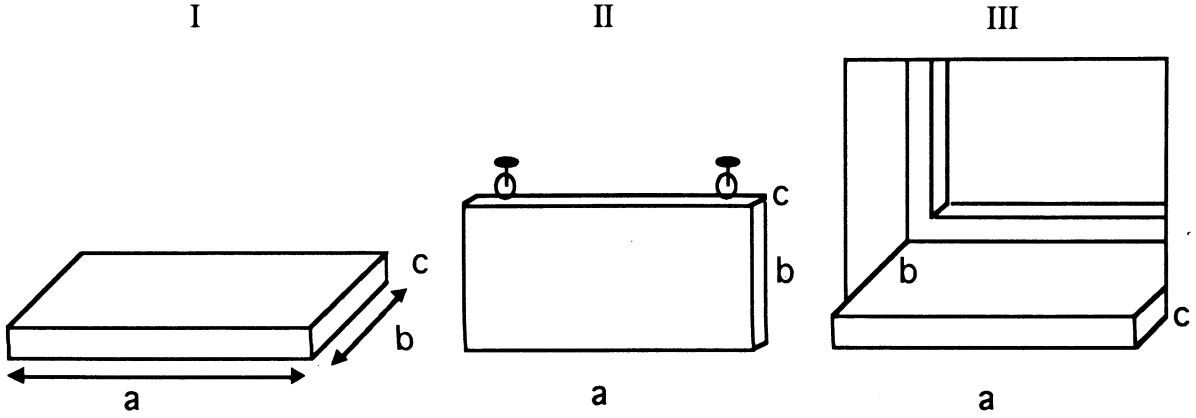
(37) Ku-nelphanci-nun kalo-ka 1 m, selo-ka 0.3 m ita.
→ Ku-nelphanci-nun kili-ka 1 m, phoki 0.3 m ita.
"this board's across axis is 1m, its observer-axis is 0.30 m"
→ "this board is 1m in length, 0.30 in width"

To conclude: as regards the lexical coverage of ACROSS, Mandarin Chinese is confined to a subset of terms that exclusively rest on proportion-based DA, Korean has two disjoint subsets of terms, one of them drawing on proportion-based DA, the other on oberver-based DA, German and English have one subset of terms on which both strategies are conflated. So much for the coarse typology sketched in (24)(a - c).

Let me just add one more example to illustrate the finer-grained sub-typology within the group of "conflating" languages. On the whole, the field of German DA terms shows more features of observer-basedness than of proportion-basedness, in Russian the opposite holds. This can be partially revealed by (38), the Russian terms are given as derived nominals:

---

9 This does not exclude sentences with terms from both subsets occurring in distinct syntactic positions as in:
(i) Ku-nelphanci-nun selo-ka kilta. "The observer-axis of this board is long"
(ii) Ku-nelphanci-nun selo-ka selo-pota kilta. "The observer-axis of this board is longer than its across-axis"



(38)	Russian	a = dlina	a = dlina/shirina	a = dlina/*shirina
		b = shirina	b = vysota	b = shirina/*glubina
		c = tolshchina	c = tolshchina	c = tolshchina
	German			
		a = Länge	a = Breite/?Länge	a = Breite
		b = Breite	b = Höhe	b = Tiefe
		c = Dicke	c = Dicke	c = Dicke

The relevant point is in setting III: while German opts for observer-based assignment, i.e. **b** = *Tiefe* encodes OBS and thus provides a suitable **d'** to which ACROSS in **a** = *Breite* can be hooked on, Russian sticks to the proportion-based assignment **a** = *dlina* ['length'], **b** = *shirina* ['width'], the label **b** = *glubina* ['depth'] is rejected as unacceptable. What does this tell us about the place of Russian on the scale?

First, unlike Mandarin it does not restrict ACROSS to being dependent on **d'** = MAX only, it also allows for a **d'** = VERTas witnessed by the assignments in II. Here **b** = *vysota* ['height'] (which follows independently from (22)) leaves room for either having proportion-based **a** = *dlina* ['length'] or - less preferred - for **a** = *shirina* ['width'].

Second, like in Mandarin but unlike in German and English, in Russian the term encoding OBS, i.e. *glubina*, is selectionally restricted to concave or hollow objects (which a board apparently does not belong to, even if integrated in the window niche in setting III).

So, the range of ACROSS axes to be covered by Russian *shirina* ['width'] is confined to those axes **d** that are hooked on **d'** = MAX or **d'** = VERT whereas German and English have an additional option for **d'** = OBS, the latter being obviously linked to the fact that *Tiefe / depth* are not selectionally restricted to concave or hollow objects but encode OBS without restrictions.

This observation reveals an interesting facet of cross-linguistic variation in DA terms: languages that in accordance with (19) have the same number of lexical primes covering MAX, SUB, VERT, and OBS, respectively, may still differ as to the selectional restrictions imposed on these terms. So, selectional differences should be included in the list of typological parameters.

In the next section we will scrutinize the lexical field of DA terms for further aspects of the typology outlined in (24) while at the same time presenting another useful test setting.

4. Typological variation in the structure of the lexical field of DA: more details

In section 3 we focussed on discussing what strategies take care of covering the parameter ACROSS and on pinpointing the effects of the various options on the structure of the lexical field of DA. Now, there is a lot of evidence that the choice between, or the interaction of, proportion-based and observer-based DA has more effects on determining the lexical coverage of DA terms. Space limitations prevent me from presenting the full particulars that might be adduced to back up this claim. Instead, I will proceed like this: 4.1 presents a list of characteristics of observer-basedness and proportion-basedness, respectively; in 4.2 these criteria will then be illustrated by applying them to data that were obtained from another test setting; 4.3 will summarize the aspects discussed so far by presenting some full lexicalization patterns of DA terms.

4.1 Characteristics of observer-basedness vs. proportion-basedness

The following sets of characteristics were abstracted from a large amount of data. They are listed as informal descriptions of features that are symptomatic of the impact on the lexical field of the observer-based or the proportion-based strategy in DA. To easy later reference, I will number them, L is a variable for languages. To begin with, the impact of the **observer-based strategy** on the internal organization of the field of DA terms in L can be diagnosed from (O-1) - (O-4) below, which - as far as I can see - do not display any intrinsic order.

(39) The term(s) that cover(s) OBS in L determine(s) the use and the range of interpretations of the term(s) that cover(s) ACROSS in L. (= O-1)

(40) The term(s) that cover(s) OBS in L is/are <u>not</u> selectionally restricted to concave or hollow objects. (= O-2)

(41) For a specified class of objects C, L allows for a ' high' - ' deep' alternation in designating some object axis **d**. (= O-3)

(42) In L, verticality assignment can absorb maximality assignment, that is, the parameter-combination MAX-VERT is covered by the terms of L. (= O-4)

In view of the basics discussed above, only (O-3) and (O-4) deserve some comments. The ' high' - ' deep' alternation mentioned in (O-3) occurs when an object's (primary or contextually induced) vertical axis [' height'] is contextually specified as being aligned to the observer-axis. Technically, if in the object's OS an entry *vert* is contextually combined into *vert-obs* thus specifying that the axis **d** referred to as vertical is being looked at in the opposite direction (OBS at $\measuredangle$ 180°to VERT), cf.:

(43) Der Topf ist zu <u>hoch</u>, um ins Regal zu passen, aber nicht <u>tief</u> genug für die Pute
The saucepan is too <u>high</u> to fit into the shelves but not <u>deep</u> enough for the turkey

The delimitation of the class of objects C is dependent on whether or not (O-2) and (O-3) jointly hold, as is the case for German, where the range of the ' high' - ' deep' alternation is rather wide.

(O-4) refers to the alternative left open if (22) ("the Vertical prevails") applies. If the maximal axis of an object happens to be also the vertical (e.g. a corner cupboard of 2m height), L has to decide

(a) whether the term assigned to this axis is reserved for verticality alone, maximality being transferred to the secondary axis and accounted for in a separate assignment or

(b) whether the term assigned to this axis designates verticality while absorbing maximality in such a way that no other axis is available for the term that assigns maximality.

The option (a) is what we observe in Mandarin or Russian, where the cupboard is assigned *vysota* ['height'], and independently *dlina* ['length'] and *shirina* ['width'], whereas option (b) is typical for German and English, yielding the assignments *high* , *wide* (**long*) and *deep* , or even *tall*, *wide*, and *deep*, where in *tall* the absorption of maximality by verticality is explicitely lexicalized.

Next, let's look at the characterics of **proportion-based strategy** being operative. As one might guess, (P-1) will be supplementary, while (P-2), (P-3), (P-4) will be complementary to the symptoms (O-1) - (O-4) and hence self-explaining.

(44) The term(s) that cover(s) MAX in L determine(s) the use and the range of interpretations of the term(s) that cover(s) ACROSS in L. (= P-1)

(45) The term(s) that cover(s) OBS in L is/are selectionally restricted to concave or hollow objects. (= P-2)

(46) If L allows for a ' high' - ' deep' alternation in designating some object axis **d**, then the class of objects C for which it holds is more constrained than in (O-3) (= P-3)

(47) In L, verticality assignment is separated from maximality assignment. (= P-4)

Having the two sets (O-1) - (O-4) and (P-1) - (P-4) at our disposal, we may now take further steps towards a finer-grained specification of typological variation in the lexical field of DA terms. As regards the internal structure of this field, a language L can be scaled by stating which of (O-1) - (O-4) and/or of (P-1) - (P-4) can be proved to apply in L. Moreover, the characteristics expounded above provide us with a suitable means to account for ambiguities and for lexical gaps. This will be shown in the next section where I proudly present another elicitation test.

4.2 The *staircase* case

The experimental design of this naming task is again quite simple. Subjects were presented the picture of a staircase in Fig. 5 below and asked (a) to name the extensions of the first step (shadowed), (b) to name the dimensions of the staircase as a whole. Subjects were hinted to answer twice, while imagining themselves (i) going upstairs, (ii) going downstairs (big arrows).

The test is valuable in several respects: it allows to check the availability in L of the ' high' - 'deep' alternation, it reveals the degree of spatial integration of a step into the macro-object "staircase", and - based on this - it is a diagnostic for inferences that draw on DA inheritance from parts to wholes and vice versa.

4.2.1 Coarse Typology. In order to show how the typology in (24) can be reconstructed in terms of (O-1)-(O-4), (P-1)-(P-4), I will first present data from Mandarin, Russian, German, and Korean. The data are arranged in Fig. 5 in such a way that it reflects the underlying typology.

(48) Mandarin

a = gāo ['high']
b = cháng ['long']
c = kuān ['wide']

(49) Russian

a = vysota ['height']
b = dlina ['length']
c = shirina ['width']

(50) German

a = hoch	a = tief
b = breit	b = breit
c = tief	c = ???

(51) Korean

a = nophi ['height']	a = kiphi ['depth']
b = kalo ['across-axis']	b = kalo ['across-axis']
c = selo ['observer-axis']	c = selo ['observer-axis']

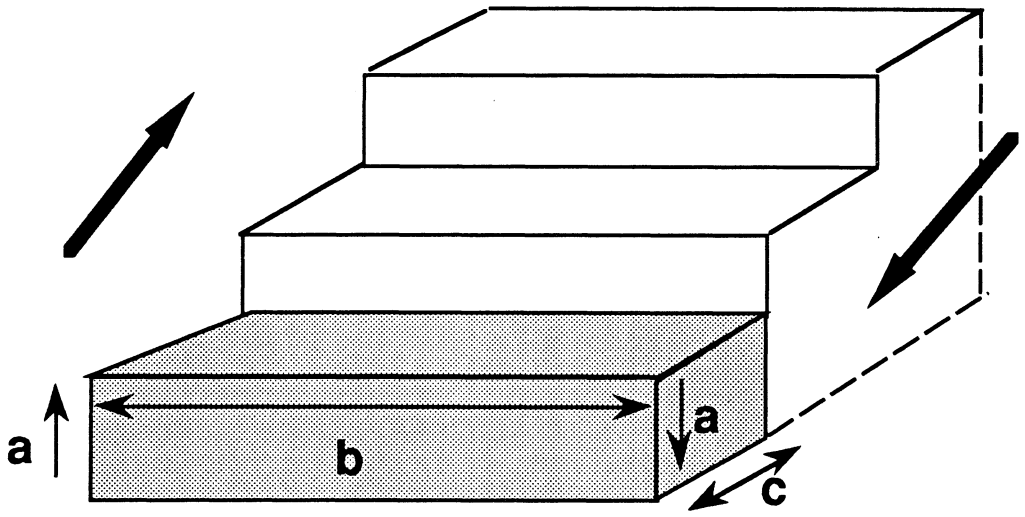


Fig. 5 Naming the extensions of a step of a staircase

Note that (48) marks the only way in which Mandarin can dimensionize a step of a staircase. This provides additional evidence for the claim that Mandarin DA terms are strictly proportion-based. Let's check this by running through the Ps: (P-1) yes, (P-2) yes (a step is not a concave object), (P-3) yes (no 'high' - 'deep' alternation with non-concave objects), (P-4) yes (verticality assignment [*gāo*] is separated from maximality assignment [*cháng*]. So all Ps but no O apply. What is more, in such a clear-cut case of proportion-basedness we would not expect any ambiguities or gaps, and in fact there are none. Russian is similar except that (P-1) is supplemented by (O-1), i.e. the option to assign ACROSS relative to VERT — see (38) II), which may apply in (49), too.

The German assignments in (50) are as expected if assigned going upstairs (left set) but reveal an hitherto undetected gap if assigned going downstairs (right set c = ???). How come? Checking the Os we obtain (O-1) yes, (O-2) yes (both proved by the left set); (O-3) yes (witnessed by **a** = *deep* in the right set), (O-4) yes (not evidenced by Fig. 5 but easy to imagine: if **a** were the maximal the steps would be uncomfortable but the assignments would be the same). So what's wrong?

The gap in (50) is caused by two independent but converging factors. First, as will be recalled, German *tief* covers OBS and OBS-VERT, the latter being the basis for the 'high' - ' deep' alternation we observe in (50), the former causing an interference of (O-1) and (O-3) w.r.t. *tief*. This is the language-particular factor.[10] Second, the Uniqueness Constraint (23) is operative and thus prevents *tief* from occurring twice in (50) despite the fact that it would refer to distinct extensions. This is the universal factor which expectedly prevails.

Finally, (51) shows the observer-based subset of Korean DA terms in two versions. How does this subset score? (O-1) yes (that is the gist of the *selo-kalo* correlation), (O-2) + (O-3) yes (step is in class C), (O-4) no (verticality assignment is separate from maximality assignment). Why are there no ambiguities or gaps? Well, unlike in German, in Korean (O-3) does not interfere with (O-1) as the 'high' - ' deep' alternation [*nophi - kiphi*] is independent of the assignment of *selo-kalo* . So no ambiguity arises nor is there room for a gap.

4.2.2 Finer-grained typology. Another set of staircase data is that in (52) - (55). The point here is the variations we find among languages that are near cognates in i.a. using terms which etymologically share the same common Slavonic origin. Russian, as shown in (48), is high on the scale of proportion-basedness, whereas Polish and Slovak, while using the same adjectival terms, display a tendency towards observer-based assignments — presumably under the influence of German.

(52) Russian

a = vysokij	['high']			
b = dlinnyj	['long']			
c = shirokij	['wide']	/*glubokij	['deep']	

(53) German

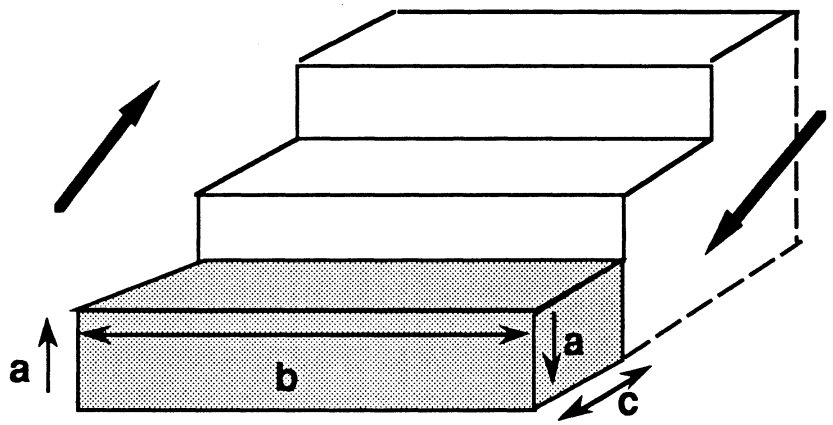
a = hoch	['high']	a = tief	['deep']
b = breit	['wide']	b = breit	['wide']
c = tief	['deep']	c = ???	

(54) Polish

a = wysoki	['high']	a = głęboki	['deep']
b = długi	['long']	b = długi	['long']
c = szeroki	['wide']	c = szeroki	['wide']

(55) Slovak

a = vysoký	['high']		a = hlboký	['deep']
b = dlhý	['long']	/široký ['wide']	b = dlhý	['long']
c = široký	['wide']	/hlboký ['deep']	c = široký	['wide']



[10] I dare say that English works the same way. But as my informants are still quarreling on whether or not Britannia rules the gaps, I shall not quote English step data before being officially notified of the results.

The gap in (50) is caused by two independent but converging factors. First, as will be recalled, German *tief* covers OBS and OBS-VERT, the latter being the basis for the 'high' - ' deep' alternation we observe in (50), the former causing an interference of (O-1) and (O-3) w.r.t. *tief*. This is the language-particular factor.[10] Second, the Uniqueness Constraint (23) is operative and thus prevents *tief* from occurring twice in (50) despite the fact that it would refer to distinct extensions. This is the universal factor which expectedly prevails.

Finally, (51) shows the observer-based subset of Korean DA terms in two versions. How does this subset score? (O-1) yes (that is the gist of the *selo-kalo* correlation), (O-2) + (O-3) yes (step is in class C), (O-4) no (verticality assignment is separate from maximality assignment). Why are there no ambiguities or gaps? Well, unlike in German, in Korean (O-3) does not interfere with (O-1) as the 'high' - ' deep' alternation [*nophi - kiphi*] is independent of the assignment of *selo-kalo* . So no ambiguity arises nor is there room for a gap.

4.2.2 Finer-grained typology. Another set of staircase data is that in (52) - (55). The point here is the variations we find among languages that are near cognates in i.a. using terms which etymologically share the same common Slavonic origin. Russian, as shown in (48), is high on the scale of proportion-basedness, whereas Polish and Slovak, while using the same adjectival terms, display a tendency towards observer-based assignments — presumably under the influence of German.

(52) Russian

a = vysokij	['high']		
b = dlinnyj	['long']		
c = shirokij	['wide']	/*glubokij	['deep']

(53) German

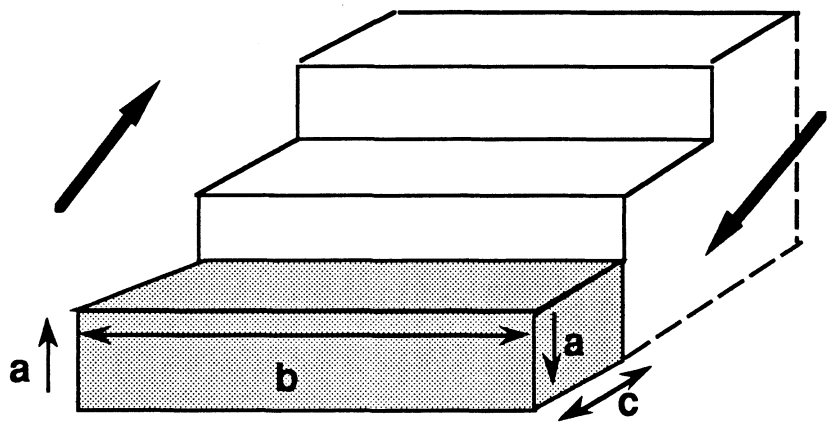
a = hoch	['high']	a = tief	['deep']
b = breit	['wide']	b = breit	['wide']
c = tief	['deep']	c = ???	

(54) Polish

a = wysoki	['high']	a = głęboki	['deep']
b = długi	['long']	b = długi	['long']
c = szeroki	['wide']	c = szeroki	['wide']

(55) Slovak

a = vysoký	['high']	a = hlboký	['deep']
b = dlhý	['long'] /široký ['wide']	b = dlhý	['long']
c = široký	['wide'] /hlboký ['deep']	c = široký	['wide']



[10] I dare say that English works the same way. But as my informants are still quarreling on whether or not Britannia rules the gaps, I shall not quote English step data before being officially notified of the results.

Let's look at how Polish and Slovak stepwise deviate from Russian and approach German. This can best be shown by the following table:

(56)	Mandarin	P-1	P-2	P-3	P-4	
	Russian	P-1	P-2	P-3	P-4	O-1
	Polish	P-1	P-2	O-3	P-4	O-1
	Slovak	P-1	O-2	O-3	P-4	O-1
	German	P-1	O-2	O-3	O-4	O-1

In Mandarin, only Ps apply. In Russian, as we observed in (49), (P-1) - (P-4) apply supplemented by (O-1). Now, (54) shows that Polish has replaced (P-3) with (O-3) - thus allowing for a 'high' - 'deep' alternation (*wysoki* ['high'] – *głęboki* ['deep']), the rest remaining unchanged.[11] Slovak has replaced (P-3) with (O-3) too, but in addition it has replaced (P-2) with (O-2). This is a decisive step towards observer-basedness. Semantically, this is to say that while Russian *glubokij* ['deep'] maintains the selectional restriction to concave or hollow objects, Slovak *hlboký* ['deep'] has loosened it to the effect that the class C of objects to which (O-3) applies resembles that in German.

