Specificity-driven Syntactic Derivation

Antje Lahne

1 Introduction

Specificity is arguably one of the main basic concepts of morphological theory. It is typically invoked in form of the Subset Principle (e.g. Halle 1997) to resolve competitions between the markers of a language which arise due to lexical underspecification of inflectional markers.[1] The Subset Principle has the effect that whenever a more specific marker competes for a syntactic context with a less specific marker, then it is always the more specific marker that is inserted.

It seems to be a recurring pattern in syntactic derivations, too, that structure-building operations are ordered in such a way that elements that are more 'specific' in one sense or another are are prefered by the probing element over less specific elements. This pattern emerges in a number of seemingly unrelated environments, such as long-distance agreement (e.g. Polinsky & Potsdam 2001, Bobaljik & Wurmbrand 2005, Bhatt 2005) and Mahajan's Generalisation (Mahajan 1994, 1997, Sauerland 1995, Bittner & Hale 1996), which were previously accounted for by independent principles.
The goal of this paper is to propose a new approach to syntactic structure-building that accounts for these data in a uniform way. The underlying hypothesis is that specificity is a much more powerful underlying syntactic principle than was previously thought. I propose that syntactic derivations are driven by a specificity principle on merge: when more than one constituent is a candidate for an agreement relationship or displacement operation, the operation will involve the more specific constituent. This specificity principle yields most effects of the Minimal Link Condition (MLC), apart from the superiority case, which is shown to be an phase effect. In addition, specificity also accounts for a number of Anti-MLC effects, such as oder-preserving movement and anti-superiority effects, which were previously accounted for by independent, partially overlapping principles. The MLC and related locality conditions (Equidistance, A-over-A-Condition) can therefore be dispensed with in such a system.

2 Theoretical Background

Following Manzini (1994), Epstein & Seely (2002), Müller (2004) and Lahne (2008c), I adopt the null hypothesis that each phrase is a locality domain; in other words, each phrase is a phase. Consequently, the search space of the current head X at a stage of the derivation comprises the next lowest head Y and the edge of Y, but not the material

1. Several version of this principle are known in Morphology and Phonology as *Emergence of the Unmarked, Specificity Condition, Elsewhere Principle, Proper Inclusion Principle, Blocking, Paṇīnī's Principle, or Proper Inclusion Condition*, among others; e.g. Kiparsky (1973), Di Sciullo & Williams (1987), Anderson (1992), Lumsden (1992), Williams (1994), Halle (1997), Stump (2001).

embedded under Y. Furthermore, this analysis employs a clausal structure consisting of the core categories C – I – v – V.

The underlying argument structure is such that v assigns internal case, and I assigns external case. The internal argument is merged as the sister of V, the external argument as Spec,v (Murasugi 1992, among others).

I propose that syntactic derivations are driven by a specificity principle on merge. The underlying principle is given in (1).

(1) *General Specificity Principle:*
A Probe undergoes a syntactic operation with the most specific matching goal. Specificity is determined by cardinality of morpho-syntactic features: a set Q is more specific than a set H iff $|Q| > |H|$.

This principle is an extension of the *Maximize Matching Effects* constraint.

(2) *Maximize Matching Effects (MME; Chomsky 2001:15):*
If local (P,G) match and are active, their interpretable features must be eliminated at once, as fully as possible; partial elimination of features under Match, followed by elimination of the residue under Match, is not an option.

Principle (2) operates at the points when the currently active head could potentially deal with its unhandled features in several ways. It has the effect that an agreement relation between a probe and a goal involves handling the maximal number of matching features.[2] This becomes crucial when there is more than one potential action that the current phase head can trigger. For example, in (3), the C head contains a movement-inducing [•wh•]-feature.[3] Wh-movement can be decomposed into an Agree relation (Agree $C_{[wh]}$, $what_{[wh]}$) and a Move operation (Merge C', *what*). One possibility is now that C agrees with the wh-element *what* in the feature [wh], but then moves another constituent, for instance the subject, to its edge, while the wh-element stays lower in the structure. This, however, results in an unacceptable sentence, as (3b) shows. The desired result is that C agrees with *what* in [wh] and then attracts *what* to its edge, as shown in (3a).

(3) a. What had the boy seen?
b. The boy (*what) had (*what) seen (*what)? (No echo question)

This result is only possible if the concurrence of syntactic operations is forced by an underlying specificity principle such as the MME.

The effect of the GSP is identical to that of MME in constellations where a *single* potential Goal for an active Probe is involved: Agree between P and G involves handling of the maximal number of matching features. However, MME and the GSP differ in one crucial respect: MME does not have an effect on the choice of Goal if there is more than one

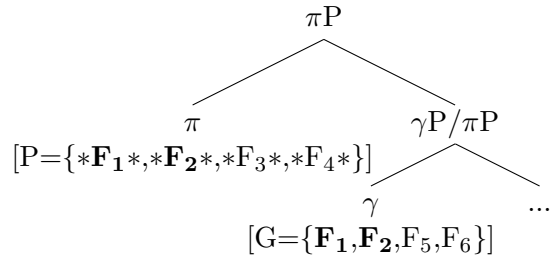
2. The MME Principle is a more concise version of the Earliness Principle (Pesetsky 1989), which is an independently well motivated assumption underlying minimalist derivations.

(i) *Earliness Principle:*
An uninterpretable feature must be marked for deletion as early in the derivation as possible (Pesetsky 1989, Pesetsky & Torrego 2001).

3. The notation of structure-building features as bulleted is adopted from Heck & Müller (2006).

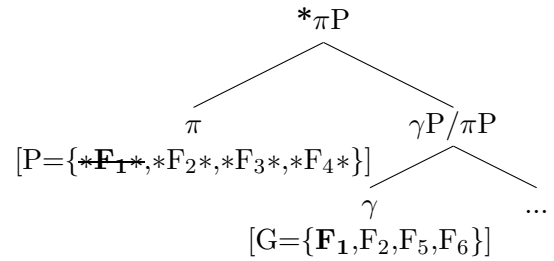
potential Goal in the search space; due to the Minimal Link Condition (MLC), it is invariably the closest goal that the Probe agrees with. The GSP, on the other hand, has the additional effect that with more than one potential Goal being available in the search space of a Probe, the Probe agrees with the Goal that has the highest number of matching features. These two cases are best demonstrated by means of an abstract example. In a possible derivation, there is a goal γ in the search space of a probe π, and $\{*F_1*,*F_2*\}$ contained in π's feature set P match γ's features $\{F_1,F_2\}$.

