

Cognitive Transitivity

The motivation of basic clause structures

Wolfgang Schulze (University of Munich) © 2011

1. Introduction: Some simple observations

A seemingly valid language universal concerns the structural properties of basic clauses. Accordingly, any basic clause in natural language is marked for a verb-like segment and at least one obligatory noun-like segment. This minimal requirement, however, is often violated especially with respect to the presence of the 'noun-like' segment, more rarely with respect to the 'verb-like' segment. These violations occur (among others) with zero-anaphora (1a), certain semantic verb classes (1b), and zero-verbs (1c), compare:

- (1) a. a^3nie^1 a^3ba^1 lou^4-die^1 bo^3 hu^3ri^1 luo^4 nie^3 se^3ke^1 pe^2ti^1 ri^1
mother father other-PL IO work person ASSOC field land work
'[My] parents work[ed] for others, [they] planted[ed] (lit.: work) others' field (and) land...' [Tujia; Brassett et. al. 2006:160]
- b. 下雨[xià yǔ]
rain(ing) ~ rain=drop(ing)
'It is raining.' [Chinese]
- c. *Ich* *trug* *den* *Tisch* *und* *Paul* \emptyset *die* *Stühle*
I:NOM carry:PAST:1SG ART:M:SG:ACC table and Paul ART:PL:ACC chair:PL
'I carried the table and Paul [carried] the chairs.' [German]

It should be noted that from a structural point of view, instances of the absence of 'noun-like' elements are restricted to dependent marking or isolating languages. Otherwise, residues (or 'echoes') of the noun-like element are usually present in the verbal complex, such as:

- (2) a. *doždi-l-o*
rain-PAST-N:SG
'It was raining' [Russian] [alternating with *dožd' idet* 'rain goes']
- b. *niev-a*
snow-PRES:3SG
'it is snowing' [Spanish]
- c. *c'vim-s*
rain-3SG
'it is raining' [Georgian]

Naturally, one may argue that the referential residues in these verbs are purely etymological and not structural, but the presence of e.g. a dummy-NP (German *es regnet* 'it is raining' etc. suggests that agreement in present (see Schulze & Sallaberger 2007 for the notion of referential echoes).

Whereas the overt 'noun-like' elements may occasionally lack in cotext-free utterances¹, zero-verbs seem to be restricted to a preceding cotext except for exclamative utterances such as *goal!* etc. From this we can infer that the presence of a verb-like element is the minimal requirement for a basic clause. The maximal size of a basic clause, however, is more difficult to describe. It heavily depends on the question of whether we are dealing with spoken or written texts and which role is played by the capacity of short term memory when producing and processing such texts. In written texts, the number of 'noun-like' elements is often higher than in spoken utterances, because here, the visuospatial sketchpad supports the phonological loop (see Baddeley & Hitch 1974, Baddeley 2000) and hence the memorizing of larger referential clusters. Starting from the hypothesis that written language is based on a secondary and relatively recent extension of cognitive capacities, we may assume that if we are looking for the motivation of basic linguistic structures, we should primarily refer to the structure of spoken language (and oral traditions). It is reasonable to assume that spoken language comes closer to phylogenetically older layers of (linguistic) cognition than written language. A first clue to the question of how many noun-like and verb-like constituents are typical for (non-situated) spoken language is given by a simple calculus that counts the constituents at issue in oral (traditional) text. For instance, a coarse-grained corpus analysis of the three languages German, Udi (Southeast Caucasian) and Navajo (Diné, Athapaskan) yields the following results (number of tokens: German: 33.830, Udi: 27.511, Navajo: 70.312; see Schulze 2004 for details on these data):