Note that this is a subtle but interesting case of interference that has gone unnoticed so far. The explanation sketched here seems quite plausible. Despite the fact that in any L I checked so far, DA terms belong to the core lexicon (in this case to the common Slavonic core lexicon), which is normally resistant to lexical or other familiar types of borrowing, they can nevertheless be subject to interference. So, while keeping its place in the common Slavonic core lexicon, the field of Slovak DA terms has undergone subtle changes as to the lexical coverage of DA parameters. The replacement of common Slavonic (P-2) with (O-2) - supposedly under German influence - is a case in point. The subtlety of this change is neatly accounted for by being attributed to the loosening of a selectional restriction of a particular term - as seems to be the case with Slovak *hlboký* ['deep'].

Moreover, also the pragmatic underpinning of this change sounds quite reasonable. Due to geographic, political, and cultural reasons, the West Slavonic languages were and still are in close contact with German. The loosening of the selectional restriction on a term such as *hlboký* ['deep'] is induced by widening the class C of objects to which *hlboký* ['deep'] becomes applicable. This, in turn, is a sort of process that is set going or at least facilitated by the culturally and linguistically intertwined everyday communication that has been charcteristic for this area for centuries.

Unlike sortal constraints, which are a matter to be accounted for on the Conceptual Level, selectional restrictions have to be accounted for on the Semantic Level as they are linked with lexical packaging and hence are subject to language-particular variation (see Lang 1994). So I would like to suggest to enrol selectional restrictions in the list of typologically relevant parameters.

---

[11] As regards the selectional restrictions on *głęboki* ['deep'], Polish seems to be in a transitory state where (P-3) and (O-3) compete with each other. So my informants accepted assigning *głęboki* to e.g. a cupboard "only if it is open".

4.2.3 Inferences. The reconstruction in (56) of the data (52) - (55) is reflected (and hence confirmed) by looking at the sets of inferences in (57) - (59) that draw on the inheritance of dimension assignments from parts to wholes. (57) shows that height if referring to the canonical vertical axis of an object is both part-whole and whole-part inheritable. This is but another aspect of the dominance of the Vertical mentioned in (5) and of the universal principle in (22), so we would not expect any variations across languages as regards the validity of this inference.

(57) (a) The height of the staircase is composed of the heights of the steps
(b) Die Höhe der Treppe ergibt sich aus den Höhen der Stufen [German]
(c) Vysota lestnicy sootvetstvuyet summe vysoty stupenej [Russian]
(d) Wysokość schodów wynika z wysokości schodków [Polish]
(e) Výška schodiště sa skladá zo súčtu výšiek schodov [Slovak]

The inheritability of depth assignments is subject to different conditions. Due to physical unavailability, primary depth assignments - irrespective of covering OBS or OBS-VERT - are hardly inheritable. So if depth enters part-whole inheritance at all, it is confined to cases of OBS-VERT emerging as contextual specification of the vertical. The OBS covered by (53) c = *tief* is not inheritable to a staircase as a whole either. So the staircase talked about in (58) should be one leading downwards into the cellar. With this proviso, the data in (58) seem to prove that depth assignments are part-whole inheritable only in languages where (O-3) applies (hence * for Russian):

(58) (a) The depth of the staircase is composed of the depth of the stairs
(b) Die Tiefe der Treppe ergibt sich aus den Tiefen der Stufen [German]
(c) *Glubina lestnicy sootvetstvuyet summe glubiny stupenej [Russian]
(d) Głębokość schodów wynika z głębokości schodków [Polish]
(e) Hl'bka schodiště sa skladá zo súčtu hl'bok schodov [Slovak]

Things are still different with 'length' (MAX) and 'width' (ACROSS) assignments, which - for the class of objects a staircase belongs to - are not part-whole inheritable. For languages that adhere to proportion-basedness, we will expect an obligatory change of terms, e.g. when stating that the 'length' of the steps equals the 'width' of the staircase. That is what we observe in Russian - cf. (59)(c). In languages like English and German, the observer-based strategy (O-1) - (O-4) prevails due to being applicable to both steps and staircase and thus making a step's width inheritable to the staircase - cf. (59)(a,b). What about the West Slavonic languages in between ?

(59) (a) The width of the steps is equal to the width of the staircase
(b) Die Breite der Stufen ist gleich der Breite der Treppe [German]
(c) Dlina stupenej sootvetstvuyet shirine lestnicy [Russian]
Dlina lestnicy sootvetstvuyet summe shiriny stupenej
*Shirina stupenej sootvetstvuyet shirine lestnicy
(d) Długość schodków odpowiada szerokości schodów [Polish]
Długość schodów odpowiada sumie szerokości schodków
*Szerokość schodków odpowiada szerokości schodów
(e) Dl' žka schodov udáva šírku schodiště [Slovak]
?Šírka schodov udáva šírku schodiště

Towards a revision of the lexical subcategorization features

Renate Steinitz
FAS

1. The history of the subcategorization features [V] and [N]

The classification of the four major lexical categories Verb, Noun, Adjective and Preposition[1] by means of the features [V] and [N] with two values has done its job for about twenty years.

Advantages:

- The minimal set of distinctive features allows the wanted discrimination without redundancies,
- It allows reference to similar cross-categorial properties of lexical items and generalizations of so called natural classes. (HAIDER (1993:34ff) defines e.g. the parametrization in the direction of government over the feature [N], that is nouns and adjectives behave the same in respect to the direction of government in German and so do verbs and prepositions, [+N] and [-N] define natural classes.)

Disadvantages:

- The features themselves are not motivated independently and their interpretation is somewhat ad hoc and differs among the various accounts.
- With respect to markedness theory P [-V,-N] is a maximally unmarked and A [+V,+N] a maximally marked category. This is clearly not a desirable description of the hierarchy of marked lexical categories in any respect (if we compare the status of the various lexical categories in one language or between languages).

The dissatisfaction with the classic inventory is very much in the air. I refer to proposals in HALE/KEYSER (1994) and WUNDERLICH (1994) in some footnotes. The idea of interpreting lexical categories in the frame work of markedness is due to Wunderlich in his cited papers.

2. Properties of linguistic entities as candidates for subcategorization features

In my work on non-verbal predication and related topics in STEINITZ (1988, 1990, 1991, 1993, 1994) I consider the property that is responsible for assignment of a lexical category to a lexical item.

Neither the morphology of inflection can do the work, if we want to have a definition independently of specific languages with rich morphology, that is if we claim a universal foundation of categorization. Nor can the morphology of derivation; otherwise, simplex words would not be mentioned. I think morphological marking is an epiphenomenon of more basic properties. It cannot be the syntactic functions either, because they don't map one to one into the lexical categories.

---

1 For the moment I have to ignore a fifth subcategory, that is real adverbs such as *gern, sehr, glücklicherweise* etc. A subclass of adverbs such as *hier, dort, so, dann,* etc are pro-adverbials and described as PPs in STEINITZ (1969) already, in terms of current generative theory they are maximal projections of P, or intransitive Ps. To establish a fifth category adverb a third feature is necessary, and the problem is then how to interpret the eight possible combinations. Later on I will try to incorporate the results of ZIMMERMANN (1988a, 1988b, 1994) in my framework. Zimmermann introduces a third feature [Adv] to account for subcategories such as adverbia and attributiva tantum. For my present purposes I keep to two features and interpret all adverbs as intransitive prepositions. The compromise is somewhat fishy. In section 10 it becomes disturbing, because HENGEFELD (1992) do not take PPs but only manner adverbials into account, which would hardly be intransitive prepositions.

In the UNITYP model "semantic parameters" such as "time-stable:time-instable", "related:relational"(cf. e.g HOPPER/THOMPSON 1984, BROSCHART 1991) motivate the distinction of nouns and verbs. Such a parameter is thought of as a bipolar scale, a continuum between the two complementary poles with segments occupied by prototypical nouns and verbs respectively.

I do not deny the affinity of certain categories of meaning to certain lexical categories. Verbs typically denote time instable entities, that is, activities. Nouns typically denote time stable entities, that is, objects and so on. But this level of concrete semantics is not of much help, because once established in the framework of grammar the grid of categories holds for the untypical cases as well. Derivational forms mostly change from a typical into a nontypical lexical category. The categories are now grammatically based independently of their lexical meaning. (But see my attempt to incorporate the model of semantic parameters in the grammatically based one in section 9). It is likely that all of the mentioned properties take part in lexical subcategorization, but I think they are epiphenomenal rather than basic.

3. Semantic type: lexical-syntactic categories

I suppose Formal Semantics, the **theory of structure of meaning**, to be the base for a grid, in which all syntactic and lexical categories in natural languages arrange. There are two distinct categories with further subcategorization, cf. CHIERCHIA (1985):

(a) **Basic, simple categories**

These categories are mapped onto the ontological universe:

1. **Propositions** refer to situations, their semantic type is **<t>**, e.g. they have a truth value.
2. **Individual terms** (definite descriptions and proper names) refer to entities (object entities, situation entities etc.); they are of type **<e>**.

The standard correspondence in syntax:

- (Independent declarative) sentences (Inflection phrases **IP**s,) correspond to the semantic type <t>.
- (Definite) Determiner phrases **DP**s correspond to the semantic type **<e>.**

(b) **Derived, complex categories**

These categories are n-place predicates <e, t>, <<e,t>e> etc, and they are mapped into properties, states, activities etc.

The standard correspondence in syntax:

1. The set of **lexical categories** and their projections. In German and the other Indo-european languages these are V, N, A, P. In the Semantic Form (SF) of a lexical entry there is no discrimination between the lexical categories, they are all n-place predicates, cf. the entries in (ii), section 5, for illustration.
2. The set of **functional categories** corresponds to special semantic predicates: **I(nflection)** corresponds to an operator over situations, which takes a predicate, selects a section of its extension and localizes it in time. (I ignore for my purposes here the category Complementizer or other functional categories). **D(eterminer)** corresponds to an operator over entities, which takes a predicate and by selecting a part of its extension specifies its reference in the world.

The categories in (a) are basic, primitive, saturated and referential; and those in (b) are derived, complex, unsaturated and non-referential.

4. The basic assumption

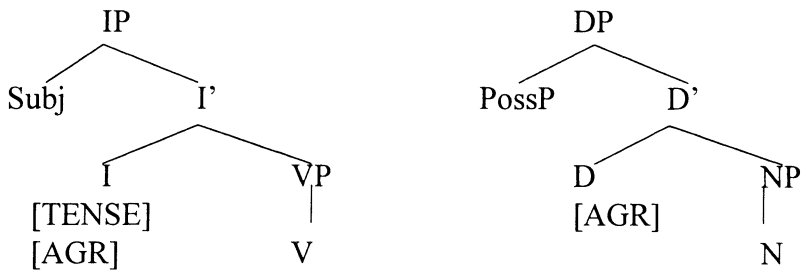
The structures in semantics and syntax are autonomous but systematically related. Every expression in a natural language has to belong to one of the three semantic categories <e>, <t> or a derivation of them. We assume that universally distinct semantic categories are encoded systematically in all languages such that they are distinct at one syntactic level at least. The problem is at which level the distinction has to be fixed.

Referentiality: the basic categories <t> and <e> are both ontological categories, no other category is independently referential. Other categories are referentially dependent predicates. Operators cause the referential anchoring of the predicates, cf. the comments above.

As syntactic correlates to the semantic basic categories <t> and <e> the syntactic phrases IP and DP are the only two categories containing referential arguments. The operation referentially specifying the predicates is **theta-binding**. A functional category I or D as head of IP and DP respectively theta-binds the referential argument of its complement. Theta binding is a selectional constraint, the only complement of I is VP, the only complement of D is NP. The other way round, if there are two distinct lexical categories V and N with a referential argument in a specific language (cf. in (ii) the variable underlined twice in AS) the referential argument of V can only be bound by I and the referential argument of N can only be bound by D.

ABNEY (1986) represents the similarities between sentences (IPs) and Noun phrases (DPs) in the structures (i) (the changes here with respect to other functional categories are not crucial for my point here):

(i)



All projections of lexical categories are semantically predicates, and the functional categories map them onto the basic categories <t> or <e>.

5. Lexical entries in German

The lexicon in the grammar is a model of the lexical knowledge of a native speaker and a lexical entry represents all linguistically relevant information a lexical unit entails. Using the notation of lexical entries e.g. in BIERWISCH (1986, 1988) I give the examples in (ii) for illustration only:

(ii)

		SF	
PF	LC	AS	PAS
1. ähneln	+V -N	λy λx λw	[w INST [SIMILAR x,y]
2. ähnlich	+V +N	λy λx	[SIMILAR x,y]
3. Ähnlichkeit	V +N	(λy) λx λw	[w INST [SIMILAR x,y]]
4. schlafen	+V -N	λy λx	[x INST [SLEEP y]]
5. schlaf	-V +N	(λy) λx	[x INST [SLEEP y]]
6. schläfer	-V +N	λy	[x INST [SLEEP y]]
7. unter	-V -N	λy λx	[[LOC x] BELOW loc y]

6. The argumentation

(1) In the Semantic Form SF of a lexical entry all lexical units of any subcategory are represented as **category-neutral n-place predicates**. Specific theta roles are assigned to the n arguments, depending on the meaning of the unit.

(2) Argument structure (AS) contains information with respect to the way in which the argument variables are saturated. Internal arguments are saturated by **theta marking** only (if not left unbound). In syntactic terms the arguments get addresses with respect to assigned case, if they don't have structural case, cf. the difference in (iii):

(iii)

	Anna	*(ähneln)*
	(mit) Anna	*(ähnlich sein)*
(die Ähnlichkeit)	*mit Anna*	

(3) The **highest argument** in the hierarchy causes the relevant categorial discrimination of lexical units. It is **a referential or an external argument,** never an internal one.

(a) The hierachically highest argument of the lexical entries 1, 3, 4, 5 and 6 in (ii) is **referential** (underlined twice). Referential arguments are saturated by **theta-binding,** i.e. binding by a functional category I or D. A lexical category with an argument identified as referential in the AS forms a referential category in configuration with a functional category. The **type** of the functional category binding the referential argument determines whether the resulting category refers (via proposition) to a situation or (via individual term) to an object. This is the difference between 1, 4 on the one hand and 3, 5, 6 on the other, i.e. between V and N.

(b) The hierarchally highest argument of the lexical entries in 2 and 7 is **external** (underlined once). Lexical units without referential arguments are referentially dependent, they need a referential anchor to participate in the denotation of a situation or an object. Units of this sort are designated to the **modificational** function. Remaining n-place predications
- they modify situations: adverbal modification; or
- they modify objects: adnominal modification.

A and P have external arguments only and are first of all heads of modifiers. This may surprise the reader, what about the predicative position? I ask him or her for patience until section 8.

Here we have the third kind of saturation of an argument in AS: the external argument of an A or P is **theta-unified** with an argument of the referential phrase, modified by the maximal projection of this A or P.

(4) A referential argument distinguishes V and N from A and P. In STEINITZ (1993), 86 I postulate the feature [Ref]erential to be the first candidate for a revised lexical subcategorization. WUNDERLICH (1993, 1994) incorporates lexical subcategorization into markedness theory. He supposes the complement feature **[Dep]endent** to be unmarked, which then gets the minus value per default in the case of V and N, whereas A and P are [+Dep]-categories. I adopt Wunderlich's proposal.

V and N differ in the **type of reference** they are transmitted to by theta-binding. Candidates for an appropriate feature are the two basic semantic categories <t> and <e> respectively. Again, we have to decide which category is the unmarked one, V or N, and in this case it is not so obvious. In an earlier version of this paper I decided on V to be the maximally unmarked category, probably the only category which appears in every language of the world, I based evidence from SASSE (1988) and others. But the interpretation of the data is not quite convincing yet. For the time being I think that N is the maximally unmarked category. With this I agree with WUNDERLICH (1994) again, whose second feature [Rel]ational as well as the feature [complement] of HALE/KEYSER (1994) gives the same subcategorization with respect to markedness. This is not surprising because we have only the two alternatives. But I do not follow the motivation, I will come back to this in section 11, then I hope to give evidence that only properties I propose are able to motivate lexical subcategorization features. For the moment I take for granted that [<t>] is the second categorization feature, and [-<t>] is equivalent to [+<e>].

Based on the assumption in section 3 and 4 with respect to the basic categories <t> and <e> and their correlates IP and DP in syntax, I conclude that at least these phrases must differ in a grammatically relevant way. That is, V and N have to be differentiated no higher than the projections of the functional heads with which they combine.

The syntactic categories IP and DP are universal categories. The two lexical categories V and N are probably distinct in the default case as well. Note that this assumption is not stated from a European-based view as Sasse supposes, but on the line of my argumentation hitherto.

(5) The modificational and referentially dependent category is not subcategorized so far. In languages with only one modificational category, that is enough. But what to do with clearly distinct categories as the German A and P, although they behave much the same in their syntactic functions? They can both be in adnominal or adverbal position, with a lot of exceptions, which may be relevant, as ZIMMERMANN (1988:284) argues.

I want to take A and P as two multifunctional predicate expressions with certain preferences in their syntactic positions. In the unmarked case A modifies a projection of N, it is theta-unified with the referential argument of N. Therefore A has the features [+Dep], [-<t>]. In the unmarked case P modifies a projection of V, it is theta-unified with an argument of V and has

the features [+Dep], [+<t>].[2] P is determined to be the most marked category, a welcomed result so far.

7. The new feature candidates [Dep] and [<t>] for subcategorization

In contrast to the classic version the base of this subcategorization is not the definiendum itself (V has a V-feature, N has an N-feature etc., whatever these features are). The features are defined independently of the categories they classify, by properties of a kind, which I suppose to be fundamental to the aim of subcategorization.

The features do not rely on specific classes of meaning to which the lexical units belong either, such as 'activity' 'thing' or 'time-stable vs. time-instable' or the like. Nor do they belong to semantic properties such as 'relational'. They rely on that part of AS in SF which collects the formal semantic content, the structure of meaning, and information of its specification in syntax.

In this sense lexical categories are **AS-guided**, that is, they are based in grammar. They are subcategorized with respect to their highest argument in hierarchy. The definition is based on the category specific AS of lexical items, and the specific way of saturation of the highest argument is crucial to the categorization of the items. To summarize:

(iv)	1. [-Dep]:	Theta binding of the highest (=referential) argument:	N, V
	2. [+Dep]:	Theta unification of the highest (=external) argument:	A, P
	3. [-<t>]:	Theta binding by D:	N
		Theta unification with D:	A
	4. [+<t>]:	Theta binding by I:	V
		Theta unification with I:	P

[2] Is there empirical evidence for this option? In (i) it is plausible to name the adnominal positon of PP the marked one:

(i) *Helena tanzte durch die Räume*
Die Frau in Troja

It is related to relative clauses (*die Frau, die in Troja war*), in some languages relative clauses are the only option. Other adverbial subcategories are limited to abstract nouns in the adnominal position:

(ii) *die Frau im neunzehnten Jahrhundert / damals*
die Frauen in den Städten / dort
**die Gäste wegen des Geburtstages /* deshalb*
Die Feier wegen des Geburtstages / deswegen
**Das Kind aus Unachtsamkeit /* deswegen*
Die Verletzung des Kindes aus Unachtsamkeit /?deshalb
*Der Besuch aus Vergnügen / * mit Vergnügen*

Situations or events are modified in respect of more different dimensions than objects. With respect to A the adverbial position is the marked one:

(iii) *die schöne Helena*
Helena bewegte sich anmutig

because only "manner adjectives", corresponding to manner adverbials with *-ly*affix in English, can occupy the adverbal position. Other adjectival subcategories are interpreted as predicative attributes:

(iv) *Peter kam krank nach Hause*
Eßt das Obst doch roh

The features reflect the properties lexical units have as lexical heads of syntactic categories in their structural contexts.

It would be near at hand to conclude that lexical categories is mapped onto syntactic positions in a one-to-one fashion, as in (v):

(v)

syntactic position	lexical features [Dep] [<t>]	lexical category
a) subj/obj	- -	N
b) finite VP (predicate)	- +	V
c) attributive	+ -	A
d) adverbial	+ +	P

The same features subcategorize the lexical categories and the syntactic positions, it looks like a redundant representation. But the world is not as tidy as that. Lexical category and syntactic position are related, but not identical as supposed in traditional German grammar. The lexical categories can have more than one syntactic function and a function can be met by more than one category, as we have seen in the case of A and P.Thus the situation isn't clear even in German, a language with a specialized system of lexcal categories (cf. HENGEVELD (1992)).

If we take non-Indo-european languages into account the situation seems to be quite confusing, I come back to this topic in sections 9 and 10. And above all, you can have every lexical category except the finite verb in the predicative position.

8. The predicative position

In this context I refer to phrases as in predicative position if they are complements of an overt or covert copula, and I interpret certain configurations of predicative and copula as "**analytic verbs**", which fill gaps in the verbal system. They vary according to the functional perspective of the sentence etc.[3] For illustration see (vi):

(vi)	*schlafen/einschlafen*	:	*wach sein/ wach werden*	(AP + copula)
	zusammenhängen	:	*in Zusammenhang stehen*	(PP + copula)
	---	:	*in Zusammenhang bringen)*	
	--- *ohnmächtig sein*	:	*in Ohnmacht liegen*	
	--- *ohnmächtig werden*	:	*in Ohnmacht fallen*	
	--- ---		*im Zimmer stehen/ sein*	
	--- ---		*ins Zimmer kommen/geraten*	
	regieren	:	*Regent sein*	(DP or NP + copula)
	--	:	*König sein/werden*	
	-- *katholisch sein/werden:*		*Katholik sein/werden*	
	ähneln : *ähnlich sein :*		*Ähnlichkeit haben*	
	--- *ähnlich werden:*		*Ähnlichkeit bekommen*	

A phrase XP in the position **predicative** has following properties:

- it is the complement of a copula,
- it gets no theta-role assignment from this copula,
- therefore, the phrase is not a proper argument, it is not referential,
- but like a normal verb the lexical head of the XP has a theta-marked external argument.
- A succsessfull theta-marking presuppose an argument position already opened by the copula.