(4) *Example I: "Classical" Maximize Matching Effects case*

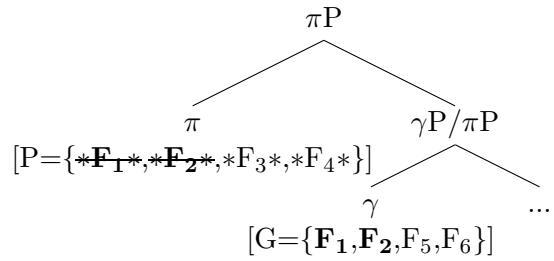


An agreement relation involving only $*F_1*$ and F_1 is not an option in this constellation. Thus, (5) is an impossible derivational step; the agreement relation must involve both $*F_1*$, F_1 and $*F_2*$, F_2, as shown in (6).

(5)

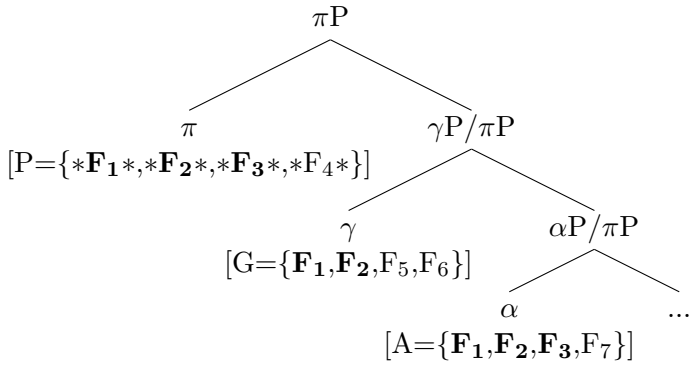


(6)



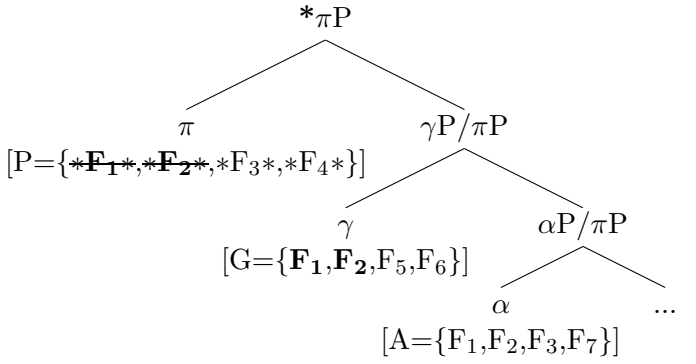
In this example, the GSP and MME have the same effect. However, the two principles differ in their effect in contexts where more than one potential Goal is available in the search space of a Probe. In derivations obeying MME, it is always the closest Goal that the Probe agrees with, due to the the Minimal Link Condition. The GSP, on the other hand, can overwrite the MLC: If the GSP is assumed to be at work, then the Probe agrees with the Goal that has the highest number of matching features, no matter if it is is the closest available goal or not. This is demonstrated in a second abstract example: In a second possible derivation, there are two potential goals, γ and α, in the search space of π. The closer goal γ has two matching features for π, and the more remote goal α has three matching features.

(7) *Example II: A case for the General Specificity Principle*

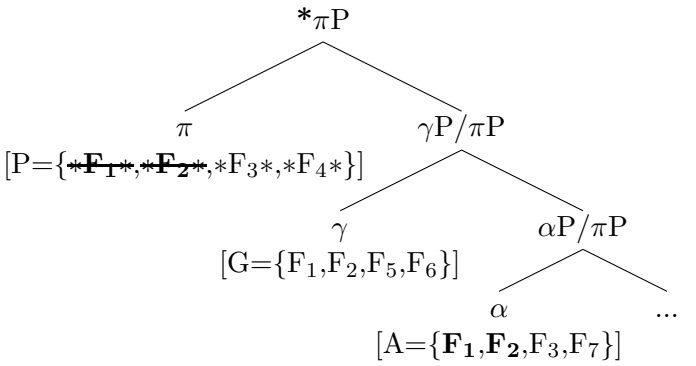


In this constellation, feature handling between π and the closer goal γ is not an option (see (8)). Likewise, feature handling involving only $*F_1*$, F_1 and $*F_2*$, F_2 is not valid, as feature handling must involve the greatest possible number of features that can be dealt with (see (9)). Thus, in a derivation obeying the GSP, the agreement relation must involve $\{*F_1*,*F_2*,*F_3*\}$ of π and $\{F_1,F_2,F_3\}$ of α; this is shown in (10).

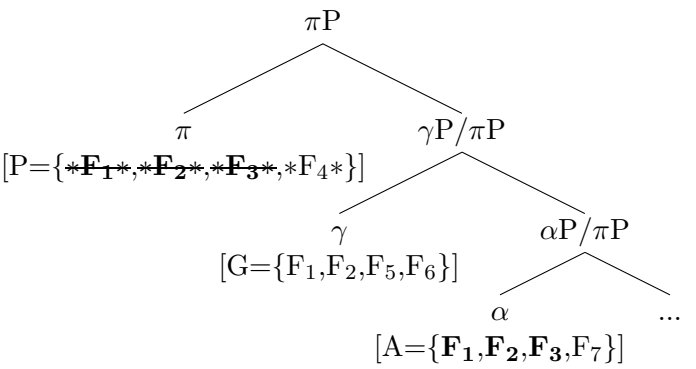
(8)



(9)



(10)



3 Deriving MLC Effects

In what follows, I sketch a possible way in which the effects of a locality constraint – the Minimal Link Condition –, can be derived by means of the General Specificity Principle.