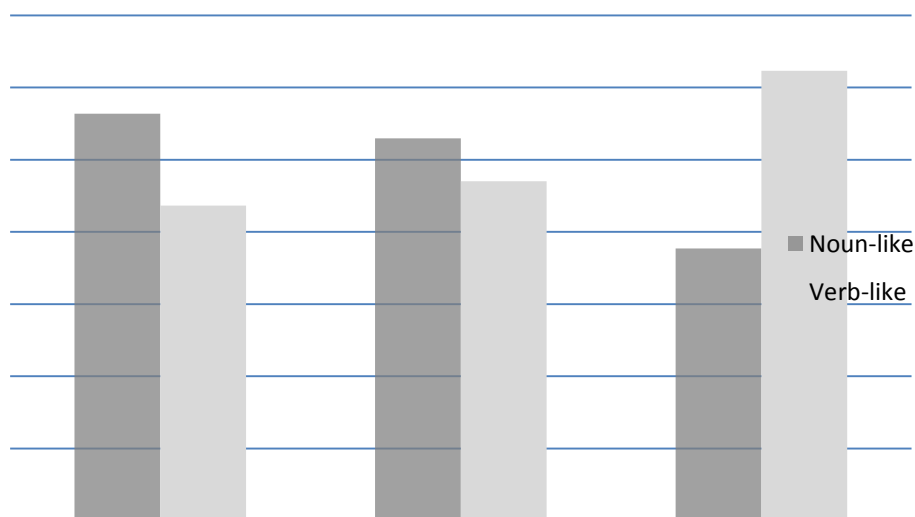


Diagram 1: *Noun Phrase and Verb Phrase in oral texts (relative proportion)*

Accordingly, for both German and Udi, each verb-like constituent is statistically marked for roughly 1.3 noun-like constituents, as opposed to 0.59 noun-like constituents in Navajo. The divergent data for Navajo illustrate that some languages (not necessarily confined to heavy head-marking languages) are characterized by a stronger tendency towards overt NP deletion, especially in discourse. It should be noted, however, that all three languages at issue have head-marking properties (monopersonal agreement in German and Udi, polypersonal agreement in Navajo). If we take agreement markers as echoes of what has been termed 'noun-like' so far, the proportion will increase in favor of the noun-like constituents (up to roughly 1.8 ~ 2.0). This calculus is corroborated for instance by a closer inspection of two Udi narratives (see Schulze 2004 for details):

¹ Note that I use the term 'cotext' to refer to the textual surroundings of an utterance. 'Context', on the other hand' refers to extra-linguistic domains (of knowledge, presuppositions etc.).

(3)

	Overt %	Inferred %	Total %
Noun-like	50,26	93,16	53,87
Verb-like	35,32	6,82	30,20

Overt and inferred units in King & Shepherd (dialect: Vartashen)

(4)

	Overt %	Inferred %	Total %
Noun-like	52,65	86,76	61,71
Verb-like	42,02	13,24	34,38

Overt and inferred units in Walking Sieve (dialect: Nizh)

The analysis considers both overt and inferred (covert) units. The degree of inference differs in the two texts: Whereas the text 'King and Shepherd', a traditional tale recorded in 1902 is marked for a ratio of 0.21 inferred constituents (per constituent), the less traditional tale 'Walking Sieve' (written down in 2002) is marked for the value 0.36. Still, the overall ratio of noun-like and verb-like constituents remains roughly alike: Each verb-like element is accompanied by approximately two noun-like constituents.

This trifurcation of basic clauses is reflected by what is traditionally called 'transitivity': Two homologous constituents are said to be bound to a non-homologous or 'relational' constituent given that one of the homologous partners presupposes (?) the existence of the other. Many definitions of 'transitivity' do not explicitly refer to the presence of two (or more) homologous partners. Rather, one partner (the *conditio* of the existential relation) is suppressed (or generalized) and only the existence of the conditioned partner is mentioned, e.g.: A transitive structure is given if a verb calls for an obligatory object. Hence, the definition is based on relational and semantic properties of the verb (see Lazard 2002, Næss 2007). In addition, it alludes to syntactic features such as (implicit) 'subject' and 'object', thus mixing up semantic and structural/functional features. In this sense, transitivity can be regarded as one of the expression types of verbs which competes with zero expressions ('intransitivity') and double expressions ('ditransitivity', see e.g. Anagnostopoulou 2003). The fact that some verbs behave as ambitransitives leads to the hypothesis that the category or 'dimension' expressed by the chain 'intransitive-transitive-ditransitive' is continuous rather than disjunct.² Ambitransitives (see Dixon & Aikhenvald 2000:4) behave in a lexically ambiguous way:³