A **copula** has the complementary properties:

- it has a referential argument just like a normal verb.
- it has positions for internal and external arguments too, but it does not theta-mark them, thus they turn to pure slots. One slot is filled by the non referential predicative complement and one is filled by the external argument of the head of this predicative complement.

To assign theta-roles to its arguments is a property of syntactic categories corresponding to semantic n-place predicates of type <e,t> , <e<e, t>> etc.; arguments are of type <e>,

---

[3] On the one hand I use the term "predicative" in a restrictive way, only dealing with non referential qualifying predicatives and not with referential identifiing ones, cf. the difference in (i):

(i) *My brother was gardener*
The murder was the gardener.

On the other hand I use the term in a expansive way as the examples in (vi) in the text illustrates.
(a) I subsume the so called periphrastic PPs + copula-like verbs (Funktionsverbgefüge) cf. *in Zusammenhang stehen/ geraten* under the term.
(b) Periphrastic DPs like *Ähnlichkeit haben/ bekommen* have in my opinion the same interpretation, the verb *haben* being a "transitive copula. Cf. the argumentation in STEINITZ (1977)
(c) In STEINITZ (1990) and (1991a) I interpret local and directional adverbials, which are complements of verbs as predicates as well, cf. *im Zimmer sein, stehen/ ins Zimmer kommen, geraten.*
(d) *zu*-infinitivals as *Er hat/ ist zu lieben* probably belong to the predicatives too.

whether they are internal, that is syntactic complements, or external, that is syntactic subjects (in kernal sentences). The copula is the only verb, which has a complement beeing not an argument but a predicate of type <e, t>. The position of the complement can be filled by AP and PP without any complication, they are semantic predicates themselves. What is going on?

In the modificational position A and P have an external argument theta-unified with an argument of the modified phrase, that is both A and P are referentially dependent, they are [+Dep]-categories. In predicative position, however, their external argument has to be theta-marked, just like the external argument of V, a [-Dep]-category. But predicatives need the configuration with a "verb maker" to be real, analytic verbs. The predicative position requires neither a [+Dep]- nor a [-Dep]-head. The feature is irrelevant. In other words, in this position it is not instantiated. That is all that happens in the case of A and P.

But the [-Dep]-category DP can fill the predicative position just as well. In this case, however. the irrelevance of a feature has consequences. In predicative position the DP loses the referential capacity with the syntactic consequence that it gets a lot of constraints with respect to article alternation, modification by relative clauses etc.

The hierachically highest argument of an N in argument position is a referential one, it corresponds to the external argument of an N in predicative position.[4] This is a "process" inverse to the nominalization process in e.g. *lehren* to *Lehrer,* where the external argument of the V corresponds to the referential argument of the N. But in contrast to this process of word formation (and even conversion), in the predicative position we don't have a change of category, an N remains an N in all positions.

The semantic type of the DP in predicative position however has to be <e, t> just like the type of AP, PP, and N projections, which belong to the class of common nouns. I will leave it open for the time being whether we have to describe these facts in terms of "defective DPs"[5] or in terms of "templates", a change in SF only.[6]

What to do with the feature [<t>] of lexical catgories in this position? The normal interpretation according to whcih theta-binding or theta-unification of the highest argument by D or I respectively (see (iv) above) can not be true, because in predicative position the highest argument has to be theta-marked. That is, the feature [<t>], just like [Dep], is not instantiated in predicative position.

To sum up, the predicative position requires no special lexical category. The subcategorization features need not be instantiated. In German all lexical categories can be inserted in this position. The only operation of saturation of an argument is theta marking of its external argument. Corresponding to this the phrases are all category-neutral predicates in SF.

The lexical distinction becomes relevant in non predicative positions:

- In modification position the external argument is unified only [+Dep]-categories can be inserted. Therefore, I called the modificational position the relevant position of AP and PP, distinguishing them from other categories.
- In the position of an individual argument only a [-Dep]-category without an external argument can be inserted, the lexical head has to be an N.

---

[4] You can describe the facts in this way. D loses the capacity of a functional category to bind the referential argument of N. The argument has to be saturated another way; it will be theta-marked, that is the argument type changes and gets a external one.

[5] STEINITZ (1988) is an attempt to interpret predicative nominals as defective NPs.

[6] Cf. DÖLLING (1992) and Blutner (1992).

9.Universality

On this background I can refine a bit what was said in section 2. I insist on the assumption that the grammatical features [Dep] and [<t>] are basic for lexical subcategorization.[7] Are the four lexical categories universal in the sense that they appear in every language as postulated in earlier GB theory? Obviously not.

But there is a minimal condition that every language must meet, this is the result of my work until now:

- Every language has to discriminate between the two semantically basic categories <e> and <t> on some syntactic level; the distinction must be universal.
- All languages are likely to have a distinct category of modification too, adding restrictive or appositive properties to the referent. This distinction too helds not necessarily on the lexical level.

What are the possible combinations and which probably will be realised in natural language? You can see from section 8, that in predicative position the four distinct categories N, V, A and P are redundant even in German, it would suffice to have a non specific category together with a "verb maker".

From that you can deduce the **minimal equipment** a language must have:

- no distinction of lexical categories at the lexical level. This language has n-place predicates as syntactic primes and in addition "functional elements" which specify the predicates.
- a "verb maker" (copula or equivalents) identifies the configuration as the predicate of the sentence.
- a "noun maker" (article, quantifier, classifier ot equivalents) identifies the configuration as the subject or object of the sentence.
- if not combined with a verb- or noun-maker, the external argument of the syntactic prime will be unified with some other argument. The configuration is identified as modification of some kind.

---

[7] My idea on how to reconstruct the obvious affinity of certain categories of meaning to certain lexical categories (cf. section 2) in the framework given is not clear enough. I discuss it briefly in this footnote. Look at the lexical entries 4, 5 and 6 of the sample (ii) in section 5. The three lexical entries have identical representation in PAS, they all are predicates with the form [x INST [SLEEP y]]. They differ in AS and the addresses to be added to their highest, that is referential argument: If bound by I(nflection) the lexical entry is a verb. If bound by D(eterminer) the lexical entry is a noun, the functor D takes a predicate and maps it onto an individual term with the syntactic correlate DP, *das Schlafen, der Schlaf.* In the word formation process the entry loses verbal properties as fixed in the verbal inflection paradigm (tense, mood etc.). Nevertheless the DP refers to the situation entity of sleeping in contrast to the DP *der Mann*, which refers to an object entity. Similarly, a DP in predicative position loses nominal properties, cf. *Peter ist ein Langschläfer*, though the lexical head remains a noun as we have seen.

A lexical entry belongs to a lexical category X according to its feature values. It is a"typical X", if the grammatical properties correlate with the semantic properties in the sense of UNITYP,section 2:

bound by ***semantic category***	***I(nflection): VP***	***D(eterminer) : NP***
situation, event	**typical verb** *schläft*	nominalization *der Schlaf*
property of an object	predicative +cop *ist ein Mann*	**typical noun** *der Mann*

In such a language the features <t> and <e> are not instantiated at a syntactic level higher than X^0.

The combinatory of the two features together with markedness conditions result in a list of possibilities natural language have in respect to their inventory of lexical categories.

(a) **Features are not instantiated**

No instantiation of a feature means no lexical discrimination by this feature, categories collapse.

(vii)

1. Neither [Dep] nor[<t>] is instantiated:	N/V/A/P
2. Only [Dep] is instantiated:	N/V, A/P
3. Only [<t>] is instantiated:	N/A, V/P
4. [<t>] is instantiated in combination with [-Dep] only:	N, V, A/P
5. [<t>] is instantiated in combination with [+Dep] only:	P, A, N/V
6. [Dep] is instantiated in combination with [-<t>] only:	N, A, V/P
7. [Dep] is instantiated in combination with [+<t>] only:	V, P, N/A

The collapse of N/P, V/A is excluded.

(b) **Features are instantiated but not marked** (= underspecified)

The distinction between two categories ist neutralized, the unmarked part of the distinction is generalized. In terms of markedness it means there is no positive value of the feature, the unmarked feature gets the value minus by default. At this point the earlier decision on N as the most unmarked category gets crucial, cf. Section 6.

(viii)

1. Neither [Dep] nor[<t>] is marked, there is one unmarked category only:	N
2. Only [<t>] is marked:	N, V
3. Only [Dep] is marked:	N, A
4. [<t>] is marked only in combination with [-Dep] or	
5. [Dep] is marked only in combination with [-<t>]:	N, V, A,
6. [<t>] is marked only in combination with [+Dep]:	N, A, P
7. [Dep] is marked only in combination with [+<t>]:	N, V, P

3 and 6 are not likely to appear in any language. If there is any lexical discrimination at all, N and V must differ in any case. The constraint to exclude them looks something like:

> If [Det] is marked (has the plus value, the unmarked part gets the minus value by default), then the unmarked part of it must be marked with respect to [<t>].

This is the situation in 5 or 7. You can combine (vii) and (viii), so the possibilities of variation increase. I will leave it as it is.

Is this representation a model to describe the situation in natural languages? Or more precisely, do instantiation and feature-marking form a grid in which differences between languages can be arranged? Although I am more sceptical than in the first version of my paper in respect to the language specific data Hengeveld uses in his book I still think the idea in HENGEVELD (1992) is interesting enough to try a reconstruction in the model I presented here.

10. Hengeveld's typology and its reconstruction in the model

Hengeveld postulates a typology with respect to the lexical categories in various languages (cf. footnote 1):[8]

(A) **Specialized languages** (four lexical categories, normally with four distinct syntactic functions).
(B) **Non-specialized languages** with the subclasses
  (a) **flexible languages** (a single part of speech may be used in different functions)
  (b) **rigid languages** (for certain functions a part of speech is missing)

The sample:

(A) **Specialized** type: German, English etc.: four categories

(Ba) Non-specialized type **flexible**:
  Quechua: **N/A** form one part of speech, which has the syntactic subject/object position of DPs and the position of adnominal modification.
  Lango: **A/Pitr** form one part of speech; the same attributive particle for adnominal and adverbal modification
  Tongan: **V/N/A/Pitr** form one part of speech in all syntactic positions. The units discriminate at a higher level of X projection. If the data are true this would be the minimal equipment I mentioned above!

(Bb) Non-specialized type **rigid**:
  Hausa: **A is missing**, the attributive function is occupied by Ptr ('a person with kindness')
  Wambon: **Pitr (manner adverb) is missing**, medial verbs (verbalized A) have the adverbial function.
  Inuit: **A and Ptr/itr are missing**, V and N only; N combined with modificational suffixes (like German *-lein* 'little') ('child-little, man-old') and nominalized V are adnominal modifiers, N with oblique case affixes are adverbal modifiers.
  Mand.Chinese: **A is missing**, V and N with relativizer (*de*) have attributive function. **Pitr is missing**, reduplicated V has adverbial function.
  Tuscarora, Cayuga: **N, A, Pitr are missing.** "Predication words" or verbs (SASSE (1988)) get different syntactic functions ('(It is) a cat, it is dead' = 'the cat is dead/the dead cat')

I see the problems, the sample is small. Many of the data are not clear neither in HENGEVELD (1992) nor in the papers by typologists. If the pair "rigid/flexible" is relevant at all and if the sample of languages is somewhat representative and if it reflects the data well (many ifs indeed), I can reconstruct "rigid/flexible" in terms of the features [Dep] and [<t>] together with statements about markedness.

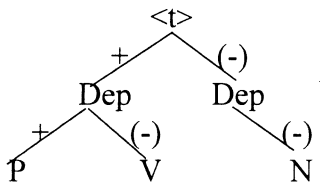
[8] As I already mentioned in footnote 1, I ignored the abdverb as the fifth category for not so good but understandable reason. But now it becomes disturbing, because Hengeveld ignores just prepositions and takes only manner adverbs into account. My comporise in this section is that I name these adverbs Pitr (intransitive P) in contrast to prepositions, which I name Ptr.

Rigid:flexible and markedness

(A) In specialized languages the whole combinatory of two features with two values holds; result: four categories.
(B) In non-specialized languages there are two different ways to take into acount:
  (Ba) In flexible languages at least one feature remain uninstantiated, with the result: less than four categories with mixed properties (two or more of the general possible categories collapsed).
  (Bb) In rigid languages all categorization features are instantiated, but at least one is unmarked and gets the value minus by default. The result is less than four categories, but clear discrimination of each category by two features.

To illustrate this take Chinese, a rigid language in Hengeveld's typology. Chinese is said to have adjectives as a subclass of verbs, cf. GASDE (1993). In the presented model I would say adjectives are not specified at all; Chinese is a language of type (viii-7), where [Dep] is marked only in combination with [+<t>]:

(ix)



In a language with less than 4 categories it is a matter of choice which other category undertakes the job of the "lost" one. There may be preferences, but if Hengeveld is right, the job of the adjective is done:

in Hausa by P,
in Inuit it is N combined with suffixes
in Chinese it is V and N with a relativizer.

11. Is V or N the least marked lexical category?

To sum up, in earlier versions of this paper I thought V to be the unmarked category and I was supported by various descriptions in rigid languages such as Cayuga, cf. SASSE (1988) or Tuscarora in HENGEVELD (1992). They describe these languages as having one category only and they call this category a verb. There is no language with a noun only.

In the meantime I have another idea about the data in Cayuga, even if I have no knowledge of this language, I ask the specialists for discussion. The summary in sections 8 and 9 was that the predicative position can be occupied by a phrase the lexical head of which can be any of the four categories. This position do not need lexical units of any special category, a non specified "predicate" would do the job as well. It might be that the predicate in Cayuga or Tuscarora is such a lexically unspecified entity.

In languages with specified lexical categories any category can occupy the predicative position. In contrast, the argument position, the position of subject or object, can be occupied only by a phrase the lexical head of which is an N. The noun is the only category, which can occupy both the argument position and the predicative position without any change. N is in this sense the least marked category.

Note that my motivation for the decision to classify N as the most unmarked category is quite different from WUNDERLICH (1994). He motivates his decision as follows. N is the most productive category with respect to word formation, almost all other categories can transform into nouns by word formation, it is the most frequent category with the greatest semantic variation, it can be an argument of all other categories, it is learned at the very beginning of the language learning process.

I think most of the arguments are true but too diverse to be relevant criteria. The last argument is relevant but not true, as far as I can see. I suppose the first words a child learns are those like *balla, heia* 'sleep', and *atta* 'walk'. They are category-neutral predicates, i.e. they are used **predicatively**.

HALE/KEYSER (1994) and WUNDERLICH (1994) have similar but not identical subcategorization features. Wunderlich's feature [Rel]ational reflects the assumption that V and P are typically relational, whereas N and A are typically not relational. This feature is based in the PAS of a lexical item, i.e. on meaning categories, more general than the UNITYP-parameters. The feature is also based on semantic features as well. With such features, you must end up with prototypicality. But how does one handle the non-typical cases like intransitive verbs, relational nouns and adjectives? The major motivation of my proposal was to find grammatically based properties for subcategorization, valid for all categories.

HALE/KEYSER (1994) have a more general view on relationality. They interpret the feature [complement] without the notion of typicality. They assume that true verbs are [+complement] or relational units. Verbs behaving differently are taken as results of an incorporation process. So-called unergative verbs like *sneeze* which lack the transitivity alternation are the result of incorporating a nominal complement into an empty verbal head. *Clear* in *The screen cleared* is the result of incorporating an adjectival complement in an empty verbal head etc. The authors denote the structural properties of the four lexical categories with the two features [complement] and [subject], the latter one corresponds to Wunderlich's and my feature [Dep]. Both terms of HALE/KEYSER are syntactic. They suppose a universal system of lexical categories: V is [+complement, -subject], N is [-complement, -subject], A is [-complement, +subject] and P is [+complement, +subject]. In individual languages there are morpholexical processes like incorporation and by this the clarity of the system is often obscured. Though many questions remain open this is a very interesting idea.

12. Invitation

On this basis I want to go further and I need the knowledge of specialists of various exotic languages. I would like to invite them to test the idea to see whether it is supportable. In order to do this it may be necessary to have a new look at his or her data.

Bibliography

Abney,Steven (1986) Functional Elements and Licensing. Glow Conference 86.

Bierwisch, Manfred (1986) On the Nature of Semantic Form in Natural Language. In: F.Klix/H.Hagendorf (eds) Human Memory and Cognitive Capabilities - Mechanisms and Performances, 765-784, Amsterdam-New York: North-Holland Publishers.

(1988) On the grammar of local prepositions. In: Bierwisch, M., W. Motsch, I.Zimmermann (eds)

Bierwisch, M., W.Motsch und I.Zimmermann (eds) (1988) Syntax, Semantik und Lexikon (=studia grammatica XXIX) Akademie-Verlag Berlin.

Blutner, Reinhard (1992) Type shifts in dynamic Montague grammar and the definitness restriction in *there*-insertion contexts. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 1 -22.

Bok-Bennema, Reineke (1991) Case and Agreement in Inuit. Studies in generative grammar 38. Foris publications. Berlin, New York

Broschart, Jürgen (1991) Noun, verb, and PARTICIPATION. In: Jakob Seiler/Waldfried Premper (eds.) Partizipation. Das sprachliche Erfassen von Sachverhalten. Gunter Narr Verlag Tübingen, 65-137

Chierchia, Gennario (1985) Formal Semantics and the Grammar of Predication. Linguisitc Inquiry 16, 3 (1995) 417-43

Dölling, Johannes (1992) Flexible Interpretation durch Sortenverschiebung. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 23-62.

Gasde, Horst-Dieter (1993) Lexikalische und funktionale Kategorien des chinesischen Satzes. Sprachtypologie und Universalienforschung 46 (1993) 4, 239-274

Hale, Ken, & Jay Keyser (n.d) The limits of argument structure. MIT-paper

Haider, Hubert (1993) Deutsche Syntax generativ. Narr Verlag Tübingen

Hengefeld, Kees, Non-verbal predication. Functional Grammar Series 15, Mouton de Gruyter.

Hopper, Paul J./Thompson Sandra A. (1985): The discourse basis for lexical categories in universal grammar. Language 60/3, 703-752

Riemsdijk, Henk van (1983):The Case of German Adjectives In: F. Heny and B. Richards (eds.) Linguistic Categories: Auxiliaries and Related Puzzles, Vol. One, Reidel Publishing Company, 223-252

Rothstein, Deborah Susan (1985): The syntactic forms of predication

Sasse, Hans-Jürgen (1988): Der irokesische Sprachtyp. In: Zeitschrift für Sprachwissenschaft 7, 2 (1988) 173-213

Steinitz, Renate (1969) Adverbialsyntax. (= studia grammatica X) Berlin: Akademie-Verlag, 206 pp.

(1977) Zur Semantik und Syntax durativer, inchoativer und kausativer Verben. In: Untersuchungen zur deutschen Grammatik I. (= Linguistische Studien des ZISW der Akademie der Wissenschaften der DDR, Reihe A, H. 35) 85-129

(1988) Bewegungenn bei und um Subjazenz, Grenzknoten und Satzglied. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 185-218.

(1990) Prädikation, Modifikation und Adverbiale. In: Linguistische Studien des Zentralinstituts für Sprachwissenschaft der AdW der DDR, 206, 117-132.

(1991a) Durative und inchoative Prädikate und die Adverbialkomplemente von Verben. In: Hoffmann, Ludger (Hrsg.), Deutsche Syntax - Ansichten und Aussichten. Berlin, New York: W. de Gruyter. (Jahrbuch 1991 des Instituts für deutsche Sprache, Mannheim). 1992, 186-205.

(1993) Lexikalische Kategorie und syntaktische Funktion. In: Wuppertaler Arbeitspapiere zur Sprachwissenschaft 9, 1993, 79-107.

(1994) Lexikalische Kategorisierung. Talk on the annual meeting of the DGfS
Wunderlich, Dieter (1994a) Lexikalische Kategorien. Talk on the anual meeting of the DGfS.
(1994b) Lexical Categories, Universität Köln, unpublished paper.
Zimmermann, Ilse (1988a) Die substantivische Verwendung von Adjektiven und Partizipien. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 279-311.
(1988b) Wohin mit den Affixen? In: Linguistische Studien 179,157-88. AdW der DDR.
(1994) Zur Legitimierung syntaktischer Merkmale. Talk on the annual meeting of the DGfS.

Derivational Economy and the Analysis of V2 *

Chris Wilder
Max-Planck-Gesellschaft, Arbeitsgruppe Strukturelle Grammatik, Berlin

1. Introduction

Generative analyses treat verb placement asymmetries in finite clauses in the Germanic "verb second" languages as a phenomenon reflecting differences in syntactic derivations. Descriptively, V2 is handled as verb movement to C, but views on the explanation of the phenomenon (reviewed in sections 3.-4.) diverge. There are three aspects to the problem which a comprehensive account must consider:

(i) to identify a "V2-parameter" that explains why finite verbs raise to C in some languages and not in others,
(ii) to account for the language-internal asymmetry: why V-to-C occurs in some clause types in but not in others,
(iii) to account for the "reverse side" of V2, i.e. the distribution of initial constituents in V2-sentences.