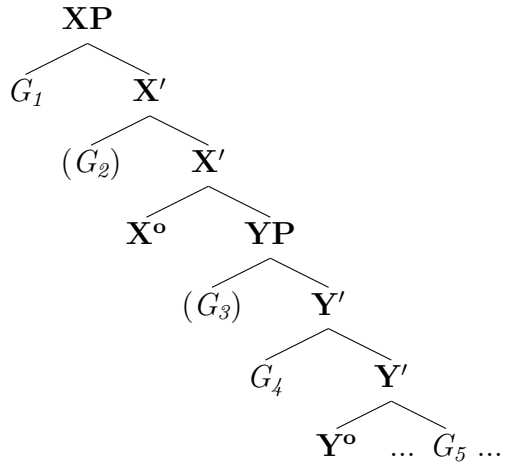
The strongest version of the Minimal Link Condition is given in (11).

(11) *Generalised Minimal Link Condition (e.g. Fitzpatrick 2002)*:
In a structure $X_{[\bullet F \bullet]}$ ... $[Y_{[F]}$ ... $[Z_{[F]}]$], movement to [•F•] can only affect the category bearing the [F]-feature that is closer to [•F•].

As Müller (2004, 2006a) notes, the effect of the MLC is limited by the Phase Impenetrability Condition, as the MLC presupposes search space, while one of the conceptual reasons for phases is to reduce the derivational complexity and thus relieve active memory by limiting the search space (Chomsky 2004).

In what follows, I will refer to subportions of a structure in which X is a (phase) head, and Y the next lowest (phase) head; G_1-G_5 are potential goals:

(12)



There are four possible constellations that the MLC is to derive:

(I) *The superiority case:*
A probe X has to decide between a goal at the edge of Y and a goal located in the complement of Y (G_4 vs. G_5).

(II) *The Equidistance case:*
X has to decide between two goals from the edge of the lower head Y (G_3 vs. G_4).

(III) *The A-over-A case:*
X has to decide between Y(P) and an element from the edge of Y (Y vs. G_4).

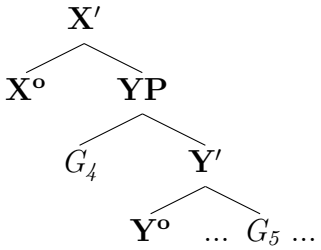
(IV) *The Two-Edges case:*
X has to decide between two goals, the one at the edge of X, the other at the edge of Y (G_1 vs. G_4).

An MLC effect of type (I) is e.g. the well-known subject-object asymmetry with multiple wh-questions in English shown in (13).

(13) a. Who_i t_i saw what?
b. * What did who see t_i?

The relevant portion of the derivation is given in (14).

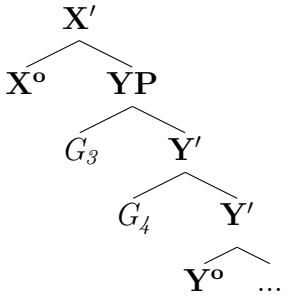
(14)



Here the MLC has the desired effect that an operation cannot involve X^o and G_5 due to the presence of the higher possible goal G_4. However, for this case, even under the widest definition of the PIC, there are MLC effects that are ruled out by it. In a minimal-static system, the observed effect follows automatically: G_5 is inaccessible to X as it is already spelled out when X is merged (Müller 2006b). Thus, if each phrase is a phase, then the MLC is not needed to derive this case, and should thus be eliminated if possible (see also Rizzi 2006).

The only domain where the MLC could have an effect in such a system is thus the search space of the current head. The relevant portion for the equidistance case is given in (15).

(15)



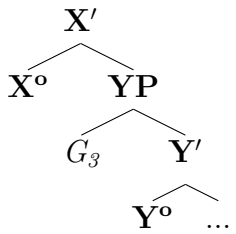
Here the MLC makes the prediction that the presence of G_3 blocks operations involving X and G_4. This is the wrong result: it is in exactly this constellation where Anti-MLC effects like order-preserving movement occur (Richards 1997, Starke 2001, Williams 2003). If the MLC is still assumed to be at work, then it must either be blocked by a prioritised Equidistance principle (Chomsky 1993, 2001), or its effect must be later repaired, e.g. by Tucking-In (Richards 1997).

(16) *Equidistance (Chomsky 2001:27):*
Terms of the edge of HP are equidistant from probe P.

The question is thus, what is the status of the MLC at phrase level? Müller (2006a) concludes that the MLC is not at work at all at phase level. The consequence is that there is *a priori* no preference for either of the two goals. That is a desired result, as it allows for order-preserving movement without invoking Tucking-in or Equidistance (note that the trigger for order-preserving movement is at any case an independent, additional factor); also, the Equidistance principle is not needed any more.

However, in the two remaining cases (III) and (IV), an MLC-less system seems to make the wrong predictions. Let us examine at A-over-A effects first. The relevant structure is given in (17).

(17)



Let us examine three examples that exhibit this constallation. The first example involves Unambiguous Domination effects in German (Takano 1994, Koizumi 1995, Kitahara 1997, Müller 1998, Sauerland 1999): Scrambling of a whole VP is possible, but scrambling of an NP inside the VP followed by remnant scrambling of the VP leads to an unacceptable result. This is illustrated in (18).

(18) a. dass [$_{vP}$ [$_{VP_2}$ das Buch$_1$ zu lesen] keiner t$_2$ versucht hat]
that the book.ACC to read no.one.NOM tried has
'That no one tried to read the book'

b. * dass [$_{vP}$ [$_{VP_2}$ t$_1$ zu lesen] [$_{NP_1}$ das Buch] keiner t$_2$ versucht hat]
that to read the book.ACC no.one.NOM tried has

The second example comes from Breton agreement (Jouitteau & Rezac 2006). Breton shows a complementarity effect in that the ϕ-features of a phonologically null NP are coded by ϕ-agreement morphology on the verb (='rich agreement'), whereas the ϕ-features of a phonologically overt NP are not coded by ϕ-agreement morphology on the target ('invariant agreement' [=frozen 3SG agreement or bare stem]). This is illustrated in (19).