In this paper I develop an analysis of finite verb placement in German within the Minimalist framework of Chomsky (1993) that addresses the issues in (i-iii).

A key role is assigned to the economy factor ("least effort/last resort") in determining syntactic derivations. Movement, triggered by abstract morphosyntactic features of functional categories, may apply before or after the "Spell-Out" point that determines the input to the phonological component. Word order variation is a consequence of derivational asymmetries induced by universal economy principles, which interact with a parametrized property of triggers to determine whether a given movement happens in the syntactic derivation before or after "Spell-Out". In section 5., I show how the integration of economy into syntactic explanation yields a conceptually elegant analysis of the basic V2 asymmetry (ii). I propose that a finiteness feature F that is shared between finite complementizers and finite verbs is implicated in movement of finite verbs to C. Phrasal movement (iii) is driven by other independently parametrized features of C. The model permits a unified account of cross-linguistic and language-internal variation (i-ii) in terms of parametrization of F.

In section 6., I consider Zwart's (1993) arguments for an asymmetric approach to declarative V2, in which the verb lands in an Infl-head (AGR_S) lower than C in subject initial sentences (SU-V2). The claim is supported by the asymmetric properties of preverbal pronouns in declarative V2 which argue against a "generalized topicalization" treatment. Like Zwart, I consider economy to provide the key to understanding this asymmetry. However, Zwart's implementation of the asymmetric V2-analysis relies too heavily on unmotivated stipulations. In section 7., I suggest an alternative approach to declarative V2 in terms of a structural analysis for SU-V2 which unifies "symmetric" and "asymmetric" analyses of V2. Borrowing Haider's (1989) idea of "Matching Projections", SU-V2 clauses are viewed as single projection of two categories C and Infl (AGR_S), whose head hosts the finite verb. In derivational terms, simultaneous projection of functional heads in SU-V2 reflects the pressure of economy on operations that project structure.

2. Word order and derivation

The problem of word order variation is often thought of in cross-linguistic terms: how to account for the fact that one language has a certain order, e.g. SVO, of constituents that appear in a different order in another, e.g. SOV. Early approaches captured such variation by attributing the grammar of a language with a phrase structure rule governing the linear order of a verb and its object:

(1) a. VP –> V NP b. VP –> NP V

With the insight that phrase structure should be viewed as governed by universal principles (X'-theory) interacting with specific properties (parameters) and without the intervention of a rule component (cf. Chomsky 1981, 1986), recourse to (1) is no longer possible.

In the Principles-and-Parameters framework, this type of variation is attributed to "parametric variation" - grammars vary with respect to a specific property (parameter) which interacts with fixed principles of UG to yield the different serializations. The VO / OV variation is attributed to a "directionality parameter" that permits language-specific choice of the direction in which a head licenses its complement. The parameter might affect the government relation that licenses Θ-role and Case-assignment (Koopman 1984, Travis 1984); alternatively, left- and right-headedness are just properties of (classes of) lexical items in a language (Chomsky 1993). Although these alternatives replace rules with parameters, the difference between OV and VO languages is still encoded in basic phrase-structural configurations, as in (1):

(2) a. $[_{VP}$ V NP] b. $[_{VP}$ NP V]

A second aspect of the problem concerns word order variation within languages - words appear in different sequences in different constructions. One of the best studied examples is the the main / subordinate clause asymmetry in the placement of finite verbs in German and other V2-languages:

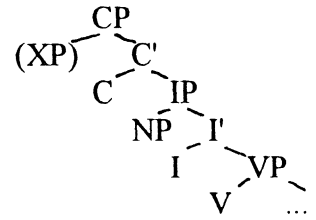
(3) a. *Sie <u>schreibt</u> heute ein Gedicht.*
she writes today a poem
b. *...daß sie heute ein Gedicht <u>schreibt</u>.*
that she today a poem writes

The treatment of this type of variation differs from the approach to cross-linguistic variation in terms of basic configurations just sketched. Most generative analyses of V2 start from the assumptions (4):

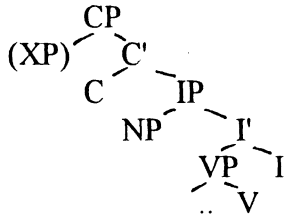
(4) a. The phrase structure of finite clauses does not vary language-internally.
b. The finite verb in a V2-clause is the same item as the finite verb in the corresponding verb-final clause.

Since Chomsky (1986), (5a) has become established as a working hypothesis about sentence structure of English. The basic clause structure of other languages may differ with respect to the order of VP-internal (2) or other major constituents. A common assumption about German is that I (Infl) follows rather than precedes its VP-complement (5b):

(5) a.



b.



But in each language, embedded and main clauses will have the same structure.

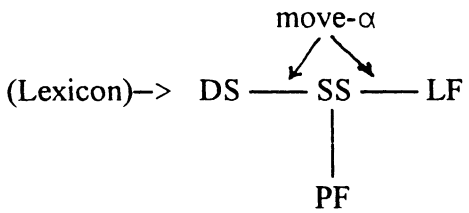
In the standard description of V2, the finite verb in V2-clauses occupies the C-position, while in verb-final clauses, it is dominated by V (or maybe I). The placement asymmetry reflects the linear order imposed by the hierarchical structure dominating the relevant positions in (5b). The V2 / V-final asymmetry would be accounted for by assuming that the finite verb in V2-clauses is an item of the category C, while the finite verb in verb-final clauses is a V (or I) element, so that the placement asymmetry follows from the location of the relevant category in the tree. However, this means adopting two categorizations, effectively two lexical entries, one for a "second position" verb, and one for a "final" verb, in conflict with (4b). So phrase-structure is not solely responsible for the V2-effect.[1]

Given the assumptions (4) and a structure (5), V2 is a classic example of a phenomenon that needs to be handled in terms of movement of constituents in a derivational model. The structure of VP in (5) expresses the idea that "underlyingly", German is an OV-language. Assuming that finite verbs have the category V, and corresponding constraints on the structural projection of lexical information, the finite verb in both examples (3) enters the phrase structure (5) in the position of the V-node. So, the finite verb of a V2 sentence can only reach the C-position by undergoing movement in the syntactic derivation. Hence, subordinate clause order is considered the "underlying order"; V2 is a "derived order".

3. V2 as V-movement to C

In the "T-model" of grammar assumed in Chomsky (1981,1986), the representation (structural description) of a sentence is viewed as a quadruple representation {D,S,L,P}, with each member satisfying level-specific principles. Three of these (D,S,L) are syntactic phrase-markers. The syntactic derivation, the mapping between DS and SS, and SS and LF, is mediated by movement operations ("move-α"):

(6)

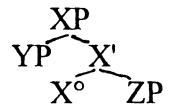


V2-structures arise through movement of the finite verb from the V-position (which it has to occupy at D-structure) to C before S-Structure; the position of the finite verb in the PF-string reflects its position in the S-Structure representation (crucially, the mapping from SS to PF does not involve move-α). The derivation DS-to-SS in subordinate clauses on the other hand does not involve movement of the verb to C: so the finite verb appears in final position in the PF-string.

3.1 X'-Theory and move-α

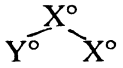
The derivational analysis of V2 depends on assumptions about phrase structure and the movement operation involved in syntactic derivations. X'-theory constrains the structural properties of representations at all levels. Following Chomsky (1993), each lexical category forms the 0-level head of a two-level projection (7). The sister of X° is its complement; the sister of the intermediate projection X' is the specifier of the phrase:

(7)

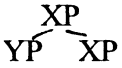


Other non-head constituents are permitted in adjunction structures (8), whereby only a head may be adjoined to a head, only a phrase to phrase:

(8) a.



b.



Other possibilities are excluded, including (9), which plays an important role in the account of the V2-effect:

(9) Adjunction to X' is excluded.

The most important constraint on movement derives from the assumption that move-α leaves a trace, and that α and its trace must form a well-formed chain. In particular, a moved category must c-command its trace, and may not be "too distant" from its trace. X'-theory must be respected by the "output" configuration, and so constrains moved categories and their landing sites. A phrase may either substitute into an XP-position (a specifier), as in *wh-* and NP-movement constructions; or adjoin to another phrase. A head may move only to another head position, by substitution or adjunction.

3.2 The standard description

Under what I shall call the symmetric analysis of V2, V-movement to C, an instance of head movement, is implicated in all clause-types in which the finite verb has fronted (see Den Besten 1983, Holmberg 1987, Schwartz & Vikner 1989, Vikner & Schwartz 1991). The initial phrasal constituent, which may be a subject (10a), a non-subject constituent such as the adverbial in (10b), or a *wh*-phrase (10c), is assumed to occupy the specifier position of CP. V1-orders (10d) either involve unfilled SPEC,CP, or as is more commonly assumed, an abstract (phonologically null) phrase - a "question operator" or "conditional operator" - in SPEC,CP.

(10) a. *Sie <u>schreibt</u> heute ein Gedicht*
she writes today a poem
b. *Heute <u>schreibt</u> sie ein Gedicht*
c. *Was <u>schreibt</u> sie heute?*
what writes she today
d. *<u>Schreibt</u> sie heute ein Gedicht?*
e. *..., daß sie heute ein Gedicht <u>schreibt</u>*

Where V-fronting has not occurred (10e), the finite verb stands in a head position inside IP (I in most versions). So there are exactly two positions in which a finite verb may surface - a clause-initial position (C), and a clause-final position (I). The standard symmetric analysis

contrasts with the "asymmetric" analysis (Travis 1984, Zwart 1993) on the description of subject-initial declaratives (10a), which is discussed in §§6-7 below.

The strength of the V-to-C hypothesis is perceived to lie in its potential to account for two main properties of V2: the fact that the finite verb is always the first or second constituent in the clause, but (with certain qualifications) never the third; and the fact that fronted finite verbs are in complementary distribution with lexical complementizers. I discuss these in turn.

3.3 V1, V2, but never V3

That the finite verb in V2 may come second, after an initial phrasal constituent, but never third, following two phrases, is accounted for fairly easily under the V-to-C hypothesis. If the finite verb is in C, then there is only one position available inside the clause for a phrasal constituent to the left of the verb - SPEC,CP. This result depends on the unavailability of (i) adjunction to C' (cf. 9) and (ii) adjunction to CP.

With regard to (ii), provision needs to be made for left-dislocation structures (LD). In German, a V3 order may arise where an LD-phrase is initial, followed by a pro-form (bearing identical case) preceding the verb:

(11) *Den Bürgermeister, den mag wahrscheinlich keiner*
the mayor (ACC) pron-ACC likes probably noone
"As for The mayor, probably noone likes him."

The dislocated phrase is outside the domain of the "second effect" (maybe adjoined to CP); the verb is "second" with respect to the pro-form.

(12) * *den Bürgermeister, den wahrscheinlich mag keiner*

LD is restricted to certain referential phrases, such as the definite expression in (11). A negative quantifier, for example, cannot be dislocated (13). LD can thus be distinguished from topicalization to SPEC,CP (14), which is free of such restrictions (Haider 1990):

(13) * *Niemanden, den mögen wir*
noone-ACC pron-ACC like we

(14) a. *Den Bürgermeister mögen wir nicht*
the mayor like we not
"The mayor, we don't like"
b. *Niemanden mögen wir*
"We like noone"

From (9), it follows that the SPEC,XP must be adjacent to the intermediate projection X'. SPEC,XP generally precedes X', so if the head X precedes its complement, SPEC,XP will be string-adjacent to X. Evidence for the need to assume (9) comes from a variety of sources. In cases where a language permits an overt complementizer to cooccur with an overt *wh*-phrase or relative pronoun in SPEC,CP (Dutch, Swedish), no further element may ever intervene between the two. In direct *wh*-questions in English, the fronted finite auxiliary is obligatorily adjacent to the *wh*-phrase in SPEC,CP. Like verbs in V2, the auxiliary is assumed to move to C:

(15) *Which car (*probably) did John buy?*

With respect to (9), the "second effect" in V2 robust: virtually all cases of "V3" are best analysed with the second constituent in SPEC,CP (X - SPEC - C° rather than SPEC - X - C°) [2]. Adverbs may not intervene:

(16) * *Hans wahrscheinlich hat mich gesehen*
H. probably has me seen

Apparent cases of SPEC - X - HEAD, such as English (17a), must therefore be analysed so that the finite auxiliary does not occupy the head of the phrase in whose SPEC the subject stands (Kayne 1989, 1993):

(17) a. *John probably has not seen Mary*
b. * *Jean probablement n'a vu Marie*

French does not permit adverbs to intervene between subject and finite verb (17b), which supports the assumption that the French finite verb raises to the head of the category (AGR_S - see below) whose SPEC hosts the subject.

3.4 Complementary distribution of complementizers and fronted verbs

Wherever a lexical complementizer appears (declarative daß, interrogative ob, conditional-temporal wenn), verb-fronting is not attested, either to a medial position following the first phrase, or to an initial position following or preceding the complementizer:

(18) a. *daß sie ihn oft sieht*
that she him often sees
"that she often sees him"
b. * *daß sie sieht ihn oft*
c. * *daß sieht sie ihn oft*
d. * *sieht daß sie ihn oft*

A frequently cited argument for the V-to-C hypothesis is that it permits an account for the complementary distribution of overt complementizers and fronted V. This follows if the only position which verb-fronting can target is the C-position. The presence of an overt complementizer in C effectively blocks movement of the verb there.

In declaratives, overt complementizers alternate with V2-order; in interrogatives and conditionals, the alternation is with V1-order:

(19) a. *Wenn er kommt, (dann gehe ich.)*
if he comes then go I
b. *Kommt er, (dann gehe ich.)*

(20) a. ... *ob sie klug ist*
if she clever is
b. *Ist sie klug?*

The strict alternation between V1 and ob is neatly illustrated in comparison clauses introduced by als (Den Besten 1983):

(21) a. ... *als ob er betrunken wäre.*
as if he drunk were
b. ... *als wäre er betrunken.*

Similar paradigms can be created for all V2-languages. Even in English, where V-to-C is confined to finite auxiliaries in restricted contexts, complementary distribution of complementizer and fronted verbs can be observed in conditionals:

(22) a. *If she had read this book, ...*
b. *Had she read this book, ...*

4. Triggers and the role of economy

In conjunction with X'-theoretic constraints and a concrete hypothesis about clause structure, the analysis of V2 as V-movement to C is able to capture word order facts, providing a descriptive basis for an explanatory account of verb-placement asymmetries in V2-languages.

However, in the derivational model, movement is in principle free to apply (within general constraints indicated in §3.1). How to block overgeneration is a fundamental issue for a derivational model as complex as (6). Even restricting attention to the derivation of S-Structures from D-Structures, move-α allows for massive overgeneration. If the V-to-C hypothesis is to be of any use in explaining the finite clause word order asymmetry, an account is needed of why V-raising takes place when it does, and why it does not in other cases.

An account is also needed for what Zwart (1993) calls the "reverse side" of the V2-phenomenon. The constraint on what can precede the finite verb in V2-sentences - i.e. at most one phrasal constituent - follows as discussed above. Under usual assumptions, the constituents that turn up in SPEC,CP are not base-generated but moved there; and in general, some constituent not only may, but must occupy SPEC,CP at S-Structure.

So the derivational analysis of V2 must provide answers to (23):

(23) a. Why does the finite verb have to move to C in some contexts?
b. Why can't the finite verb move in other contexts?
c. What enforces the movement of some constituent into SPEC,CP in V2-clauses?

The questions (23a/c) are essentially about the triggering of movement; (23b) addresses the blocking of movement. Taken together, they concern the general issue of preventing overgeneration.

4.1 Complementizers and root clause status

Most attempts to answer (23a) involve reference to the complementary distribution of fronted verbs and complementizers, by seeking to tie V-fronting in V2-languages to the absence of a lexical complementizer in C.

However, while an overt complementizer never cooccurs with a fronted verb, the generalization does not quite go through in the other direction. A phonologically empty C does not automatically trigger V-fronting - in indirect interrogatives, the verb cannot front even though C contains no overt material (the *wh*-phrase is in SPEC,CP):

(24) a. *Er fragte, wer ihn gesehen hatte.*
he asked who him seen had
b. * *Er fragte, wer hatte ihn gesehen.*

The solution often favoured is to postulate a lexical complementizer with syntactic feature content, but no PF-content. Supporting evidence comes from the fact that the postulated abstract element sometimes finds overt expression, as in Dutch (of) and Swedish (som):

(25) a. *Ik weet niet, wie of het gedaan heeft* Dutch
I know not who C° it done has
b. *Jag vet inte vem som har gjort det* Swedish
I know not who C° has done it

This allows the trigger for V-fronting to be related to empty C-positions, with the abstract *wh*-complementizer counting as filling C.

Although the absence of a lexical complementizer seems at first sight to correlate with the root status of the clause, it is generally agreed that the analysis of V2 should not be tied to the root / non-root distinction. The correlation is not exact in either direction. Clauses with V-final order (and overt complementizers) can be independently uttered, thus appearing to count as root clauses. They receive "indirect" interpretations, as an exclamation, echo question, or similar:

(26) a. *Daß du mir das nicht wieder tust!*
that you me that not again do
"Don't you dare do that again!"
b. *Ob ich ihn gesehen habe?*
whether I him seen have

The most reasonable conclusion is that a root clause may contain a lexical complementizer, and that this is responsible both for "indirect" interpretation, and for blocking V-fronting. Complement clauses lacking a lexical complementizer in German have the V2 order: [3]

(27) a. *Er hat gemeint* [*Hans habe dort einen Löwen gesehen*]
he has thought H. has there a lion seen
b. *Er hat gemeint* [*dort habe Hans einen Löwen gesehen*]

Where a constituent is extracted out of a V2-complement, V-1 order is obligatory: V2-order leads to strong ungrammaticality:

(28) a. *Was hat er gemeint* [t *habe* [*Hans dort* t *gesehen*]] ?
what has he thought has H. there seen
b. * *Was hat er gemeint* [*sie habe* [*dort* t *gesehen*]] ?
c. * *Was hat er gemeint* [*dort habe* [*sie* t *gesehen*]] ?

This provides further evidence for assuming that the fronted verb is in C. Supposing that the extracted phrase must leave an intermediate trace in SPEC,CP, the ungrammaticality of (28) results from the fact that the preverbal phrase must be in SPEC,CP, blocking that position for the intermediate trace of the extracted phrase. This derives the V1-order in (27c).

4.2 Empty C as trigger

Various attempts have been made to derive the idea that a finite verb must move to C prior to S-Structure if no lexical complementizer is inserted at D-Structure, though none has become established. Most proposals share the idea that some principle forces the C-position to be "lexicalized" at S-structure. One influential approach (Koopman 1984, Platzack 1986) relates verb movement to Case Theory. Assuming that Nominative Case for subjects is

licensed from the C-position in V2-languages, the "lexicalization requirement" is interpreted as a requirement on Case-assigning categories; V-fronting saves (29) from a Case Filter violation by the subject. [4]

(29) * [C ø] *sie heute ein Gedicht schreibt*

Whatever the theoretical account, V2 has to be derived with the aid of some parametrized property, given that C remains unfilled in main clauses in other languages like English and French. There must be a "V2 parameter". [5]

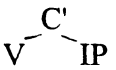
The answers to the questions (23) that emerge are that V2 and V1 orders are the result of some principle forcing movement of the finite verb into an empty C-position (23a); and the impossibility of V-fronting in other cases (23b) is due to the presence of lexical material in C° (through insertion at D-structure), which "blocks" the movement of the verb to C. In this way, the complementary distribution of complementizers and fronted verbs is directly integrated into the V-to-C analysis. On (23c), see §5.5 and §6. below.

4.3 Complementary distribution again

However, this way of integrating the complementarity of complementizers and fronted verbs turns out on closer inspection to be dubious. To ensure that the presence of a complementizer in C is sufficient to block movement of the verb to C, it must be assumed that verb-fronting is head-movement by substitution: only then will the position targeted by V-fronting be occupied by the complementizer.

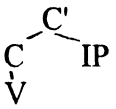
This assumption is questionable. Treating V-fronting as head movement by substitution gives an output - V directly dominated by the intermediate projection of C - that contravenes X'-theory (assuming that the fronted category is an Infl-head to which V is adjoined does not alter the argument):

(30)



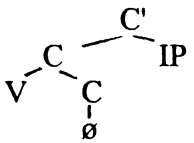
To avoid this type of violation, basic assumptions outlined in §3.1 must be modified. Rizzi & Roberts (1989) suggest that V-to-C creates a new type of "substitution" structure (31):

(31)



In the system outlined in Chomsky (1993), head movement by substitution is only permitted where the landing site heads a projection of the same category of the moved head. [6] V-to-C involves a landing site of a different category, so adjunction is the only possibility:

(32)



C dominates an element with no phonological content (but maybe syntactic feature content). The verb adjoins to this element.

It might be objected that treating V-fronting as adjunction loses the insight which a substitution analysis gives into the striking paradigms showing the complementary distribution of complementizers and fronted verbs. But to make the argument go through,

the substitution analysis needs an additional stipulation to the effect that V-fronting can only be head-movement by substitution. Otherwise, a string in which V-fronting cooccurs with an overt complementizer will have a licit derivation in which V adjoins to C. [7]

4.4 Movement as Last Resort

In his "Minimalist Program" (1993), Chomsky proposes a conceptually simple and intuitively appealing solution to the problem of overgeneration. The idea is that movement is only possible if necessary, that derivations are constrained globally by a fundamental principle of economy. In a given derivation, movement operations are not "optional", free to apply or not (within general limits), but are fully determined, or caused.

As in earlier approaches, there is assumed to be a set of "triggers", defined with respect to well-formedness conditions on the output representations, which necessitate applications of move-α in a derivation. If a necessary movement fails to take place, an ill-formed representation results. In such a case, move-α can be thought of as a "last resort" measure. If a movement that is not required for the well-formedness of a representation takes place in its derivation, the derivation will count as "uneconomic", even if the output were well-formed. The general principle (33) suffices to exclude such a derivation:

(33) Economy of Derivation (Last Resort)
"No unnecessary movement"

To illustrate the notion of Last Resort, consider how in the model of Chomsky (1981), the need for a representation to satisfy the Case Filter at the level of S-Structure acts as a trigger for movement of a NP in a Case-less position (John in (34a)) into a position where it can be assigned a Case-Feature (34b):

(34) a. e *was told John that it was raining*
b. *John was told* t *that* ...
c. * *It was told John that* ...

The NP cannot remain in its base position, as the ungrammaticality of (34c) shows.

The Last Resort nature of Case-driven NP-movement is shown by comparing (34) with (35). Even though the subject position of the verb seem is a legitimate target for NP-movement (as in Subject Raising examples), and even though English permits preposition-stranding under NP-movement (John was talked to), movement of the NP John to the subject position in (35b) results in an ungrammatical structure:

(35) a. e *seemed to John that it was raining*
b. * *John seemed (to)* t *that* ...
c. *It seemed to John that* ...

The NP can satisfy the Case Filter by remaining in its base position (35c); movement is thus unnecessary, and Last Resort ensures that the derivation leading to (35b) is ungrammatical.

In this model, the fundamental task in analysing a phenomenon like V2 is to identify the trigger for movement. Overgeneration - in the form of derivations which involve unmotivated movement - will be excluded by Last Resort.

The substitution-based account of the blocking of V2 in the presence of complementizers is made redundant by the adoption of (33). Given that a "trigger" for V-fronting must be independently assumed for V2-sentences, to ensure that movement takes place, (33) provides both a necessary and a sufficient account for the blocking of V-to-C where the trigger is not present. The complementary distribution of complementizers and

fronted verbs becomes a side effect of the distribution of the trigger: present in sentences where C is covert, and absent where there is a lexical complementizer. No particular problem arises with embedded interrogatives, where a covert C fails to trigger V-fronting. The trigger is simply absent in such cases.

It is important to realize the extent of the explanatory potential of this principle: even with respect to verb movement alone, a host of other cases of potential overgeneration are captured. Consider for instance the ungrammatical result of raising a participle to C in German:

(36) a. * *er* [$_C$ *getrunken*] *das Bier* t *hat*
he drunk the beer has
b. * *er* [$_C$ *getrunken hat*] *das Bier* t t

Example (36a), where the participle crosses the finite auxiliary on its way to C, might be excluded independently of Last Resort, by appealing to a locality principle for X°-movement, the Head Movement Constraint (HMC), which excludes one head from moving across another. (36b), where the participle has followed the auxiliary to C, cannot be so excluded. Indeed there are languages, such as Croatian, in which such movement patterns are attested: [8]

(37) *Pio je pivo*
drunk has beer-ACC
"He drank the beer"

Although (36b) is permitted by locality conditions on move-α, its ungrammaticality can be explained as a violation of Last Resort - there is no trigger for the movement of the participle (on the trigger for participle raising to C in Croatian, see Wilder & Cavar 1994b).

A further reason for wanting to reject the substitution-based interpretation of the "blocking" effect of overt complementizers will become clear in the analysis to be developed below. It will be claimed that the finite verb does in fact move to C in embedded clauses, in the derivation of LF from SS. Under this account, it cannot be that the presence of a lexical complementizer blocks the C-position as a landing site for verb movement.

A more revealing interpretation of the relation of lexical C to finite verbs is as that of an expletive to a contentful element (cf. Law 1991 for discussion). A lexical complementizer like daß can be seen as an expletive element that occupies a position otherwise occupied by a contentful element (a verb), much as the expletive form there occupies the position otherwise taken by a nominative subject:

(38) a. *there seem to be some dogs in the garden*
b. *some dogs seem to be* t *in the garden*

Just as the expletive there is always associated with a local nominative subject, so the finite complementizer is always associated with a local finite verb.

Chomsky (1993) suggests that the relation between there and NP be interpreted derivationally: the nominative NP in (38a) moves and adjoins to the expletive after S-Structure, to yield the LF (39):

(39) [*some dogs* [*there*]] *seem to be* t *in the garden*

An analogous treatment of the complementizer-finite verb relation in terms of LF-raising of the finite verb will be outlined below. The complementarity of finite verbs and

complementizers in C generalizes to the case (38) of English sentences with and without there: if the expletive appears in the target position, the associate is prevented from moving there in overt syntax.

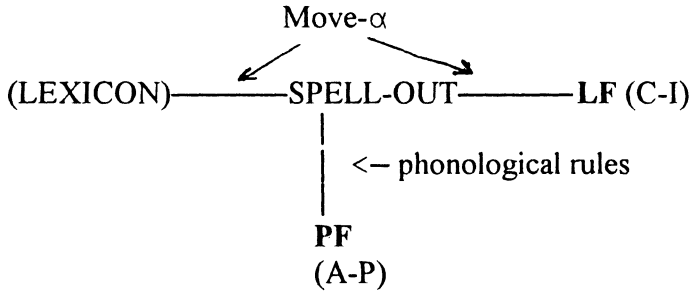
In the next section, I introduce key concepts of Chomsky's Minimalist Program, as a prelude to further development of the economy-based analysis of the V2 asymmetry.

5. V2 in the Minimalist Program

5.1 The Minimalist model

The D- and S-structure levels of earlier models (cf. (6) above), are eliminated from the revised T-model (40). The remaining levels LF and PF are the minimally necessary interfaces with non-linguistic performance systems - articulatory-perceptual (A-P) and conceptual-intentional (C-I) systems:

(40)



S-structure reduces to an intermediate stage (SPELL-OUT) at which the phonological derivation branches from the derivation to LF. D-structure is completely dissolved, so LF and PF are directly related to the lexicon by the derivational system. The representation of a sentence is simply the pair {L,P}.

In models with a D-Structure level interfacing with the lexicon, the mapping of lexical items onto positions in phrase markers is performed by a "lexical insertion" operation; tree construction is independent of lexical insertion, positions for "insertion" pre-defined. This approach is replaced with one in which the introduction of lexical items into derivations is interspersed with other operations involved in building syntactic trees.

Lexical insertion itself is reinterpreted as one such structure-building operation, which I call "Project-α". This operation projects a simple X'-tree from an item taken from the lexicon, the category of the lexical item determining node-labels in the tree. It may create a structure of any bar-level, the choice in a given case depending on subsequent operations to be performed on that tree:

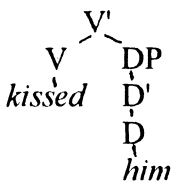
(41) a. V
kissed

b. DP
D'
D
she

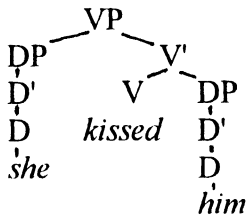
c. DP
D'
D
him

Subtrees are brought together by a binary tree-joining mechanism, "Generalized Transformation" (GT), which adds a branching node at the join. GT operates on the output of Project-α, and on its own outputs. To perform the union of a head with its complement, GT needs an X°-tree and a phrasal tree - e.g. (41a) and (41b), giving (42). Union of a specifier with an intermediate projection requires an X'-tree and a phrasal tree - (42) and (41c), giving (43). [9] As with Project-α, the output of GT must be licensed by X'-theory. The labelling of the node created is automatic in these cases:

(42)



(43)

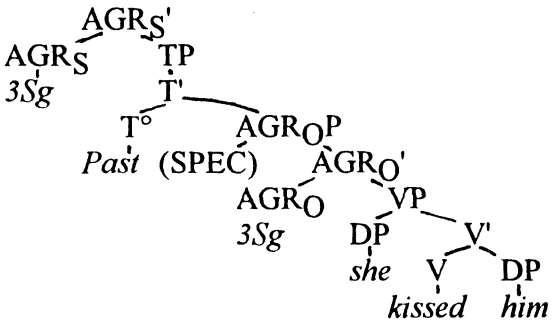


To build a sentence out of (43), further operations of Project-α yield simplex trees containing functional elements, each being united by GT with the nascent clausal structure as its complement. The clause structure Chomsky adopts is (44), incorporating the "split Infl" hypothesis:

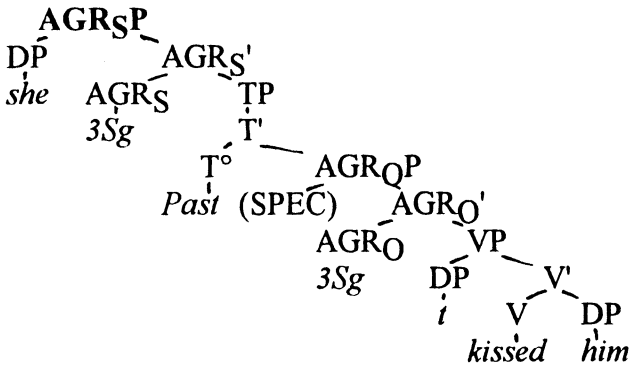
(44) $[_{CP} .. C [_{AGRsP} .. AGR_S [_{TP} .. T [_{AGRoP} .. AGR_O [_{VP} .. V ..]]]]]$

GT and Project-α interact with move-α, the third structure building mechanism. Move-α takes a subtree of an existing tree, and treats the two as trees to be joined (in the same manner as GT). Thus, a new branching node is created, dominating the moved subtree and the tree from which it was taken. (45) illustrates the movement operation that raises a subject from SPEC,VP to SPEC,AGR_S, which maps AGR_S' onto AGR_SP: [10]

(45) a.



b.



Projection from the lexicon and tree-extension cease at the Spell-Out point which inputs to PF; but move-α continues to apply after Spell-Out to derive LF-representations.

5.2 Economy of representation and morphosyntactic triggers

Economy principles play a fundamental role not just in constraining overgeneration (Last Resort), but also in explaining why movement is triggered. Movement in a derivation is necessary to ensure that the resulting interface representations meet the Full Interpretation requirement, understood as a representational economy principle:

(46) Principle of Full Interpretation (FI)
Interface representations may not contain superfluous symbols.

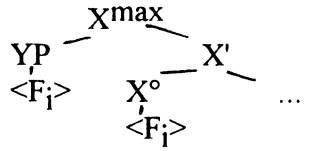
FI presupposes a set of legitimate objects in PF- and LF-representations. For PF, these are phonological symbols that are capable of phonetic interpretation; for LF, syntactic entities that are capable of semantic interpretation - i.e. chains (categories, together with their traces of movement, if any). The arrangement of these objects in the LF and PF of an sentence results from derivational manipulation of items projected from the Lexicon. FI states that interface representations must be fully interpretable: symbols that are superfluous in this sense are symbols that are not interpretable by the external performance systems.

Morphosyntactic features are entities entering syntactic derivations from the lexicon which are not externally interpretable. It is the need for such features to be eliminated during the derivation that triggers movement operations. Elimination presupposes "checking" of feature correspondence, and checking depends on movement. Movement in the derivation is in this way made necessary by requirements of representational economy.

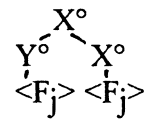
Triggers depend on morpho-syntactic features (m-features) as properties of lexical items. Case-features, Φ-features, Tense-features, and "operator-features" such as WH appear pair-wise in a structure: the Accusative feature borne by a verb and the Nominative feature of Tense (T°) will each match a Case-feature in a DP, Tense-features in the verb will match those in T°, etc. Chomsky proposes that morphosyntactic correspondence requirements (Case, Agreement, etc.) exhaust the set of triggers for movement. [11]

M-features may only be eliminated once matching pairs have been checked. The local configurations in which feature-checking is possible are the SPEC-head configuration for checking between a phrase and a head (47a), and head-adjunction for head-head checking (47b):

(47) a.



b.



Generally, one member of a checking pair needs to move to enable checking: a phrase or a head must be raised into the checking domain of a functional category. A functional head may contain a "phrasal feature" (N-features in (48)) checked by a phrase in its SPEC, and "head-features" (V-features) checked by lower heads that adjoin to it:

(48)

DP	AGR_S/AGR_O	T	V
<N->	<N-Φ>	<V-Tns>	<V-Tns>
<N-Case>	<V-Φ>	<N-Case>	<V-Φ>
<N-Case>			

The Φ-features of AGR_S match those of the finite verb and those of the DP with which the verb agrees. Thus AGR includes two sets of Φ-features, "N-Φ" and "V-Φ", each needing to be "checked" with a different element. [12]

In (45b), the subject has moved to a higher specifier, permitting checking of N-features in Infl (NOM in T and N-Φ in AGR_S). [13] This is the Spell-Out tree that feeds the phonological component. Before LF, all features must have been checked to permit deletion. So following Spell-Out, the verb must raise as far as AGR_S, to permit checking of all "V-features"; and the object must raise to SPEC,AGR_O to check N-features.

5.3 Economy of derivation: Procrastinate

The well-known contrast in the position of finite main verbs in English and French results from the fact that the English verb stands in V° at Spell-Out, while in French it has raised to AGR_S: [14]

(49) a. *John often kisses Mary*
b. * *John kisses often Mary*

(50) a. *Jean embrasse souvent Marie*
b. * *Jean souvent embrasse Marie*

In both languages, the subject raises to a functional specifier position before Spell-Out, while object raising is deferred to the covert syntax. Given that all features must be "checked" in a well-formed LF, overt raising of finite verbs in French is mirrored by covert raising of finite verbs in English.

The word order variation reflected in PF-representations is explained by the factors that determine whether a given movement operation applies before or after Spell-Out. Chomsky's account rests on the assumptions (51):

(51) a. Certain m-features are "invisible" in PF-representations.
b. Movement after Spell-Out is preferred.

FI constrains PF as well as LF, so we might expect all movement to apply prior to Spell-Out, to ensure all features are eliminated by PF. Chomsky proposes that some m-features have a special property enabling them to be present in PF without causing a violation of FI. Such features are weak ("PF-invisible"), contrasting with m-features that have the property of being strong ("PF-visible"), requiring to be eliminated prior to PF. Only strong features force move-α to apply prior to Spell-Out.

The V-features in Infl-heads in English are weak: the verb does not have to raise and check these features prior to Spell-Out (49a). In French, a V-feature in AGR is strong, forcing the finite verb to raise prior to Spell-Out (50).

The fact that English verbs may not raise before Spell-Out (49b) follows from (51b). Applying Last Resort to the preference for movement "late" in the derivation yields a specific derivational economy principle:

(52) PROCRASTINATE
A given application of move-α must occur after Spell-Out, unless earlier movement is necessary to ensure fulfillment of FI at PF.

Movement in the overt syntax is a "last resort" option, used only to avoid a violation of FI at PF. Overt V-raising in the derivation of French (50a) is permitted, since delaying it leads to an FI-violation in PF (50b). Overt raising of a finite verb in English (49b) violates Procrastinate.

A highly restrictive theory of parameters emerges. Depending on the degree of uniformity of clause structure and of the distribution of features in functional heads, LF-representations will vary little if at all across languages. Word order variation may then reduce to differences concerning the way move-α feeds Spell-Out, as determined by the interaction between strong m-features and Procrastinate. Parametrization then consists solely in the "weak" and "strong" options for m-features. This hypothesis may be further restricted, if only functional elements are parametrized (cf. Borer 1984, Chomsky 1991,

Ouhalla 1991). In that case, the marked value for m-features will be "strong", and all major category items will bear weak m-features. [15]

5.4 Economy of derivation: Greed and altruism

The effect of the fundamental derivational economy principle (Last Resort) is to block gratuitous operations; only operations needed to satisfy interface principles can apply. The paradigm originally used to substantiate Last Resort (cf. (35) above), actually motivates a stronger principle:

(53) a. * *Seemed to John that it was raining*
b. * *John seemed (to)* t *that ...*
c. *It seemed to John that ...*

The ungrammaticality of (53a) is due to the strong Nominative Case feature in Infl that is unchecked. Movement of the DP John into the subject position would be one possibility for rescuing the sentence, and in this sense, would not be a "gratuitous movement". But movement is blocked, and pleonastic insertion is forced, to enable NOM to be checked. The reason why the movement option is not available is that the DP itself can satisfy its own Case requirement in situ. [16] (54) ensures that movement of α must be "self-serving":

(54) GREED
Move-α must result in satisfaction of requirements of α.

Given Greed, there can be no purely "altruistic" movement: if α does not need to move to position X to check features of its own, then movement of α to X is illegitimate, even if this operation would satisfy requirements of another item ß, in accord with the weaker formulation of Last Resort.

As pointed out in Wilder & Cavar (1994b), "altruism" *per se* is not excluded in this system. "Greedy" movement may (and generally does) have "altruistic" side-effects: the movement of α to position X leads to the checking of m-features of some ß ≠ α. The raising of a verb α to an Infl-position X enables m-features of both α (Greed) and a functional element ß (Altruism) to be checked. Moreover, if strong features that enforces overt movement are located in the higher functional heads (ß), and not in the element α that moves, then overt movement is generally "early altruism": α raises before Spell-Out to permit checking of ß's strong feature, in order to save the derivation from FI-violation at PF, even though α itself would only need to raise by LF.

A characteristic of early altruistic movement is that a constituent-type has different distributions within a language. That a nominative DP in English may appear lower than SPEC,IP (55a) is a result of the fact that the features it bears and checks in SPEC,IP are weak - so movement of the subject raises to SPEC,IP in overt syntax (55b) is early altruistic movement:

(55) a. *There seems to be a man in the garden*
b. *A man seems to be* t *in the garden*

Similarly for *wh*-movement: the fact that *wh*-phrases can remain in situ (in multiple questions) indicates that the features that they check against C, when in SPEC,CP, are weak, so that overt *wh*-movement represents early altruism, to enable checking of a strong feature in C. [17]

5.5 The trigger for V2

So far little has been said about the feature content of the functional head C in this system; this forms the basis of the following analysis of V2, which builds on the proposal made in Wilder & Cavar (1994b).

V2 involves two processes: V-raising and XP-raising, both to the checking domain of C. In each V2-clause, there is one trigger for V-fronting and another for XP-raising (cf. (23) above). Also, there are at least two XP-fronting processes involved. The initial phrasal constituent in *wh*-questions with V2 must be a *wh*-phrase (bear the feature WH). In declarative V2, the initial phrase may be of any category, but must not be a *wh*-phrase.

In line with many previous approaches, I suppose that the complementizer in C in direct *wh*-questions contains an m-feature WH. The WH-feature is a "phrasal" feature, and is strong, forcing pre-Spell-Out movement of a *wh*-phrase to SPEC,CP.

In declarative V2-clauses, C will not contain the WH-feature. Instead, I suppose that a different complementizer with a phrasal feature "TOP" heads the clause. TOP is an m-feature, which may be checked by any phrase with "topic properties". [18] Crucially, however, TOP is incompatible with the WH-feature. A German V2-clause whose initial constituent is a *wh*-phrase cannot be interpreted as a declarative topicalization. The incompatibiity is clearer in English, where subject-auxiliary inversion occurs with *wh*-fronting in direct questions, but not with preposed topics. A main clause like (56c) is not interpretable as a non-interrogative topicalization - it is simply ungrammatical:

(56) a. *Who did she see?*
b. *John, she saw.*
c. * *Who, she saw.*

Further properties of topicalization are discussed in §6. below.

In embedded clauses, the declarative complementizer appears not to bear any phrasal feature, in contrast with the main clause declarative C. The complementizer heading interrogative CP-complements will bear a strong WH-feature. As in direct questions, this complementizer is not usually associated with a phonological form. The major difference concerns whether the finite verb raises to C before Spell-Out or not.

The concept of early altruistic movement provides the key to understanding language-internal word order variation, including the finite verb-placement asymmetries in V2-languages. Movement of the verb to C must be triggered by a strong m-feature that the verb itself has to check. The fact that a finite verb may remain inside IP (embedded clauses) indicates that, whatever its identity, the feature the verb checks in C in main clauses is weak on the verb itself. V-raising to C in V2 clauses must be early altruism. Moreover, in embedded clauses, the verb that remains in IP at Spell-Out must check its weak feature by LF. So the corresponding feature in C is weak, and the verb raises to adjoin to C after Spell-Out. This conclusion is more or less forced, if Greed is adopted, and a single lexical specification for finite verbs is assumed (cf. §2.(4b) above).

The simplest hypothesis on the identity of the trigger for V-raising to C, is that C hosts a V-feature which it shares with a finite verb. The need to check and eliminate this feature will force adjunction of the finite verb to C. If the feature is strong, V-raising will occur before Spell-Out. If it is weak, Procrastinate will delay V-raising to the covert syntax.

In Wilder & Cavar (1994b), it is proposed that the trigger for V2, and more generally for all cases of raising of finite verbs to C, is a finiteness feature (F) residing in finite clause complementizers. Finiteness can be viewed as a property shared by finite complementizers and some head in a finite IP, which is reflected in the selectional relation that obtains between finite complementizers and finite verbs, and which explains why a finite complementizer does not govern an infinitive complement.