(19) a. Gant o mamm e karf-ent /*karf-e *pro* bez-añ
with their mother R would.love-3PL /*would.love-3SG 3PL be-INF
'They would like to be with their mother'

b. Gant o mamm e *karf-ent /karf-e Azenor ha Iona bez-añ
with their mother R *would.love-3PL /would.love-3SG Azenor and Iona be-INF
'Azenor and Iona would like to be with their mother'

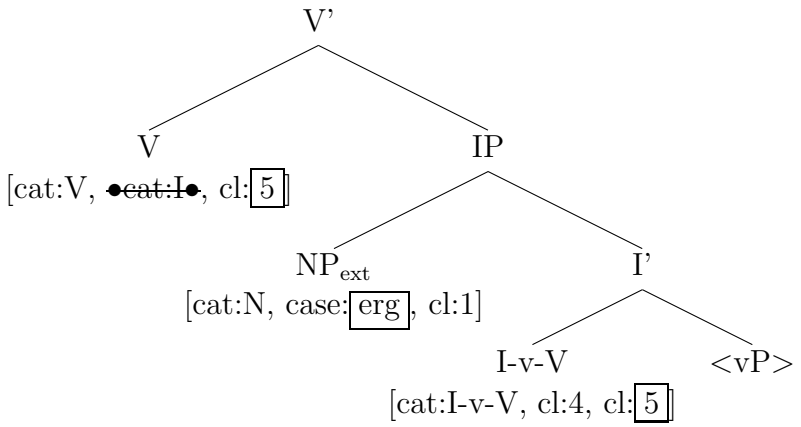
Jouitteau & Rezac (2006) analyse the complementarity as a locality effect. The starting point of the analysis is the observation that v in Breton has nominal properties. It is thus assumed to bear interpretable 3SG ϕ-features. Consequently, when I probes for ϕ-features in its search space, then v intervenes between I and the external argument, which is contained in the vP. I must therefore value its unvalued features with ϕ-features of v, which results in 3SG ('frozen') agreement on I. If, on the other hand, the external argument is an affixal *pro*, then it incorporates into T and thus contributes its ϕ-features to the feature set of T (i.e., it becomes a bound pronoun), which surfaces as rich agreement. A very similar constellation can be found in the context of long-distance agreement (LDA), which is an agreement relation that seems to hold across the boundaries of locality domains. Basically, the configuration of LDA is such that the verb of a root clause agrees with the internal argument of an embedded clause. It occurs e.g. in Tsez, Kwarshi, Kutchi Gujarati, Hindi-Urdu, Blackfoot, Chukchee, and Itelmen (Polinsky & Potsdam 2001, Khalilova 2007, Grosz & Patel 2006, Bhatt 2005, Bobaljik & Wurmbrand 2005, Bošković 2007, among others). An example from Khwarshi is given in (20).

(20) a. Išet'ul **y**-iq'-še goli [uža **bataxu** **y-**acc-u]
mother:OBL,LAT **cl5**-know-PRS COP [boy:ERG **bread(cl5)** **cl5**-eat-PSTPRT].CL4
'Mother knows that the boy ate bread'

b. Išet'ul **l**-iq'-še goli [uža bataxu y-acc-u]
mother:OBL,LAT **cl4**-know-PRS COP [boy:ERG bread(cl5) cl5-eat-PSTPRT].**cl4**
'Mother knows that the boy ate bread' (Khwarshi; Khalilova 2007)

Butt (1993, 1995), Chomsky (2001), Legate (2005), Lahne (2008b) analyse LDA as cyclic agree; i.e., the matrix verb agrees with the embedded verb, which has previously agreed with the embedded object. The relevant portion of structure is shown in (21). The gist of the analysis proposed in Lahne (2008b) is that LDA takes place if v-V moves to I and thus contributes the class features of the embedded object, which it receives by agreement. If, however, v-to-I movement does not take place, then matrix V only finds embedded I's inherent class feature [cl:4].

(21)



The important point for the discussion is that the external argument at the edge of I never intervenes in the agreement between matrix and embedded verb (see Bobaljik 2006:29). Thus again, the data suggest that a probe "prefers" agreeing with head Y to agreeing with an element at the edge of Y.

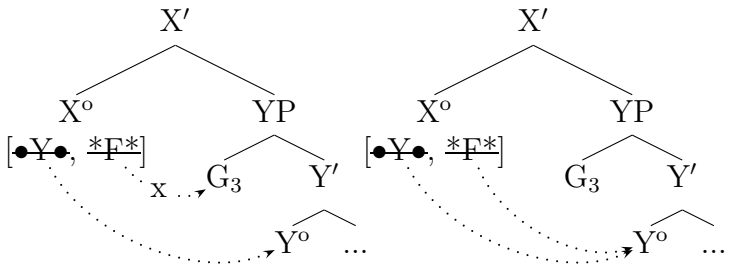
A first sight, these data are not accounted for if the MLC is completely abandoned. They can, however, be derived from a weaker version of the MLC in which the condition for intervention is dominance (Müller 2004):[4]

(22) *A-Over-A Condition (Chomsky 1964):*
In a structure [A ... [A ...] ...], an operation can only affect the higher, more inclusive category A.

The A-over-A Condition correctly derives A-over-A cases, and is at the same time not defined for the Superiority and Equidistance cases, which is the desired result (Müller 2004). However, even this last residue of the MLC can be derived by the Generalised Specificity Principle, see (1): It has the effect that the current probe X has to handle its selectional feature [•cat:Y•] by merging with a head Y; due to GSP, agree between X and Y must involve handling of the maximal number of matching features. Thus, if X has another feature [*F*] that it can value with a feature of Y, then it must value it with Y. This is illustrated in (23).

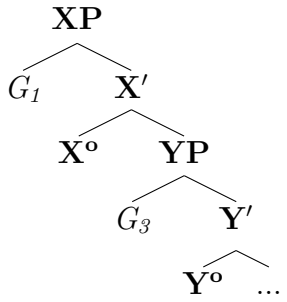
4. Note that this runs counter to Pesetsky & Torrego (2001), where head and specifier are taken to be equidistant.

(23) a. b.



Let us now consider the remaining two-edges case, recall (IV). The relevant structure is shown in (24).