I assume that the second F-feature that matches F in C is a property of finite verbs. In principle, F might be located in some Infl-head containing a strong V-feature. Then, V-to-C would be a side-effect of I-to-C. But where V-features in Infl-heads are weak, as seems to be the case in Swedish (cf. Wilder & Cavar 1994b), F must be a property of the verb itself. [19]

This approach identifies different complementizers not only for embedded clauses, where they have different phonological shapes, but also for different main clause types. As we have seen, it is necessary to distinguish main clause C-types at least with respect to the phrasal feature they contain. It is also possible for the value of F in main clause C to vary language-internally. This is what is responsible for the basic difference between V2 languages and English: declarative main clauses involving a strong TOP-feature. In German, the relevant complementizer has strong F, triggering overt V-fronting (57), while in English it has weak F, so inversion does not occur (56b):

(57) *Dieses Buch hat er gelesen.*
this book has he read.

English has another main clause C with strong WH and strong F, which causes auxiliary-fronting in English direct questions (56a). [20]

This analysis places the "lexicalization requirement" on finite C in V2-clauses in a different perspective. Main clause complementizers are not radically empty in root clauses, as supposed in previous analyses (cf. §3 above.). Instead, C always has feature content, and fronted verbs adjoin to, rather than substitute into, the C-position, as discussed in §4. The trigger for overt V-fronting is dissociated from the phrasal feature of the complementizer that determines "clause-type": the relation is indirect, mediated by lexical idiosyncrasy. Whether F in C is strong or weak is a (more or less arbitrary) property of lexical entries.

Two generalizations are not explained in this account: the tendency for overt V-to-C (strong F) to be restricted to main clauses, and the fact that complementizers with independent phonological form usually block overt V-raising (have weak F). Neither of these is exceptionless. Embedded V2 is found in German; and it may be that VSO order in embedded clauses of VSO languages involves overt V-to-C. Also, some languages appear to possess complementizers with phonological content which can attract finite verbs, i.e. "affixal" complementizers such as the *wh*-complementizers in Croatian (Wilder & Cavar 1994a) and Berber (Ouhalla 1992). This type of variation can be accommodated. Another combination which this system makes possible would be a language in which only subordinate complementizers have strong F, yielding "reverse V2". This seems not be attested at all (H. Haider, p.c.). It may prove necessary to incorporate the notion of "subordinate complementizer" in some way to account for this type of restriction on possible variation. I leave the issue open here.

The "V2-parameter" thus reduces to the value of F and phrasal features in complementizers. Languages in which main clause complementizers have strong F and strong phrasal features are V2-languages. Languages in which only a subset of main clause complementizers have strong F, such as English, will be "restricted V2 languages".

There is a close relation between this treatment of V2 and Chomsky's proposal concerning cross-linguistic variation (English vs. French) in finite V-raising to Infl. There, variation follows from different values of a V-feature in an Infl-head for which there is only one exponent in each language. In the case of V2, we distinguish a number of different exponents of a functional category (C) in a single language, which makes language-internal variation possible. In this sense, the V2 asymmetry is driven by a "language-internal parameter", and we have a unification of the accounts of cross-linguistic and language-internal word order variation. [21]

6. Subject-initial declaratives

With respect to the account of phrasal movement in V2-sentences, a problem arises. Zwart (1993) argues that not just two but (at least) three processes are involved - in addition to the distinction between WH-V2 and declarative V2, subject-initial declaratives (SU-V2) need to be separated from declaratives with fronted non-subjects (TOP-V2). The main argument for distinguishing SU-V2 from the rest (X-V2) comes from theory-internal considerations.

6.1 Economy and SU-V2

It should be possible to form declarative sentences with a "neutral" reading, in which no constituent is topicalized. In a language without generalized V2, like English, this would be the characterization of ordinary subject-initial main clause declaratives. The subject occupies its canonical position (SPEC,IP), and C contains neither a strong F-feature nor a phrasal feature, so no phrase precedes the subject, and the finite verb (or auxiliary) does not raise to C.

Sentences with similar interpretational properties in V2 languages are SU-V2 clauses. Zwart bases his argument on (58):

(58) Subjects in SU-V2 do not (necessarily) have topic properties.

The symmetric analysis claims the finite verb is in C in SU-V2, so the subject must occupy SPEC,CP. If a phrase with a TOP-feature must have "topic properties", and the subject does not have these properties, then movement of the subject to SPEC,CP is not triggered by TOP. It should raise no further than SPEC,IP, where it checks Case and Agreement (irrespective of whether C hosts a TOP-feature). Movement of the subject to SPEC,CP, which the data force us to assume, is unmotivated, and should be a violation of Last Resort (Greed). It is unclear why the unmarked order for root declaratives is not VSO.

The alternative is to suppose the subject in SU-V2 is not in SPEC,CP but in SPEC,IP. Since the finite verb follows the subject in SU-V2, F in C must be weak. This is the situation found in English for unmarked declaratives:

(59) [$_{CP}$ C *she kissed him*]

Now, the question is why the finite verb does not remain in its final position - why neutral subject-initial main clause declaratives do not show embedded S O V order:

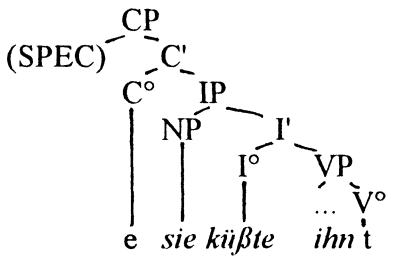
(60) * [$_{CP}$ C *sie ihn küßte*]

The S V O order forces the conclusion that the finite verb occupies an intermediate head position in SU-V2. This is the claim made by Zwart, following an earlier proposal by Travis (1984).

6.2 The asymmetric analysis of V2

In the "asymmetric analysis", a third position where finite verbs can surface is identified. While the finite verb is in C in X-V2 sentences, as in the standard approach, in SU-V2 it occupies an Infl-position (61). This analysis is clearly incompatible with the view that Infl in German is at the end of the clause, following VP. At least one Infl-head must precede VP to act as a landing site for the finite verb in SU-V2. [22]

(61)



One of the original arguments for the uniform V-to-C analysis of V2 lay in the account it made available for the strict second position effect (cf. §3.3). Second position effects are just as strong for SU-V2 as in X-V2: but this does not necessarily count against the asymmetric approach, as the explanation depends on the verb occupying the head of the projection whose specifier the first phrasal constituent occupies, independent of whether C, or some Infl head is involved. If the subject is in SPEC,AGR$_S$ and the verb in AGR$_S$, the strict requirement for adjacency between subject and verb in SU-V2 can be still explained by appeal to X'-theory. SU-V2 sentences will be like French subject initial declaratives in this regard. [23]

The fact that SU-V2 complements form "islands" for extraction that are just as strong as X-V2 (cf. (28), §4.1 above) has been used as an argument against the asymmetric analysis (Schwartz & Vikner 1989):

(62) * *Was hat er gemeint* [t [C [*sie habe* [*dort* t *gesehen*]]]] ?
what has he thought she has there seen

For the asymmetric analysis, (62) shows that where the verb has only raised to I, the SPEC,CP position in (61) is not available for intermediate trace of extraction. If each stage of a successive-cyclic movement operation is locally motivated by a strong TOP or WH-feature, which then brings the trigger for V-to-C with it, the impossibility of (62) can be attributed to an unchecked strong F in C.

I consider next some of Zwart's (1993) empirical arguments for the need to distinguish initial subjects from other initial constituents in declarative V2, which concern the asymmetric behaviour of pronouns and properties of agreement morphology. [24]

6.3 Pronouns

Various authors have pointed out that SU-V2 and TOP-V2 clauses behave differently with respect to whether an unstressed pronoun may occur in preverbal position. Dutch permits reduced forms of personal pronouns to appear in first position when the pronoun is a subject (62a-b), but not when it is an object (63c-d):

(63) a. *Ik zie'm* "I see him"
b. *'k zie hem* "I see him"
c. *Hem zie ik* "Him, I see"
d. * *'m zie ik*

The contrast can be illustrated for German using the neuter pronoun es ("it", NOM/ACC), which seems to be inherently unstressed. Es preceding the verb may only be interpreted as Nominative - both examples (64) must be interpreted as subject-initial (Travis 1984). The intended interpretation - "It (the horse) has eaten the grass" - is successful for (64a) but impossible for (64b), which does not mean "the horse has eaten it (the grass)", but can only be interpreted with es as subject ("something has eaten the horse"):

(64) a. *Es hat das Gras gegessen.*
it has the grass eaten
b. (*)*Es hat das Pferd gegessen.*
it has the horse eaten

With other personal pronouns, the asymmetry reveals itself in whether a pronoun can be prosodically weak or not. The accusative pronoun ihn ("him") may appear in initial position, but only if it is stressed, indicated by upper case (65). In postverbal position it can be, and usually is, unstressed. This behaviour contrasts with that of nominative er ("he") which may be unstressed in preverbal position (66):

(65) *IHN (*ihn) habe ich gesehen*
"Him, I have seen"

(66) *ER / er hat mich gesehen.*
"He has seen me"

If a pronoun is contained in a larger topicalized constituent such as a fronted verbal projection, it does not have to bear stress:

(67) a. [*es geLESen*] *habe ich nicht*
it read have I not
b. [*ihn geSEHen*] *habe ich nicht*
him seen have I not

In this case, however, an accent falls on the verb. This suggests a generalization that a topicalized constituent must contain a prosodic stress peak, which must be realized on a pronoun, if this is the sole constituent of the topic.

Obligatory stressing in initial position is not specific to non-nominative pronouns, as becomes clear by considering examples where the subject of an embedded clause (a V2-complement) has undergone "long" topicalization to the initial position of a main clause. The examples (68) are cases of TOP-V2 - the local subject follows the fronted verb, which is in C. The same effect emerges. Nominative es cannot undergo long topicalization:

(68) a. * *Es, glaube ich, hat das Gras gegessen.*
it believe I has the grass eaten
b. *Das Pferd, glaube ich, hat das Gras gegessen.*
"The horse, I believe has eaten the grass"

Long-moved masculine nominative er must be stressed:

(69) *ER (*er), glaube ich, hat mich gesehen*
"HE, I believe has seen me"

Zwart's claim (cf. also Travis 1984) is that if obligatory stressing can be identified as property of topicalized pronouns, the fact that preverbal subjects do not have to be stressed can be used as evidence that a preverbal subject has not necessarily moved to SPEC,CP.

This asymmetry is also found in English. Unlike German es, it can be stressed (70), although acceptability contrasts with unstressed it, which is perfect in subject position. Topicalization of non-subject it is marginally possible, but only with prosodic stress, and concomitant focus reading:

(70) a. *It hurt her.* b. ?? *IT hurt her.*

(71) a. * *it, she liked.* b. ?? *IT, she liked.*

Like German ihn, him can be stressed, and must be, if it constitutes a topic itself, but not, if it is contained in a larger topic-phrase:

(72) a. * *him, she liked*
b. *HIM, she liked.*
c. ... *but* [*SEEN him*], *I haven't.*

Nominative he is can be stressed or unstressed, in subject position, but nominative pronouns can be "long-topicalized" only if stressed:

(73) a. *He hurt her.* b. *HE hurt her.*

(74) a. * *he, she claimed* t *would never hurt her.*
b. *HE, she claimed* t *would never hurt her.*

These paradigms show that there is a cross-linguistic generalization to be captured independently of the V2 question. There is no issue of whether the English subject in subject-initial declaratives is a topic; so pronouns must be stressed in initial position only if this position is SPEC,CP. The parallel with English lends weight to the claim that the position of the initial subject in German is also the canonical IP-internal one.

Further evidence concerns the distribution of expletive pronouns. An expletive like English there cannot long-topicalize from either nominative or accusative subject positions:

(75) a. * *there, he said* [t *would be someone at the door*]
b. * *there, he believed* [t *to be someone at the door*]

Like there, expletive pronouns in V2-languages, such as German es (found in impersonal passives and sentences witbh VP-internal subjects) cannot undergo long topicalization. Es freely appears in the subject position of SU-V2-clauses, however: [25]

(76) a. *Es wurde gestern getanzt.*
it was yesterday danced
"There was dancing yesterday"
b. * *Es, glaube ich, wurde gestern getanzt*

Since Chomsky (1981), expletive there is commonly viewed as a place-holder for the obligatory subject-position. In the Minimalist model, it is inserted to allow checking of a strong N-feature in Infl. The asymmetric V2-analysis allow this treatment to be extended to es in (76). Assuming that expletives cannot be topics, (75) and (76b) fall out naturally. Under the symmetric theory, the possibility for es to appear in topic position in (76a) but not in (76b) requires a different account.

6.4 Rizzi's proposal

In the symmetric theory, SU-V2 differs from X-V2 in that the phrase in SPEC,CP has an independently necessary agreement relation with the verb in C in SU-V2, but not in X-V2 (including where X is a subject from a lower clause). Rizzi (1991) suggests that this fact can

be used to explain the pronoun asymmetry in the symmetric V2-theory. He proposes that SPEC,CP counts as an A-position exactly when it agrees with its head, so the properties of the phrase may vary according to whether the position is an A- or A'-position. In particular, unstressed pronominal subjects and expletives can be licensed (only) in an A-position. [26]

This proposal does not sit easily with the treatment of Agreement in terms of independently projecting heads. Under the "split-Infl" approach, an "A-specifier" can be thought of as "A" by virtue of intrinsic features of the head of the projection: essentially, the N-Φ features of AGR. [27] To suggest that the N-Φ feature of an AGR-head can cause the specifier of another head to which it adjoins to "become" an A-position for a phrase with which it agrees, may turn out to be too liberal an extension. Consider a situation in which AGR_O raises with the verb to C. Depending on the analysis, this may be the case in V2: the finite verb raises successive-cyclically, carrying all intervening heads with it. Then the N-Φ features of AGR_O may allow the CP-specifier position to become an A-position with respect to a topicalized object, so that the intended benefit of the proposal is lost.

Nor, as Zwart points out, does Rizzi's approach to pronoun asymmetries in terms of the A-/A'-distinction solve the main conceptual problem with the symmetric theory. It is an "after-the-event" account; no explanation is offered for the "reverse side" of V2 - why SPEC,CP has to be filled in the first place.

6.5 More on TOP

Under the symmetric theory, English and German subject-initial main clauses fundamentally differ in that the initial subject in German must raise to SPEC,CP. A way of implementing the difference in the Minimalist model would be to assume that a strong phrasal feature (TOP) always coexists with strong F in main clause declarative C in German. Then SU-V2 involves topicalization of the subject, unlike English subject-initial declaratives. To account for the "unmarked" properties of SU-V2, it might be claimed that subjects are default topics, where no other constituent fulfills this role (this was proposed in Wilder & Cavar 1994b). The difference is purely formal: a preverbal subject in SU-V2 must bear a TOP-feature, while the subject in otherwise equivalent declaratives in English doesn't have to. This solution means that the differences dividing local subjects from objects and long-moved subjects cannot be related to the presence of a TOP-feature, and an alternatve account is needed for expletive es.

The advantage of the asymmetric analysis lies in the account it permits for pronoun asymmetries. Assuming a correlation between "topic properties" (including obligatory stressing for pronouns) and TOP, the asymmetry can be related to structural position (SPEC,IP or SPEC,CP) via the presence or absence of a TOP-feature in a principled way.

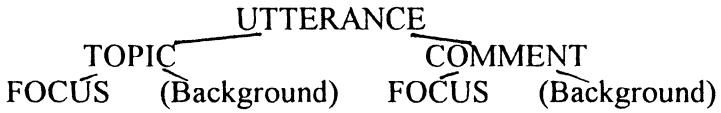
As it stands, the pronoun argument is based on descriptive generalizations; to be convincing, a theory is needed that relates "topic properties", including the obligatory prosodic peak, with the TOP-feature that forms the basis of the syntactic account of topicalization constructions. [28]

A theory of topicalization must refer to the notion of "well-formed information structure" in semantic interpretation. A topicalization construction imposes a bipartite structure on an utterance, dividing it into a "topic", which has a special relation to the discourse in which the utterance is embedded, and "comment". The topic corresponds to the phrase bearing TOP that surfaces in SPEC,CP, the comment to the remainder (the C'-constituent). It seems that both topic and comment must contain a focussed subpart to be well-formed from an information-structure viewpoint.

Suppose that focus is represented by a feature FOCUS associated with a syntactic node. This feature will be relevant both for phonological rules, which must assign a prosodic peak to a terminal element dominated by FOCUS, and for semantic interpretation.

Subconstituents not dominated by FOCUS (if any) will form a "background" with respect to the focussed constituent. The following hierarchical structure results:

(77)



The requirement for the comment to contain a focussed constituent may be what underlies an observation of Haider (1989), that topicalization of verbal projections may not empty C' (the COMMENT) of everything except a finite auxiliary:

(78) a. * [*Ein Außenseiter gewonnen*] [$_{C'}$ *hat* _]
an outsider won has

b. * [*Getanzt*] [$_{C'}$ *wurde* _]
danced was

c. [*Ein Außenseiter gewonnen*] [$_{C'}$ *hat hier noch NIE* _]
an outsider won has here yet never
"An outsider has never won here before"

d. [*Getanzt*] [$_{C'}$ *wurde hier noch NIE* _]
danced was here yet never
"There was never any dancing here before"

Being "semantically" empty, the auxiliaries in (78a,b) may be incompatible with the semantic requirements of focussing. In (78c,d), the appearance of additional material in C' that is capable of being focussed saves the examples (see also note 34).

If we suppose that a phrase bearing a TOP-feature must create a TOPIC in this sense, then the obligatory stressing facts follow. The fact that certain elements (reduced pronominals, expletives, and perhaps also *wh*-phrases - cf. §5.5)) cannot bear a TOP feature may be due ultimately to their interpretative incompatibility with the semantics of FOCUS.

Under the asymmetric theory, a subject-initial main clause (or embedded) declarative has no TOP-feature in C, and no constituent of the clause needs to form a TOPIC. In particular, the initial subject does not have to contain a FOCUS, so that no prosodic stress is required.

6.6 Agreement Morphology

Zwart adduces further evidence from an asymmetry of a rather different nature that turns up in the distribution of agreement morphemes.

Some West Germanic dialects show alternating morphological agreement on finite verbs, which seems to depend on the position of the finite verb with respect to the subject. If the subject precedes the verb (SU-V2 and V-final clauses) one form is chosen, if it follows the verb (X-V2) the other is chosen. The alternation is found in 2nd person singular forms in standard Dutch (examples from Zwart 1993):

(79) a. ..., *dat jij naar huis gaat*
that you to home go-2SG
"that you are going home"

b. *Jij gaat naar huis* (SU-V2)

c. *Vandaag ga jij naar huis* (TOP-V2)
today go-2SG you ...

d. *Waneer ga jij naar huis?* (WH-V2)
when go-2SG you ...

The alternation correlates with the position of the fronted finite verb under the asymmetric analysis: if the verb is in C, it takes the form ga, and gaat elsewhere. This structural correlation is not available to the uniform theory. [29]

Zwart proposes to relate this alternation to the phenomenon of C-agreement (complementizers inflected with morphemes marking person-number agreement with the finite verb and subject found in various West Germanic dialects - cf. Bayer 1984 on Bavarian, Haegeman 1990 on West Flemish). Zwart proposes that the special form of finite verbs in X-V2 in Dutch is an instance of the same phenomenon as C-agreement morphemes. These morphemes are generally homophonous with the agreement morphemes appearing on finite verbs, but in a few dialects, some differences can be detected. Standard Dutch does not have inflecting complementizers, but in dialects where verbal and complemetizer agreement morphemes have different forms, the claim is borne out directly. Fronted verbs carry C-agreement in X-V2 and V-agreement in SU-V2.

The way in which this paradigm can be used to support the asymmetric V2-theory depends on the interpretation of the distibution of the different morphemes. One possibility is that C can be inherently specified for Φ-features, which are checked against the Φ-features of V after it has raised to C (overtly in X-V2, at LF elsewhere) - cf. Law (1991). Alternatively, C-agreement morphemes are PF-forms associated with an independent agreement head (Shlonsky 1994). This is approximately the view Zwart takes (see below).

Whichever solution is adopted, if C is projected in SU-V2 clauses, special provisions are needed to account for the fact that C-agreement is not found in SU-V2.

7. Accounting for the Asymmetry

The conceptual advantage of the asymmetric analysis lies in the resolution of the problem of the "reverse" side of V2. The question of why the subject in SU-V2 must raise to SPEC,CP does not arise: the subject in SU-V2 is in SPEC,AGR_S, where it must move to check Φ-features, just as in X-V2.

However, a new problem arises concerning V-raising: three positions for the finite verb need to be accounted for, instead of two. Although V2 in X-V2 can still be handled in terms of strong F in C, the V2 / V-final asymmetry cannot be due solely to parametrization of the F-feature in C. A trigger for V-fronting to AGR_S in SU-V2 needs to be identified, and the account for subordinate V-final order in terms of weak V-features in Infl-heads and C may need to be revised.

The simple F-based account can be retained if it is assumed that in SU-V2 (only), F is projected in the same head as the Φ-features of AGR_S. Then, no additional triggers need to be identified: V2 signals strong F, V-final weak F (by Procrastinate). The main task for this strategy is to derive the restriction of cooccurring (strong) F and AGR_S-features in the same head to SU-V2 only - with F being a feature of C, separate from AGR_S, elsewhere. I attempt this below.

A second approach, followed by Zwart (1993), is to seek an additional trigger for V-raising to AGR_S in SU-V2 clauses which is absent (or "neutralized") in V-final clauses.