(24)



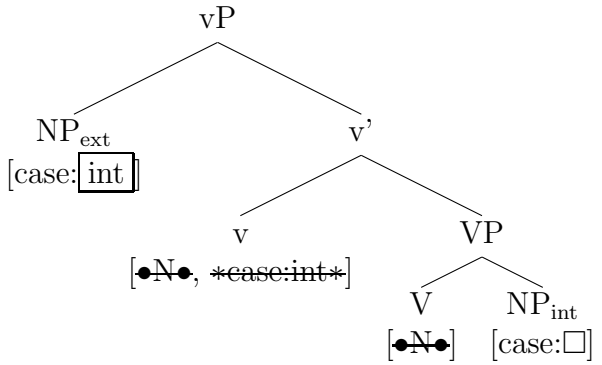
It was already argued that the original MLC cannot be retained. Crucially, if only the A-over-A Condition is retained, then G_1 does not block agreement between X and G_3. There are, however, data that suggest that the presence of a higher specifier does indeed block a value operation between a head and an element at a lower phase edge. In what follows, I present two examples.[5]

One such example is the emergence of the ergative case pattern proposed in Müller (2008): In a certain class of languages (i.e., in those languages that finally show an ergative case pattern), v merges the external argument before it assigns case.[6] As a result, the presence the external argument forces v to assign internal case, i.e., accusative or ergative, to the external argument, and not to the internal argument, which is located within the VP (i.e., within v's complement).

5. A third example is presented in Lahne (2008c), where verb inversion in the context of successive-cyclic movement (e.g. Henry 1995, Kayne & Pollock 1978, Torrego 1984, Rizzi & Roberts 1989, Rizzi 1990) is analysed as as a result of the early timing of Phase Balance (a constraint that synchronises the current make-up of the workspace with the shape of the current phase): wh-movement to the edge of I happens before I can deal with its own EPP and ϕ-features. The wh-element then blocks Value/Agree between I and elements at the edge of v. As a result, I must move to C by a last-resort operation in order to satisfy these features.

6. Reason: When v is merged, then an indeterminacy in rule application arises, as v has two unhandled features, a case feature [case:int(ernal)], and a selectional feature [•N•], which introduces the external argument. The next operation could thus either be Value(v, NP_{int}), or Merge (v, NP_{ext}). The Earliness Principle, an independently well-motivated principle underlying syntactic derivations (Pesetsky 1989, Pesetsky & Torrego 2001), demands that if Merge and/or Agree are possible at a stage of the derivation, then they must apply at once. This constitues a conflict, as only one operation can apply at a time. This conflict is resolved by giving one operation priority of the other. Languages thus differ from each other with regard to which operation applies first: In one class of languages, the case is assigned before the external argument is merged; in another class, the external argument is merged before case is assigned.

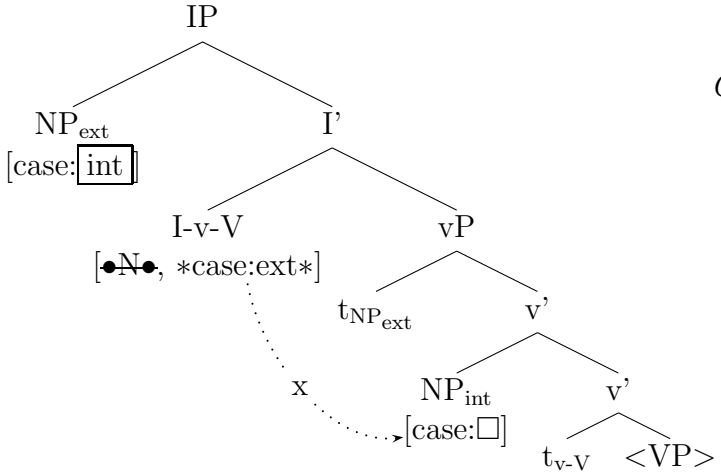
(25)



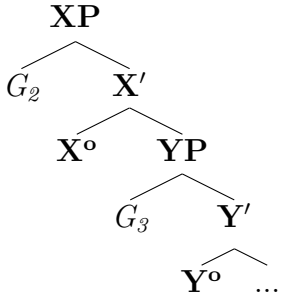
In other words, the presence of the external argument forces v to assign its feature [case:internal] to NP_{ext}, and not to the internal argument, which is located within the VP (i.e., within v's complement).

A second example of specifier intervention is Lahne's (2008a) account of Mahajan's Generalisation (i.e., the typological generalisation that ergative languages exhibit SOV and VSO order, but not SVO order; see Mahajan 1994, 1997). The account works along the same line of argumentation as Müller (2008): The feature set of I contains a case feature [∗case:ext∗], and can contain an EPP feature [•N•]. In a certain class of languages (i.e. those languages that end up with an ergative pattern), the order of operation application is Merge ≫ Value. Thus, if I has an EPP feature (=[•N•]), the order of feature handling is [•N•] ≫ [∗case:ext∗]. Consequently, I first attracts NP_{ext} to its edge. This would yield SVO, but the derivation crashes in the next step: I cannot value the case feature of NP_{int} as the presence of NP_{ext} at the edge of I prevents I from assigning case to NP_{int}:

(26) a. Relevant portion of structure:



b. Abstraction:



Mahajan's Generalisation can now be derived: In languages where Merge precedes Agree (which finally yields an ergative case pattern), the presence of a G_2 blocks case assignment to G_3. In other words, in ergative languages, the presence of specifier of I blocks agreement between I and an element at the edge of v. Thus, in ergative languages, those derivations in which external and internal argument are specifiers of different heads do not converge; only those derivations converge in which both external and internal argument are specifiers of the same head. Consequently, in ergative languages, there is no position between external and internal argument that the verb could possibly fill.

These examples suggest that in structures of type (IV), intervention by closeness does seem to play a role after all. There are two possible responses to this. One is to define an intervention condition that jumps in only in constellations of type (IV), and that has the

effect that specifiers of X are closer to X than specifiers of Y, while the specifiers "within" an edge do not block each other. This constraint must be defined in such a way that it only holds for Value/Agree relations, and not for structure building, as derivation by phase unquestionably involves successive-cyclic movement from the edge of Y to the edge of X even if X already has one or more specifiers.[7] I will however not follow this path, as it will presumably amount to a definition that is simply stipulated, which would be an unsatisfying result. A second and arguably more desirable solution would be to explain the apparent MLC-effect in case (IV) by means of an independent principle. There is indeed such a principle: the General Specificity Principle (GSP). In the two examples discussed, the intervening specifier is merged due to a selectional feature [•N•]. The GSP has the effect that Agree between a probe and a goal must involve the maximal number of matching features. Hence, if X satisfies a selectional feature by merging a specifier G_2, and if it can satisfy more features with G_2, then it must satisfy these features with G_2. Hence, Value/Agree operations involving X and G_3 are blocked. Merge operations, on the other hand, are not blocked, as a second specifier is merged by an edge feature or a second [•N•]-feature, which cannot be satisfied by internally merging G_1 again within the edge of X.