7.1 Zwart's Proposal

Zwart makes the assumption that finite verbs raise to C in any case at LF (this is necessary, given Greed - cf. §5.5). He does not identify the feature involved, and furthermore, proposes an independent trigger for overt V-to-C in X-V2. So we may take Zwart's LF-trigger to be F, weak in all cases.

In embedded clauses, the finite verb must check features in Infl heads, including AGR_S, and in C at LF. The questions that arise therefore concern triggering and blocking of overt movement:

(80) a. What is the trigger for V-movement to C in X-V2?
b. What is the trigger for V-movement to AGR_S in SU-V2?

(81) Why is there no trigger for V-movement...
a. ... to C in SU-V2?
b. ... to C in V-final clauses?
c. ... to AGR_S in V-final clauses?

How (81) is approached depends on the answers given to (80).

What ultimately causes a finite V to raise to AGR_S (80b/81c) is the need to check V-Φ-features, but the position of the verb in embedded clauses indicates that V-Φ in AGR_S is weak: if V-Φ in AGR_S were "strong", then finite verbs would be expected to raise to AGR_S in embedded clauses as well as main clauses. This difficulty could be avoided by positing different lexical specifications of AGR_S - one specific to V-final clauses (V-Φ weak), the other to V2-clauses (V-Φ strong). But such ad hoc multiplication of lexical specifications would be no improvement on the uniform V-to-C theory.

Zwart explores a different idea - that what drives V2 is the presence of a strong "phrasal" feature in a functional head (in AGR_S for SU-V2, in C for X-V2). To implement this, Zwart proposes an extension to the class of triggers: in addition to feature-matching with the relevant phrase, a necessary condition for the successful "checking" of the strong feature is that the functional head be associated with "lexical" material. In some cases, V-movement is the only means of providing a functional head with such lexical support.

Assuming that the N-feature in AGR_S is strong accounts for the appearance of the subject in SPEC,AGR_S - preverbal position in SU-V2, immediate post-C position in X-V2 and V-final clauses. V-raising in SU-V2 is necessitated by a "lexicalization" requirement on the strong N-feature in AGR_S:

(82) AGR_S must be lexically supported.

C has a strong phrasal feature in X-V2 clauses (TOP, WH, etc.). Zwart proposes that these features also need lexical support to permit checking, which triggers V-to-C in X-V2:

(83) C containing a strong operator feature must be lexically supported.

The account presupposes that a lexicalization requirement may be met by an inherent property of the head which bears the strong feature. Subordinate complementizers and finite verbs have the the property in question, root complementizers do not. The exact status of this property is unclear; I refer to it as [+L].

Since (82) and (83) are prerequisites for checking of strong m-features, they must be fulfilled before Spell-Out. But they make movement necessary only if [+L] is not already present in the head in question. This derives the lack of V-to-C in embedded clauses (81b). Where a head lacking [+L] also has no strong phrasal feature, the lexicalization requirement is inoperative. This derives the lack of V-to-C in SU-V2 (81a). [30]

Strong features associated with root complementizers are not supported by an inherent [+L]-feature, so that (83) must be achieved through movement. Raising of the finite verb is the only option. This accounts for X-V2 (80a).

AGR_S also lacks [+L]. There are two ways for it to satisfy (82):

(84) a. Adjunction of AGR_S to C, if it has [+L]
b. Adjunction of the finite verb (always [+L]) to AGR_S

These options are ordered by economy: the option involving the shortest movement path is preferred. Assuming the distance between AGR_S and C to be shorter than that between the finite verb in final position and AGR_S, then option (05a.) will be preferred wherever it is available. [31]

In clauses in which C has [+L], AGR_S can satisfy (82) by raising to C. In embedded clauses, C is inherently specified for [+L], so AGR_S raises to C. This accounts for the neutralization of AGR_S as a trigger for V-fronting in V-final clauses (81c).

In X-V2 clauses, there is an independent trigger for V-raising to C, so C receives [+L] from the verb. Since the verb must move to C anyway, raising to C is the only option for AGR_S (maybe the finite verb picks up AGR_S en route to C). If C does not have [+L], the cheaper movement option is not available. This only arises is SU-V2, where (82) triggers the more costly option of raising of V to AGR_S. This accounts for (80b).

The trigger for V-raising to AGR_S arises by default from the unavailability of AGR_S-to-C. Independent evidence for this proposal comes from the distribution of C-agreement morphemes (see §6.6). Zwart proposes that these morphemes are morphological reflexes of AGR_S in C, so the distribution of C-agreement follows from the proposal about movement of AGR_S to C. The morphemes appear on a lexical head in C, so in the one clause-type where C is not "lexicalized" - SU-V2 - they are absent.

7.2 Problems with [+L]

In seeking a trigger for V-raising to AGR_S in SU-V2, Zwart is led to propose an extension of the notion of "trigger", so that not only "strong" m-features, but also "lexicalization" requirements such as (82/83) can force movement before Spell-Out. However, V-fronting still counts as "early altruism" (cf. §5.4 above); the set of "LF-triggers" is not extended, and can be restricted to m-features as before.

As triggers for move-α, m-features strongly select the element that can "satisfy" them - only the paired element (e.g. a verb with a weak M) can satisfy the trigger (strong or weak M in a functional head). A [+L]-requirement is not selective in this sense, as it can be satisfied by any [+L]-element, so that it becomes possible for a trigger for V-movement to be present without always triggering movement of the verb. V-raising is forced only if "cheaper" solutions are unavailable. Nevertheless, movement in response to a [+L]-requirement can only be "early altruism": Greed ensures that it will have to reflect LF-movement triggered by m-features, excluding such "wild" movements as extraction of a preposition for satisfaction of the [+L]-requirement of AGR_S:

(85) * *Hans mit*[+L]-AGR_S [[PP t *mir*] *einen Kaffee trinkt*]
H. with me a coffee drinks

So the extended trigger set remains highly restricted. In this respect, Zwart's proposal has the desired qualities.

The "lexicalization requirement" on which it is based raises doubts, however, as to its source and scope. Does it apply to all functional heads with strong phrasal features, or only to some? Is the notion "lexical" to be understood as something more than a primitive property (a feature [+L])? Is there a correlate?

Not every head with a strong XP-feature has to be phonologically overt. Two null heads with a strong XP-feature that fail to trigger overt movement are C in English topicalizations and in embedded *wh*-interrogatives in languages like English and German:

(86) a. *John* C *she likes* C = null, TOP+, weak F
b. *(ich weiß) wen* C *sie mag* C = null, WH+, weak F
c. *(I know) who* C *she likes* C = null, WH+, weak F

A zero embedded *wh*-complementizer in German and Dutch has to be inherently [+L], to act as host for AGR_S and so prevent V-fronting. So [+L] does not correlate with "overtness". No other correlate is apparent either. It must therefore be a primitive - a property indirectly relevant to PF, like a strong m-feature, but with the different "unselective" characteristics outlined.

Concerning the scope of the [+L]-requirement, there are two options. If it holds of every head with a strong m-feature, then any such head lacking inherent [+L] either triggers movement or has to move itself. In the English topic construction (86a), C must be inherently specified as [+L]. Alternatively, if [+L]-requirements are specific to certain heads, perhaps C in (86a) is exempted. But inherent [+L] is still needed for the empty C in embedded questions in German (86b), so that it can support AGR_S and block V-fronting. Either way, we end up resorting to lexical stipulation to derive the trigger for V2. In German, embedded C with strong WH (86b) is [+L], but main clause C with strong WH is [-L]. German AGR_S with strong N-Φ is [-L]. English C with strong TOP (86a) is either [+L] or [-L], depending on other decisions.

This problem is noted by Zwart, but left open. As it stands, the analysis is not clearly superior to the symmetric V-to-C solution it replaces. Even granted the necessity for an asymmetric solution, it may be just as plausible to derive V2 in SU-V2 by stipulating that V-features in AGR_S are strong in main clauses and weak in embedded clauses.

7.3 Avoiding Redundancy

The [+L]-requirement is problematic in two respects: it involves an extension of the set of triggers; and as no reasonable correlate for [+L] is forthcoming, the cost in terms of lexical stipulation seems unwarranted. A third objection to the introduction of the L-feature concerns its redundancy with respect to V-features.

Lexical entries for finite C are cross-classified by phrasal features (WH, TOP, or none), and the main/subordinate distinction (there is no TOP-subordinate). Given that Zwart must assume an LF-trigger for V-fronting to C, every entry is specified for weak F. Additionally, specification [+L] / [-L] is needed for the account of overt V-fronting. The feature specification of complementizers looks like (87) in Zwart's account (an m-feature marked "+" is strong): [32]

(87) a. Root:

i	declarative	English:	<F>	[-L]	
		German:	<F>	[-L]	(= SU-V2)
ii	topicalization	English:	<F, TOP+>	[+L]	
		German:	<F, TOP+>	[-L]	
iii	interrogative	English:	<F, WH+>	[-L]	
		German:	<F, WH+>	[-L]	

b. Subordinating:

i	declarative	English:	<F>	[+L]	*that*
		German:	<F>	[+L]	*daß*
ii	interrogative	English:	<F, WH+>	[+L]	*__/whether*
		German:	<F, WH+>	[+L]	*__/ob*

If overt V-movement is related directly to the value of F in C, as in §5.5, the information encoded in (87) can be reduced to a minimum. Given that F is required in any case, the distribution of fronted verbs can be related directly to parametrization of F. [-L] is interpreted as "strong F"; "weak F" replaces [-L], and takes on the function of blocking overt V-fronting in subordinate clauses. Ignoring subject-initial declaratives, we have (88):

(88)

a. Root:				
ii	topicalization	English:	<F, TOP+>	
		German:	<F+, TOP+>	
iii	interrogative	English:	<F+, WH+>	
		German:	<F+, WH+>	
b. Subordinating:				
i	declarative	English:	<F>	*that*
		German:	<F>	*daß*
ii	interrogative	English:	<F, WH+>	*__/whether*
		German:	<F, WH+>	*__/ob*

In the following, I sketch an alternative solution to the SU-V2 question based on (88), that makes no use of the L-feature or an extended notion of trigger.

7.4 F in AGR_S

The simplest way of handling V-fronting in SU-V2 is to assume the same trigger as in X-V2: strong F. To make this compatible with the assumption that the subject only checks Case and Φ-features in SU-V2, we need to assume something like (89):

(89) The node containing AGR_S contains strong F in SU-V2.

The finite verb must adjoin to AGR_S to enable checking of F before PF; and the strong N-features in AGR_S ensure that the subject is in its specifier before PF, giving the desired Spell-Out word order. Since the only trigger for V-fronting is strong F, other V-features are uniformly weak. There is no other trigger for V-fronting to consider. In clauses with V-final order, F in C is weak, and delayed V-fronting is simply a side-effect of Procrastinate.

How does (89) come about? F is a feature of every C in (88). It makes no sense to assume that F is also present in AGR_S in cases other than SU-V2. So if F is in AGR_S only in SU-V2, then the feature content of AGR_S must vary; generally it contains only Φ-features. Lexical stipulation is inadequate: the possibility for variation and the distribution of variants need explanation.

The C-position in SU-V2 has no phrasal feature, no phonological content, and maybe no semantic content either, meaning that F is the sole content of this complementizer (cf. 87a-i). But if F is in AGR_S, there is no need to assume F is in C as well: so C is a category label dominating nothing - a radically empty head. In fact, there is no reason to assume the presence of a C-projection at all. It will play no role in the internal syntactic derivation of the clause. A C-projection might be necessary for the clause to participate in "external" relations. Selectional requirements may dictate that the complement of a "bridge" verb (α in (90a)) be a CP. The parallelism requirement on conjuncts might require α to be CP in (90b.):

(90) a. *Peter hat gesagt,* [$_{\alpha}$ *es sei schön dort*]
P. has said it is nice there

b. [$_{CP}$ *In den Wald ging der Jäger*] *und* [$_{\alpha}$ *(er) fing einen Hasen*]
in the wood went the hunter and (he) caught a hare

Even this type of consideration does not force the assumption of a radically empty C-projection.

The issue rests on concrete assumptions on the nature of syntactic category labels and their relation to content, in particular, syntactic feature content. In the standard view, major category labels (N,A,V,P) are shorthand for combinations of binary lexical features (±N,±V) which cross-classify and define them. Supposing that these features are not involved in the definition of functional category labels, the defining feature of the category C may be F (the common denominator of (88)). [33] So, given (89), the AGR_S-projection of SU-V2 clauses will simultaneously count as a C-projection. There is no need to assume an additional empty projection above AGR_SP: moreover, if clausal conjuncts, complements, etc. must be CPs, then SU-V2 must not contain such a projection, since this is not a projection of an F-feature.

Instead, SU-V2 clauses can be seen as projections of a "mixed" category AGR_S/C:

(91) $[_{CP/AGRsP}$ NP C/AGR_S [...]]

X-V2 clauses differ in having independent C and AGR_S projections. So this analysis is both symmetric and asymmetric. If the fronted verb is "in C" in SU-V2, a unified account of the V2-trigger is possible (I suspect that this point is the main source of resistance to the asymmetric V2 analysis in the literature); pronoun asymmetries motivate the claim that the subject is in SPEC,AGR_S. (91) meets both demands.

7.5 "Mixed Projections": Previous Proposals

The notion of "mixed" projections has many antecedents in the literature. The idea that INFL and COMP are not independently projecting categories in V2 languages has been explored in several variants: e.g. the "CONFL-hypothesis" of Platzack (1986). The present proposal differs in attributing a mixed C/INFL-projection only to SU-V2-sentences - INFL-projections being otherwise independent of the C-projection.

Conceptually closer is the "Matching Projections" hypothesis of Haider (1989). Haider proposes two constraints on syntactic derivations and representations:

(92) a. No empty heads. b. No string-vacuous derivations.

(92a) is violated by the analysis (93a) of subject-initial declaratives in English. The type of derivation banned by (92b) is the *wh*-movement analysis of subject-questions (93b):

(93) a. * $[_{CP}$ e $[_{IP}$ *John will* $[_{VP}$ *come*]]]
b. * $[_{CP}$ *who* e $[_{IP}$ t *will* $[_{VP}$ *come*]]]

Mixed (matching) projections arise through the need to avoid situations where (92) may be violated. Where a projection containing an empty head would otherwise arise, it must be superimposed on the projection it dominates, in order to avoid violation of (92a). So in (93a), the C-projection must conflate with the Infl-projection (headed by <u>will</u>) to yield a well-formed representation (94a):

(94) a. $[_{CP/IP}$ *John* $[_{C/I}$ *will*] $[_{VP}$ *come*]]
b. $[_{CP/IP}$ *Who* $[_{C/I}$ *will*] $[_{VP}$ *come*]] ?

In subject questions, where C is also an "empty head", movement of the subject *wh*-phrase to SPEC,CP would be a string-vacuous movement. The only way to avoid violation of (92b) is for the C-projection to be superimposed on the I-projection, as in (94b).

In non-subject-questions, fronting of the *wh*-phrase is not string-vacuous, and may take place: and in order to avoid a violation of (93b.), raising of Infl to C (or insertion of do) becomes necessary:

(95) $[_{CP}$ *Who will* $[_{IP}$ *John* t $[_{VP}$ *visit* t]] ?

In this way, both the Vacuous Movement Hypothesis of Chomsky (1986), and the absence of do-support in English subject questions, are derived by (92). [34]

In the present proposal, I adopt neither of the conditions (92). Both empty heads (such as C in English topicalizations, which Haider analyses as adjunction to IP) and "string-vacuous" movement are allowed. What I do adopt is the basic idea that projections conflate where independent projections would be unnecessary.

Heycock & Kroch (1994) apply a similar idea to the analysis of SU-V2 in V2-languages. They suggest that raising of a verb to C and its subject to SPEC,CP yields redundant structural representation of "licensing relations"; Case, Agreement and Predication relations between the subject and the verb in the C-projection are doubled in the I-projection by their respective traces:

(96) $[_{CP}$ NP $[_{C}$ V+INFL] $[_{IP}$ t_{NP} $[_{VP}$ ...] t_{INFL}]]

Assuming deletion of redundant traces to be possible, and that deletion of a head leads to deletion of its phrasal projection, an "economy" principle ("each licensing relation is instantiated only once") ensures that the "empty" I-projection must delete, yielding (97):

(97) $[_{CP}$ NP $[_{C}$ V+INFL] $[_{VP}$ ...]]

Since INFL is present on the verb in C, the C-projection can be thought of as "doubling" as an I-projection with respect to licensing relations. Where a non-subject occupies SPEC,CP, C and I participate in different licensing relations, and so the trace in I may not delete.

Like others before it, this proposal offers no insight into Zwart's original problem of why the subject must raise to SPEC,CP to create the "double structure" (96) in the first place. So I shall not pursue the idea that conflated projections result from deletion in the derivation.

7.6 Economy and Project-α

Instead, I suggest that mixed projections are "base-generated" - that they arise at the point in the derivation where items from the lexicon are mapped to phrase-structures by Project-α (see §5.1).

As a derivational operation, "Project-α" is governed by economy. Only so many head positions can be created in a structure as there are items from the lexicon to be accommodated. The clause structure (44) (§5.1 above) is not rigid; extra X°-positions may needed to house "optional" functional heads such as negation or auxiliaries in periphrastic constructions. However, Last Resort forbids projection of an "extra" position if there is no item to be inserted.

Representations like (91) violate a fundamental assumption about the relation between categories and nodes in trees, namely, that each category is associated with an independent node. In the Minimalist model this translates as a restriction on projection: project-α operates on one lexical item at a time. If we loosen this assumption to allow a

single operation of project-α to access a set of more than one lexical item and project a tree from that set, then mixed projections like that in (91) will be possible. The projection of mixed categories will be automatic, to ensure licensing with respect to X'-theory. [35]

Under this perspective, it is not necessary to resort to multiple lexical stipulations for AGR_S to derive (89). Rather, a single lexical entry can be maintained (strong N-Φ, weak V-Φ). The entry for unmarked declarative C contains only strong F (this will form the missing entry in (88)).

The new assumption is:

(98) Project-α can apply to any number of lexical items at once.

The factor that forces project-α to apply to C and AGR_S simultaneously in SU-V2 is economy. Since it involves less derivational steps (applications of Project-α), joint projection of several items under one node is "cheaper" than individual projection of each item under a new node. Applying the logic of Last Resort, we get (99):

(99) Economy of Projection
Minimise applications of Project-α
i. Joint projection wherever possible.
ii. Projection of a single item under one node is a more costly option to be used only where necessary.

Given (98) and (99), the issue arises of what conditions make joint projection impossible and independent projection necessary.

If two lexical items with phonological form are inserted under a single node, there will be a linearization problem: words cannot be simultaneously uttered. So it can be assumed that joint projection of overt elements is precluded. [36] Functional heads such as AGR_S and main clause C lack PF-forms, and so escape this constraint.

The factor that forces projection of AGR_S before C in X-V2 is X'-theory: a head position only licenses a single specifier position. With respect to feature-checking, specifiers are crucial:

(100) A head has only one phrasal checking position.

If AGRS containing N-Φ and C containing TOP or WH project simultaneously, then one of the phrases is unable to check, since the other occupies its position. In X-V2, the presence of phrasal features in both heads forces the more costly option of independent projection.

If it not further constrained, this system has rather drastic consequences. Consider the options for analysing a simple sentence like (101):

(101) *He came.*

Joint projection of the categories of this sentence under a single X'-projection would be consistent with the terminal string (this is the structure (101) would have under Haider's proposal):

(102) $[_{XP}$ *he* $[_{X'}$ $[_{X^\circ}$ *came*]

The category X would be a conflation of all functional heads (C,AGR,T) with the lexical category V. The subject and the verb stand in a checking configuration, so all features can be checked in one X'- projection.

Potentially, problems arise if (102) is permitted: to name one, if adverbs such as completely are left-adjoined to VP (see Pollock 1989), then it is unclear why (103) should be ungrammatical:

(103) * *Completely John failed.*

To avoid this type of problem, I propose a further constraint on projection:

(104) Major category items must project independently.

This ensures that a verb may not check its m-features "in situ" by conflating its projection with functional heads; there must be at least one functional projection encasing the VP and the verb must raise out of VP to check its features. Subjects and objects, arguments of the verb which must be projected in VP to be assigned Ø-roles, must also raise to check m-features. So the analysis (102) is excluded. [37]

Overt complementizers lacking phrasal features (that/daß) do not project together with AGR_S in subordinate clauses - the subject occupying SPEC,AGR_S (to the left of AGR_S) at Spell-Out follows the complementizer in word order. So I propose to strengthen the condition on projection of overt items to (105): [38]

(105) Items with phonological content must project independently.

In other words, only empty heads may conflate.

The options that remain are rather restricted. In SU-V2, C and AGR_S are both functional items, neither having phonological content, so (104/105) are met. C has no phrasal feature, so joint projection creates no problem with respect to (100). Since independent projection of AGR_S before C is not necessary, the cheaper option of projecting both items at once is forced. So the fact that the subject appears before the finite verb in Spell-Out in SU-V2 reflects the pressure of economy principles on derivational operations of projection as well as movement.