To sum up this discussion, in a system where each phrase is a phase, MLC effects can be derived as a result of phase-based derivation on the one hand (=case [I]), and of the principle Maximize Matching Effects on the other hand (=cases [III] and [IV]). Neither the PIC nor the GSP hold for case (II), which is the desired result, as now no additional principle is needed to allow for Anti-MLC effects (again, note that the trigger for order-preserving movement is at any case an independent, additional factor). The strong conclusion from this discussion is that the Minimal Link Condition can indeed be abandoned in a system where each phrase is a phase.

4 Analysis: Deriving Anti-MLC Effects

4.1 Order-preserving Movement

In Bulgarian multiple wh-constructions, the order of the fronted wh-phrases is exactly the same as the base-merge order: a fronted wh_{nom}-phrase precedes a wh_{acc}-phrase, and when three verbal arguments are wh-fronted, then the order is invariably '$NP_{NOM} \prec NP_{ACC} \prec NP_{PREP}$' (Rudin 1988):

(27) a. Koj_i $kogo_j$ vižda t_i t_j? (*only possible order of wh-expressions*)
who whom sees
'Who sees whom?'

b. Koj_i $kogo_j$ na $kogo_k$ e pokazal t_i t_j t_k? (*only possible order of wh-expressions*)
who whom to whom has pointed.out
'Who pointed out whom to whom?'

I claim that this is a specificity effect. The new analysis is based on the following case feature decomposition:[8]

7. This is indeed the solution proposed in Müller (2008) and Lahne (2008a), where intervention is defined by closeness (path): a specifier of X is closer to X than an element contained in the complement of X.
8. See e.g. Caha (2006) for a similar decomposition of case into feature sets varying in size.

Table 1

Case	*Decomposition*
NOMINATIVE	[case]
ACCUSATIVE	[case object]
PREPOSITIONAL	[case object oblique]

In ditransitive contexts, V thus contains two selectional features [N], and two case features to be assigned, [case object] and [case object oblique]. The GSP has the effect that the argument which V merges with first is assigned the more specific case feature (=[case object oblique]). V's second argument then receives the less specific case feature (=[case object]). The prediction, which will be tested in the project, is thus that structurally lower arguments always receive a more specific case than higher arguments.
The same argumentation holds for the data in (27): C has three structure-building [wh]-features, and three potential goals in its search space, which are identical apart from their case specification: *koj* bears [case] *kogo* [case object], and *na kogo* [case object oblique]. All potential goals have the same number of matching features, namely [wh]. The probe therefore cannot decide which element to merge first. I propose that the C head now probes further into the feature sets of the goals, being sensitive only to those features that are syntactically relevant. Following e.g. Wiese (2003), I assume that wh-elements are underspecified with regard to person and number; they are fully specified only for category ([cat:D]), quantification ([+wh]) and case. The wh-elements are identical in [cat:D, wh]; they differ only in their case feature. Thus, even though case does not take part in the agree operation between C and the wh-items, it is now the decisive factor for the order of merge: C attracts the goal that is most specific in terms of case features. The consequence is that C merges the wh-item with the most specific case first, which is *na kogo*. C's second goal is the second most specific wh-item in terms of case, *kogo*. The least specific wh-element, *koj*, is merged last. The prediction is thus that elements that are less specific with regard to morpho-syntactic features appear higher in the structure than more specific elements.
At first glance, this prediction does not seem to hold. For example, in Bulgarian, the specificity effect interacts with an animacy effect, that is, animate arguments appear higher in the left periphery than inanimate arguments, regardless of their case marking. This is illustrated in (28) and (29).

(28) a. ?(?) Kakvo kogo e spoletjalo?
what whom is stricken'
'What struck whom?' (Billings & Rudin 1996:38)

b. Kogo kakvo e spoletjalo?
whom what is stricken
'What struck whom?'

(29) a. ? Kakvo na kogo mu xaresva?
what to whom to-him appeals
'What appeals to whom?' (Billings & Rudin 1996:40)

b. Na kogo kakvo mu xaresva?
to whom what to-him appeals
'What appeals to whom?'

The potential problem for the new analysis is that elements bearing an additional feature [+animate] are arguably more specific than inanimate elements. I will develop an analysis that derives data of this kind within the new approach. The basic idea is that syntactic derivations are indeed driven by specificity. However, C's specificity-based decision on the order of merge operations can be overwritten if it runs counter to a more important constraint, e.g. linearisation preferences linked to animacy according to which animate elements must be merged last (see Simpson & Bhattacharya 2007).[9]

4.2 Anti-Superiority Effects

Cinque (1999) and Rizzi (2002), among others, show that an intervening adverb blocks displacement of a deeper adverb to the C domain, but not if the deeper adverb is contrastively focused. Some German data illustrating this effect are shown in (30); similar effects can be found in French and Italian.

(30) a. Karl hat wahrscheinlich sehr oft gelogen
Karl has probably very often lied

b. Wahrscheinlich$_i$ hat Karl t_i sehr oft gelogen
probably has Karl very often lied

c. * [Sehr oft]$_i$ hat Karl wahrscheinlich t_i gelogen
[very often] has Karl probably lied

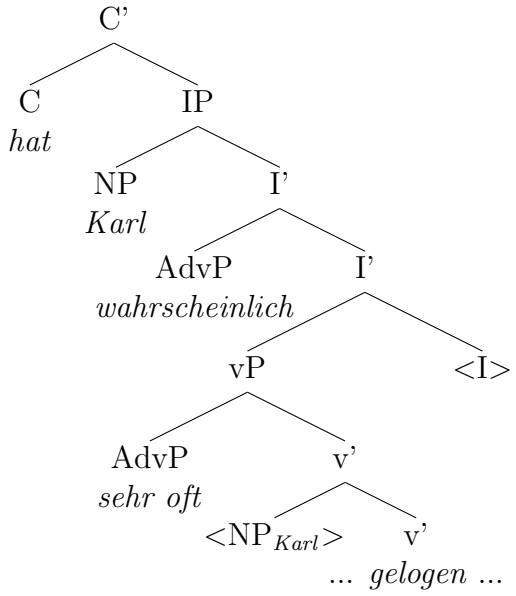
a.-c.: 'Karl probably lied very often'

d. [SEHR OFT] hat Karl wahrscheinlich t_i gelogen
[VERY OFTEN] has Karl probably lied

'Karl probably lied VERY OFTEN'

I claim that this is a specificity effect. Consider the following structure, which shows the derivation shortly before C attracts an adverb:

(31)



9. A similar effect can be found in English:

(i) Whom did what upset? (Simpson & Bhattacharya 2007)

(i) shows that in English, too, the MLC effect can be overwritten if the object wh-phrase crosses an inanimate wh-subject.

I assume that sentence adverbials are more specific than v-adverbials. A possible account for this assumption is that I contains a feature [+anaphoric], which it assigns to all adverbials that it merges (for a similar analysis cf. Lopez 2007); v-adverbials do not receive such a feature. As a result, the probing head C always prefers the sentence adverbial *wahrscheinlich*. In (30d), however, the lower adverbial bears a focus feature, which is an abbreviation for [–anaphoric +contrastive] (cf. Lopez 2007). Now *SEHR OFT* is more specific, as it now contains more features than the sentence adverbial. Due to the GSP, $C_{[\bullet \text{–anaph +contr}\bullet]}$ must satisfy both [–anaphoric] and [+contrastive] with the same element. Consequently, *SEHR OFT* is fronted.

A second type of data that can be explained by means of specificity are weak island violations that unexpectedly lead to a grammatical result: It is a well-known observation that discourse-linked wh-elements can be extracted from weak islands, as shown in (32).

(32) a. ? [Quale problema]$_i$ non sai come$_j$ risolvere t_i t_j?
which problem not know.2SG how solve
'Which problem don't you know how to solve?'

b. * Come$_j$ non sai [quale problema]$_i$ risolvere t_i t_j?
how not know.2SG which problem solve
'*How don't you know which problem to solve?'

I claim that this is a specificity effect. There is a point in the derivation where both wh-elements accessible to a higher probe. One of them now has to be moved on. The edge feature now finds two elements with interpretable wh-features, however, one of them is more specific than the other, as it bears an additional semantic feature [top] (or [anaph]). It is thus the more specific element, *quale problema*, that is moved on.

5 Summary

It seems to be a recurring pattern in syntactic derivations that structure-building operations are ordered in such a way that elements that are more 'specific' in one sense or another are merged earlier with the probing element than less specific elements. The main consequence of the new analysis is that the MLC and related locality conditions (Equidistance, A-over-A-Condition) can be dispensed with in a system in which syntactic derivation is modeled as specificity-driven. This strengthens the role of the Phase Impenetrability Condition (Chomsky 2000:108, Chomsky 2001:13).

References

Anderson, Stephen R. (1992): *A-Morphous Morphology*. Cambridge University Press, Cambridge, UK.

Bhatt, Rajesh (2005): Long Distance Agreement in Hindi-Urdu, *Natural Language & Linguistic Theory* 23(4), 757–807.

Billings, Loren & Catherine Rudin (1996): Optimality and Superiority: A New Approach to Overt Multiple-Wh Ordering. In: J. Toman, ed., *Formal Approaches to Slavic Linguistics: The College Park Meeting 1994*. Michigan Slavic Publications, Ann Arbor, pp. 35–60.

Bittner, Maria & Ken Hale (1996): The Structural Determination of Case and Agreement, *Linguistic Inquiry* 27(1), 1–68.

Bobaljik, Jonathan (2006): 'Where's ϕ? Agreement as a Post-Syntactic Operation'. Manuscript, University of Connecticut.

Bobaljik, Jonathan & Susi Wurmbrand (2005): The Domain of Agreement, *Natural Language & Linguistic Theory* 23(4), 809–865.

Bošković, Željko (2007): Agree, Phases and Intervention Effects, *Linguistic Analysis* 33, 54–96.

Butt, Miriam (1993): Hindi-Urdu Infinitives as NPs, *South Asian Language Review* 3(1), 51–72.

Butt, Miriam (1995): *The Structure of Complex Predicates in Urdu*. Dissertations in Linguistics, CSLI Publications, Stanford.

Caha, Pavel (2006): 'The Superset Principle'. Manuscript, Universitetet i Tromsø.

Chomsky, Noam (1964): *Current Issues in Linguistic Theory*. Mouton, The Hague.

Chomsky, Noam (1993): A Minimalist Program for Linguistic Theory. *In* Hale & Keyser (1993), pp. 1–52.

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

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

Chomsky, Noam (2004): Beyond Explanatory Adequacy. In: A. Belletti, ed., *Structures and Beyond. The Cartography of Syntactic Structure*. Vol. 3, Oxford University Press, Oxford, etc., pp. 104–131.

Cinque, Guglielmo (1999): *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.

Di Sciullo, Anna-Maria & Edwin Williams (1987): *On the Definition of Word*. Number 14 *in* 'Linguistic Inquiry Monographs', MIT Press, Cambridge, Mass.

Epstein, Samuel & Daniel Seely (2002): Rule Application as Cycles in a Level-free Syntax. In: S. Epstein & D. Seely, eds., *Derivation and Explanation in the Minimalist Program*. New York, Blackwell, pp. 65–89.

Fitzpatrick, Justin M. (2002): On Minimalist Approaches to the Locality of Movement, *Linguistic Inquiry* 33(3), 443–463.

Grosz, Patrick & Pritty Patel (2006): 'Long Distance Agreement and Restructuring Predicates in Kutchi Gujarati'. Manuscript, MIT.

Hale, Kenneth & Samuel Jay Keyser, eds. (1993): *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. MIT Press, Cambridge, Mass.

Halle, Morris (1997): Distributed Morphology: Impoverishment and Fission. In: B. Bruening, Y. Kang & M. McGinnis, eds., *MIT Working Papers in Linguistics*. Vol. 30, Cambridge, Mass., pp. 425–449.

Heck, Fabian & Gereon Müller (2006): Extremely Local Optimization. In: E. Bainbridge & B. Agbayani, eds., *Proccedings of WECOL 34*. Department of Linguistics, University of Fresno, CA, pp. 70–182.

Henry, Alison (1995): *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford Studies in Comparative Syntax, Oxford University Press, New York.