In X-V2, where both C and AGR_S have a phrasal feature, C and AGR_S must project independently, as the two heads need to license different phrases. Given that if C projects independently, it projects above AGR_S, the subject follows C in X-V2, while the *wh*-phrase or topic precedes C. The finite verb precedes the subject in Spell-Out since strong F is a feature of C. [39]

In fact, (100) only requires independent projections for C and AGR_S in X-V2 when the phrase that checks the phrasal feature of C is not the subject that checks the N-features of AGR_S. Where the phrase that checks TOP or WH is the main clause subject, then C and AGR_S may conflate. The prediction is that these cases should display characteristics of SU-V2 and not of X-V2, for example with respect to the agreement morphology facts mentioned in §6.6. As far as I am aware, this is borne out; in Dutch, a stressed (2nd person) pronoun subject, compatible with a TOP-analysis, cannot induce the "C-agreement" inflection on the finite verb.

This view of subject-initial topicalizations and questions is roughly equivalent to the Vacuous Movement Hypothesis of Chomsky (1986), according to which subject-phrases that only need to raise to the immediately preceding SPEC,CP do not do so in overt syntax. However, Chomsky assumed that *wh*-phrases must raise at LF. In the present

proposal, such phrases do not need to move anywhere at LF, since they already occupy the position in overt syntax in which they satisfy all triggers.

This conclusion has consequences for the status of the A/A' distinction: in the standard view, SPEC,IP cannot be an A'-position; and *wh*-phrases need to occupy an A'-position at LF. In the present model, *wh*-phrases must occupy a position in the checking domain of a *wh*-feature. Where C and AGRS conflate, this position happens to be the specifier of a head also containing N-Φ features (i.e. an A-position). In other words, the set of A-positions is not disjoint from the set of A'-positions.

I have motivated the proposal (98/99) as a way of handling the SU-V2 problem. It remains to be seen whether it supports a wider range of data. Mixed projections in root clauses are expected to arise in non-V2 languages, as well. [40] The prediction for English sentences like (101) is that although V and maybe T must project independently, AGR_S and C may, and therefore must conflate. Neither head has phonological content, and only one phrasal feature needs to be checked:

(106) $[_{C/AGRsP}$ *he* $C/AGR_S\ [_{TP}\ [_{VP}$ *came*]

So my proposal converges with Haider's on the analysis of examples with auxiliaries (94a).

The asymmetry between subject-questions and non-subject questions with respect to do-support (cf. the discussion of (94b) vs. (95)) also invites analysis in terms of conflation. The cases where inversion is required correspond with the cases in which C must project independently of AGR_S. However, it is expected that the trigger for AUX-to-C (strong F) associated with the *wh*-complementizer should be present also where C and AGR_S conflate in subject-questions, whereas subject-questions pattern with declaratives with respect to do-support. In both, do appears only under negation and emphatic denial of negation:

(107) a. *Who didn't come*?
b. *Who DID come*?

So do-support in questions cannot purely be a response to the presence of strong F. Instead, the asymmetry is reminiscent of the asymmetry in the distribution of C-agreement morphemes, which as discussed in §6.6., do not appear on verbs in SU-V2 (cf. also note 29). The issue also depends on the analysis of do-support itself, which would take us too far afield here.

Notes

* Parts of this material were presented at the colloquium "Structural and Procedural Aspects of Verb Placement" in Hamburg, December 1992, and at the 6. Wuppertal Linguistics Colloquium, October 1993. I thank the organizers and other participants of these meetings for stimulating discussions. Thanks also to Manfred Bierwisch, Ewald Lang, and Sten Vikner for helpful comments.

1. The assumptions in (4) are not shared by other frameworks: (4b) in particular is not assumed in HPSG-treatments of V2, or in categorial grammar.

2. There are nevertheless cases in German which seem to violate (9): parentheticals and discourse particles like aber ("but", "however") may intervene between specifier and verb to induce V3-order. I have nothing to offer on these cases:

i) *Niemand, wie es scheint, hat mich gesehen*
noone as it seems has me seen

ii) *Ich aber werde nicht da sein.*
I however will not there be

3. Bare V2 clauses function as complements only in German. Other V2-languages, e.g. Swedish, permit the V2-pattern directly following an overt complementizer in complement clauses - see Holmberg (1987). I ignore these cases here. The fact that the pronoun er must be disjoint in reference from the name Hans in (27) may be taken as evidence for the claim that the German construction involve genuine embedding, and not parataxis. See Schwart & Vikner (1989), Vikner & Schwartz (1991) for further arguments. Reis (1994) has recently cast doubt on this claim.

4. For other proposals, see Holmberg (1987) and the papers in Haider & Prinzhorn (1986).

5. In Rizzi & Roberts (1989) and Wilder (1989), it is suggested that root complementizers of V2-languages are affixes that need to attach to a finite verb. (29) thus represents a violation of the principle barring unbound affixes. The weakness of this idea lies in the need to assume the existence of an affix with no PF-reflex other than the movement it is intended to explain. However, by identifying the trigger for V2 as an inherent morphological property of the root complementizer, the proposal has the advantage of being compatible with the lexical parametrization hypothesis. The proposal is revived in §5.5 under a somewhat different interpretation, lacking the drawback noted, in the context of Chomsky's (1993) proposals about the triggering of movement.

6. This case arises where a VP dominates another VP and the head of the lower VP raises to the head of the higher VP, as in the multiple VP-analysis of double complement structures of Larson (1988).

7. Adjunction of heads to other heads containing lexical material is often proposed to describe processes of incorporation. While many cases involve a target containing affixal material, others patently do not - for example, Verb Cluster formation in Dutch and German. There is no independent motivation for stipulating a ban on the head-adjunction of a verb to an overt non-affixal complementizer.

8. Croatian is a pro-drop language. For arguments that (37) involves raising of both finite auxiliary and participle to C, see Wilder & Cavar (1994a, 1994b). Several authors have argued that the HMC should be loosened to permit cases of "long" verb-movement analogous to (36a) - see Rivero (1991) and references cited there. This would mean that the explanatory scope of Last Resort is even broader.

9. The model incorporates the VP-internal subject hypothesis, whereby subjects are introduced into the derivation inside the VP.

10. Movement by substitution in the structure-building stage - the overt syntax prior to Spell-Out - is assumed to underly an "extension" requirement: i.e. the new node created by move-α must be the root node of the output. Movement by adjunction (e.g. raising of V to AGR_S) will create a new node labelled with the category of the category adjoined to. Adjunction is not subject to the "extension requirement"; and head-adjunction is necessarily not "tree-extending", as the adjunction site, a head, cannot be the root of a complex tree.

11. Since both the verb-Infl relation and "Case-assignment" are handled in terms of checking theory, inflected verb forms, determiners, etc. must be formed in the lexicon and project into syntax already bearing m-features. This contrasts with earlier approaches in which inflected verbs are created in the syntax by adjunction of a verb-stem to inflectional affixes in an Infl-node; or in which DPs only receive Case-features by "assignment" in the course of the derivation.

12 . If Φ-feature correspondence between subject and finite verb were not mediated by a functional head, then checking would be possible inside VP (where verb and subject DP are already in the required configuration) without movement.

13 . Case-checking is assumed to take place in Agreement projections, following raising of the verb to AGR_O, and T to AGR_S.

14 . See also Pollock (1989). The adverb marks the boundary of VP.

15 . Cf. Wilder & Cavar (1994b) for discussion. The model outlined has considerable explanatory potential with regard to the acquisition problem. Variation follows from parametrized morphological properties of a restricted class of lexical items (whose idiosyncratic properties must be "learned" in any case), which are "visible" in, and hence deducible from, PF-representations. It remains to be seen whether such a restrictive theory of triggers and parametrization can be maintained in the face of a wider range of data. See also discussion of Zwart (1993) in §7. below.

16 . The fact that the DP John appears in a PP, where it presumably checks "objective" Case, should not in principle prevent it from checking the NOM-feature in Infl as well. Where a phonological form is not specific to a particular Case, it is reasonable to suppose that the associated abstract Case-feature is underspecified. So the Case-feature of John is compatible with, and so can check, either a NOM or an ACC-feature. Under a minimalist conception of the lexicon, it makes sense not to assume two forms John, one bearing NOM, one bearing ACC.

17 . Nominatives in there-constructions and "in situ" *wh*-phrases are assumed to raise at LF. On how feature-checking proceeds in these cases, cf. Wilder & Cavar (1994b).

18 . See also Zwart (1993). There is no overt morpheme that signals a topic phrase in German or other V2 languages, but the existence of special "topic"-markers in languages like Japanese may be claimed to support the existence of the abstract underlying feature.

19 . If C had a strong F-feature matched by F in some abstract Infl-head, and this head contained no strong V-feature, then the overt movement (I-to-C) would have no overt reflex in PF-representations. It would not be deducible from word-order alternations or any other property of PFs, and so the parameter setting would be unlearnable.

20 . For analysis of do-support and the failure of finite main verbs to raise to C in English, see Wilder & Cavar (1994b). It may be necessary to distinguish a further C to account for inversion in English declaratives with preposed negative quantifiers, only-phrases, etc.

21 . Kayne (1993) has developed a theory in which a universal order for all X'-projections - Specifier-head-complement - is imposed by UG. If true, this suggests that there can be no variation in the order of a head and its complement (or specifier), and the assumption that the OV-order in German VP is "base-generated" - cf. §2 - has to be given up. This raises the prospect that all aspects of word order variation are due to differences in the way move-α feeds PF-representations.

22 . The symmetric analysis is strictly neutral with regard to the ordering of Infl and VP. Note that the asymmetric analysis does not claim that SU-V2 clauses are IPs lacking a C-projection. Zwart (1991,1993) is careful to state that SU-V2 clauses are CPs; as such they do not differ categorially from X-V2 clauses, but only in that their C-projection is "not activated".

23 . The asymmetric analysis may have consequences for the account of V-third effects in SU-V2 of the format X-SU-V (Damir Cavar, p.c.). SU-V2 does not differ from other V2 clause-types in this respect: X may only be a left-dislocated phrase. However, there is one additional adjunction site available, AGR_SP, which is sometimes assumed to be available for adjunction in embedded clauses, in German at least, yielding "scrambling" orders like (i):

i ...*daß* [*dieses Buch* [*ich nie* t *lesen würde*]]
that this book I never read would

Movement and adjunction of a phrase to CP in X-V2, is completely out (ii), and parallel SU-V2 examples are equally as bad:

ii * *dieses Buch nie würde ich* t *lesen*
iii * *dieses Buch ich würde nie* t *lesen*

Excluding this last example might be a problem for the asymmetric theory, as a preverbal IP-adjunction site is available. Under the version of the asymmetric analysis developed in §7, the AGRS-projection in SU-V2 is simultaneously the C-projection, so this problem may be avoided. However, "scrambling" is a phenomenon that raises quite different questions for the Minimalist model, so I shall leave this issue open here.

24. Zwart (1991) develops a further argument on the basis of properties of "conjunction reduction" in Dutch and German, which I do not go into here.

25. In fact, expletive es only appears in initial position in SU-V2 sentences. It is usually assumed to be replaced in other contexts by a null expletive. The question of why the overt expletive es does not appear in SPEC,IP except in SU-V2 sentences needs an independent account. Given that the data single out SU-V2 as being special, the asymmetric approach to V2 provides basis for an account in terms of the different position of the finite verb. An alternative strategy might be to view es in SU-V2 as an expletive topic, which would make these data compatible with a "generalized topic" analysis of declarative V2.

26. Rizzi's proposal is to define A-positions as all and only θ-marked positions and all and only specifier positions that agree with (some element) in the head ("...a Spec is A when construed (coindexed) with an Agr specification in its head", further sharpened to "construed with agreement in phi-features".) A similar idea to Rizzi's is to be found in Holmberg (1987).

27. In Chomsky (1993), A-positions are reinterpreted as specifiers of "L-related heads", including AGR_S and AGR_O, but not C.

28. The following comments are based on helpful discussion with Gerhard Jäger, from whom I have borrowed (77). See also Jäger (in prep.).

29. English has an analogous asymmetry that turns up with the 1SG form of be with suffixed negation. The form aren't is available only in inversion constructions (AUX in C) (i); without inversion, only the form am is possible, which cannot take suffixed 'nt. See Kayne (1989) for discussion:

i Why aren't I allowed to do that?
ii I'm not allowed to ...
iii I am not allowed to ...
iv * I aren't allowed to ...
v * I am'nt ...

30. Zwart's proposal is thus an implementation of the earlier idea that "lexical" C blocks V-fronting in which the blocking effect is derived not via the substitution assumption but via economy principles, as I argued above (§4.3) that it should be.

31. The "shortest path" condition is another subcase of Last Resort. Zwart claims that the finite verb in clause-final position is separated from AGR_S by at least two heads - AGR_O and T, which he assumes precede their complements in Dutch and German.

32. Root complementizers are "PF-null": the choice between null and overt (whether/ob) realization in embedded interrogatives usually depends on whether the *wh*-phrase in SPEC is overt or covert.

Extension to relative clauses may require additional features; see Rizzi (1990:Ch 2) for relevant discussion. The choice of L-specification for the English topic-C depends on the scope of the [+L]-requirement, as discussed.

33. If F is binary, then +F defines finite C and -F non-finite C, which heads infinitive CPs. Not wanting finite verbs, which also bear F, to be defined as complementizers, I suppose that substantive features override "functional" or "inflectional" features in determining category labels.

34. Haider applies the idea that if one projection must be superimposed on another, but no well-formed match is found, then the string is ungrammatical. He uses this to account for the ungrammatical V-projection topicalizations (i) discussed in §6.5.

i * *getanzt wurde*
danced was

Together, V-to-C and VP-fronting are string-vacuous with respect to subordinate clause order:

ii *...weil getanzt wurde*
since danced was "since there was dancing"

In (ii), according to Haider, the verb getanzt must form a complex head-constituent with wurde in VP, while in (i), it must be contained in a phrase (VP) in SPEC,CP excluding the finite verb in C. So no match is found and (i) is impossible.

iii (*weil..*) $[_{VP}\ [_{V}$ *getanzt wurde* $]]$ / $[_{CP}\ [_{VP}$ *getanzt* $]\ [_{C'}$ *wurde* _ $]]$

Adding more material to the right of the finite verb renders topicalization and V-to-C non-string-vacuous, so matching is unnecessary, and the example becomes grammatical:

iv *Getanzt wurde hier noch nie.*
danced was here yet never

Counterevidence to Haider's proposal is found in grammatical V2-clauses with passives of the format "V + AUX + extraposed clause" (v), which are string-vacuous with respect to subordinate clause order (vi):

v *Untersucht wird, ob die Daten stimmen.*
investigated is whether the data are-correct

vi *... weil untersucht wird, ob die Daten stimmen.*

If the participle and the auxiliary must form a V-constituent in (vi), there is no matching projection CP/VP for (v); but (v) is well-formed. In §6.5, the contrast (i/iv) is attributed to the requirement that C' (the comment in topicalizations) must contain a focus, given that auxiliaries have no focussable content. Where the topic is not a V-projection, the comment minimally contains a contentful main verb, which can be "comment-focussed" if need be:

vii *Mich friert*
me-ACC freezes "I am freezing"

35. A mixed projection does not violate the principle that a phrase has a unique head: C and AGR_S do not represent two configurationally independent heads of a single phrase. Nor do they represent head-adjunction structures, in which only the head X of the complex $[_X\ Y\ X\ ]$ projects. Rather, the categories C and I jointly label a single node, and this (joint) category label is projected to the non-head projections of the phrase.

36. Two or more words can be inserted under a single node if they form an adjunction structure, since this defines a linearization. This may be the case for compounds; but such cases are irrelevant if the adjunction structure is already formed in the lexicon.

37. (100) already ensures that AGR_S and AGR_O, which license different phrases in transitive sentences, cannot conflate with each other. I leave the question open of whether T can conflate with an AGR-head.

38. Given that major category elements generally have phonological content, it may be possible to reduce (104) and (105) to a single condition. It is unclear to me how they might be explained.

39. Note that the order of projection of individually projected functional elements requires an independent explanation, just as in a model that does not incorporate (98). Given a sequence of projection, only heads that are adjacent in the projection sequence can conflate.

40. The facts discussed in Ouhalla (1992) may be relevant here.

References

Bayer, Josef (1984) "COMP in Bavarian Syntax", The Linguistic Review 3.

Borer, Hagit (1984) Parametric Syntax. Case Studies in Semitic and Romance Languages, Foris, Dordrecht.

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

Chomsky, Noam (1986) Barriers, MIT Press.

Chomsky, Noam (1991) "Notes on Economy of Derivation and Representation", in: Robert Freidin (ed.) Principles and Parameters in Comparative Grammar. MIT Press, Cambridge MA.

Chomsky, Noam (1993) "A Minimalist Program for Linguistic Theory", in: Ken Hale and Samuel Keyser (eds.) The View from Building Twenty, MIT Press, Cambridge MA.

Den Besten, Hans (1983) "On the Interaction of Transformations and Lexical Deletive rules" in: W. Abraham (ed.) On the Formal Syntax of West Germania, John Benjamins, Amsterdam.

Haegeman, Liliane (1990), "Subject Pronouns and Subject Clitics in West Flemish", The Linguistic Review 7, 333-364.

Haider, Hubert (1989) "Matching Projections", in: A. Cardinaletti, G. Cinque, & G. Giusti (eds.) Constituent Structure: Papers from the 1987 GLOW Conference, Venice, Foris, Dordrecht, 101-122.

Haider, Hubert (1990) "Topicalization and Other Puzzles of German Syntax", in: G. Grewendorf & W. Sternefeld (eds.) Scrambling and Barriers, Benjamins, Amsterdam.

Haider, Hubert & Prinzhorn, Martin (1986) Verb Second Phenomena in the Germanic Languages, Foris, Dordrecht.

Heycock, Caroline & Kroch, Anthony (1994) "Verb movement and coordination in a dynamic theory of licensing ", The Linguistic Review 11, 257-283.

Holmberg, Anders (1987) Word Order and Syntactic Features, PhD, Stockholm.

Jäger, Gerhard (in prep.) "Topik, Fokus und der Stadien-/Individuen-kontrast", Arbeitsgruppe Strukturelle Grammatik, Berlin.

Kayne, Richard (1989) "A Note on English Agreement", Ms., CUNY.

Kayne, Richard (1993) "The Antisymmetry of Syntax", Ms. CUNY.

Koopman, Hilda (1984) The Syntax of Verbs, Foris, Dordrecht.

Larson, Richard (1988) "On the Double Object Construction" Linguistic Inquiry 19, 335-391.

Law, Paul (1991) "Verb Movement, Expletive Replacement, and Head Government", The Linguistic Review 8, 253-285.

Ouhalla, Jamal (1991) Functional Categories and Parametric Variation, Routlege, London.

Ouhalla, Jamal (1992) "Subject Extraction, Negation and the Anti-Agreement Effect", Natural Language and Linguistic Theory

Platzack, Christer (1986) "The Position of The Finite Verb in Swedish", in Haider & Prinzhorn (eds.), 27-47.

Pollock, Jean-Yves (1989) "Verb Movement, Universal Grammar, and the Structure of IP", Linguistic Inquiry 20, 365-424.
Reis, M. (1994) "Gibt es Extraktion aus Verbzweitsätzen im Deutschen?", talk given in MPG, Berlin.
Rivero, Maria-Luisa (1991) "Long Head Movement and Negation: Serbo-Croatian vs. Slovak and Czech", The Linguistic Review 8, 319-351.
Rizzi, Luigi (1990) Relativized Minimality, MIT Press, Cambridge MA.
Rizzi, Luigi (1991) "Proper Head Government and the Definition of A Positions", GLOW Newsletter 26, 46-47.
Rizzi, Luigi & Roberts, Ian (1989) "Complex Inversion in French", Probus 1, 1-30.
Schwartz, Bonnie & Vikner, Sten (1989) "All Verb Second Clauses are CPs", Working Papers in Scandinavian Syntax 43.
Shlonsky, Ur (1994) "Agreement in Comp", The Linguistic Review 11, 351-375.
Travis, Lisa (1984) Parameters and Effects of Word Order Variation, PhD, MIT.
Vikner, Sten & Schwartz, Bonnie (1991) "The Verb Always Leaves IP in V2-Clauses", Ms., Geneva.
Wilder, Chris (1989) The Syntax of German Infinitives, PhD, London.
Wilder, Chris & Damir Cavar (1994a) "Long Head Movement? Verb Movement and Cliticization in Croatian" *Lingua* 93, 1-58.
Wilder, Chris & Damir Cavar (1994b) "Word Order Variation, Verb Movement and Economy Principles" *Studia Linguistica* 48, 46-86.
Zwart, C. Jan-Wouter (1991) "Subject deletion in Dutch: a difference between subjects and topics", in: Kas, Marc, Reuland, Eric & Vet, Co (eds.), Language and Cognition 1, University of Groningen.
Zwart, C. Jan-Wouter (1993) "Verb Movement and Complementizer Agreement", in: Jonathan Bobaljik & Colin Philips (eds.) Papers on Case and Agreement I (MIT Working Papers in Linguistics Vol 18), 297-341.

FAS Papers in Linguistics is intended to reflect the on-going work at the Center. It publishes papers by its staff members and visiting scholars. The publication is available on an exchange basis. For further information, please write to:

The editors
FAS Papers in Linguistics
Jägerstr. 10/11
D-10117 Berlin
Germany

Telephone: +49 30 20 19 24 04
Fax: +49 30 20 19 24 02

E-mail: faspil@fas.ag-berlin.mpg.de

The next issue of FASPIL on Optimality Theory includes the contributions from Gaberell Drachman, Ursula Kleinhenz, Sylvia Löhken, André Meinunger, Ad Neeleman/Peter Ackema, Renate Raffelsiefen, Iggy Roca and Wolfgang U. Wurzel.

Table of Contents