Jouitteau, Mélanie & Milan Rezac (2006): Deriving the Complementarity Effect: Relativized Minimality in Breton Agreement, *Lingua* 116(11), 1915–1945.

Kayne, Richard S. & Jean Yves Pollock (1978): Stylistic Inversion, Successive Cyclicity, and Move NP in French, *Linguistic Inquiry* 9(4), 595–621.

Khalilova, Zaira (2007): 'Clause Linkage: Coordination, Subordination and Cosubordination in Khwarshi'. Manuscript, Max-Planck Institut für evolutionäre Anthropologie, Leipzig.

Kiparsky, Paul (1973): The Inflectional Accent in Indo-European, *Language* 49(4), 794–849.

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

Koizumi, Masatoshi (1995): Phrase Structure in Minimalist Syntax. PhD thesis, Massachusetts Institute of Technology, Cambridge, Mass.

Lahne, Antje (2008a): On Excluding SVO in Ergative Languages, *Varieties of Competition. Linguistische Arbeitsberichte* 87, 65–80.

Lahne, Antje (2008b): 'Specificity-driven Syntactic Derivation: A New View on Long-distance Agreement'. Under review for *Syntax*.

Lahne, Antje (2008c): Where There is Fire There is Smoke. Local Modelling of Successive-cyclic Movement. PhD thesis, Universität Leipzig.

Legate, Julie Anne (2005): Phases and Cyclic Agreement. In: *Perspectives on Phases*. Vol. 49 of *MIT Working Papers in Linguistics*, pp. 147–156.

Lopez, Luis (2007): 'A Derivational Syntax for Information Structure'. Book manuscript, University of Illinois-Chicago.

Lumsden, John (1992): Underspecification in Grammatical and Natural Gender, *Linguistic Inquiry* 23(3), 469–486.

Mahajan, Anoop (1994): The Ergativity Parameter: *Have-be* Alternation, Word Order and Split Ergativity. In: *Proceedings of NELS 24*. GLSA, Amherst, pp. 317–331.

Mahajan, Anoop (1997): Universal Grammar and the Typology of Ergative Languages. In: A. Alexiadou & T. A. Hall, eds., *Studies on Universal Grammar and Typological Variation*. John Benjamins, Amsterdam, pp. 35–57.

Manzini, Maria Rita (1994): Locality, Minimalism, and Parasitic Gaps, *Linguistic Inquiry* 25, 481–508.

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

Müller, Gereon (2004): Phrase Impenetrability and Wh-Intervention. In: A. Stepanov, G. Fanselow & R. Vogel, eds., *Minimality Effects in Syntax*. de Gruyter, Berlin, pp. 289–325.

Müller, Gereon (2006a): 'Deriving MLC Effects'. Manuscript, Universität Leipzig.

Müller, Gereon (2006b): 'Extended Exponence by Enrichment: Argument Encoding in German, Archi, and Timucua'. Manuscript, Universität Leipzig.

Müller, Gereon (2008): 'Ergativity, Accusativity, and the Order of Merge and Agree'. Manuscript, Universität Leipzig.

Murasugi, Kumiko (1992): Crossing and Nested Paths. PhD thesis, MIT, Cambridge, Mass.

Pesetsky, David (1989): 'Language-particular Processes and the Earliness Principle'. Manuscript, MIT, Cambridge, Mass.

Pesetsky, David & Esther Torrego (2001): T to C Movement: Causes and Consequences. In: M. Kenstowicz, ed., *Ken Hale: A Life in Language*. MIT Press, MIT, Cambridge, Mass., pp. 355–426.

Polinsky, Maria & Eric Potsdam (2001): Long-Distance Agreement and Topic in Tsez, *Natural Language & Linguistic Theory* 19, 583–646.

Richards, Norvin (1997): What Moves Where When in Which Language?. PhD thesis, Massachusetts Institute of Technology, Cambridge, Mass.

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

Rizzi, Luigi (2002): 'Locality and Left Periphery'. Manuskript, Università di Siena.

Rizzi, Luigi (2006): 'Concepts of Locality/ On Intermediate Positions: Intervention and Impenetrability'. Lectures delivered at the XXV Curso de Verano, San Sebastián, June 2006, and the EALING Fall School, Paris, September 2006.

Rizzi, Luigi & Ian Roberts (1989): Complex Inversion in French, *Probus: International Journal of Latin and Romance Linguistics* 1(1), 1–30.

Rudin, Catherine (1988): On Multiple Questions and Multiple *Wh* Fronting, *Natural Language & Linguistic Theory* 6, 445–501.

Sauerland, Uli (1995): The *Lemmings* Theory of Case. In: J. Boro, L. Conway, P. Hsiao & E. Laurençot, eds., *Proceedings of SCIL VII*. MIT Working Papers in Linguistics, Cambridge, Mass.

Sauerland, Uli (1999): Erasability and Interpretation, *Syntax* 2(3), 161–188.

Simpson, Andrew & Tanmoy Bhattacharya (2007): Argument Prominence and the Nature of Superiority Violations. In: E. Reuland, T. Bhattacharya & G. Spathas, eds., *Argument Structure*. John Benjamins, Philadelphia, pp. 175–212.

Starke, Michal (2001): Move Reduces to Merge: A Theory of Locality. PhD thesis, University of Geneva.

Stump, Gregory T. (2001): *Inflectional Morphology*. Cambridge University Press, Cambridge.

Takano, Yuji (1994): Unbound Traces and Indeterminacy of Derivation. In: M. Nakamura, ed., *Current Topics in English and Japanese*. Hituzi Syoboo, Tokyo, pp. 229–253.

Torrego, Esther (1984): On Inversion in Spanish and Some of Its Effects, *Linguistic Inquiry* 15(1), 103–129.

Wiese, Heike (2003): Wh-Words are not 'Interrogative' Pronouns: The Derivation of Interrogative Interpretations for Constituent Questions. In: M. H. . N. Scott, ed., *Questions: Multiple Perspectives on a Common Phenomenon*. University of Liverpool Press, Liverpool.

Williams, Edwin (1994): Remarks on Lexical Knowledge, *Lingua* 92(1), 7–34.

Williams, Edwin (2003): *Representation Theory*. MIT Press, Cambridge, Mass.