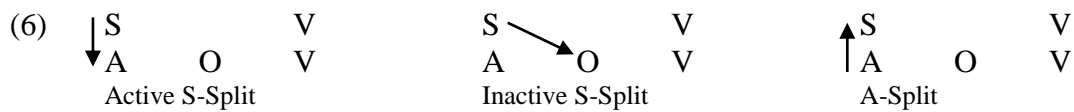
(5)	$\begin{array}{ccc} \mathbf{S=O} & \mathbf{V} & \\ \mathbf{A} & \mathbf{V} & \rightarrow \mathbf{O} \end{array}$	<i>The windows breaks</i> [unaccusative] ⁴ <i>I break the window</i>
	$\begin{array}{ccc} \downarrow \mathbf{S=A} & \mathbf{V} & \\ \mathbf{A} & \mathbf{V} & \mathbf{O} \end{array}$	<i>I win</i> [unergative] <i>I win the game</i>

² Alternatively, verbs having more than one such frame can be regarded as syntactically or semantically homonymous.

³ I use the following symbols to indicate relational primitives (see Schulze 2000, Dixon & Aikhenvald 2000): S (subjective), A (agentive), O (objective). In addition, RadEx uses IO (indirect objective) to mark an indirectly *effected* O-domain ('dative' etc.), IA (indirect agentive) to mark an indirectly *effecting* or *controlling* A-domain ('instrumental' etc.), and LOC to mark a locative domain (together with its metaphorization). The hybrid function 'causee' is indicated by AO (when strongly related to A [*let*-causation]) and OA (when strongly related to O [*have*-causation]).

⁴ See Kuno & Takami (2004) for a detailed discussion of unergativity/unaccusativity.

The intimate relationship between intransitive and transitive clause structures is also documented by the well-known Split Typology (Schulze 2000) that relates properties of one of the two 'transitive' relational primitives (A,O~IO) to S, among others:⁵ (This should be explained and illustrated)



The Active S-split is marked for the application of an A-behavior to S-constituents, whereas the Inactive S-Split copies an O-behavior. A well-know example is given by the agreement pattern of Dakota (here first Singular):

(7)

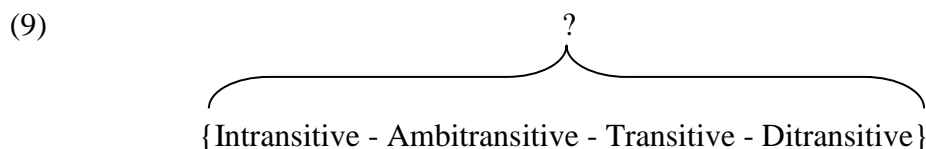
	S_A -Split	S_O -Split
S	<i>wa-</i>	<i>ma-</i>
A	<i>wa-</i>	
O		<i>ma-</i>

An A-Split copies S-properties onto the A-constituent, as in Yimas (Foley 1991:205;201):

- (8) *pu-ka-tay-Ø*
 3PL:O-1SG:A-SEE-PERF
 'I saw them'
- pu-na-tay-Ø*
 3PL:A>S-1SG:O-SEE-PERF
 'They saw me.'

In (8), the AGR-prefix *pu-* encodes the third person plural in S=O function. In the expression *punatay*, the O domain is higher in rank than the A domain, which is then demoted to an S-like constituent. This is done in order to indicate that A is construed as having a lower degree of agentivity with respect to O.

Accordingly, the four marks on the transitivity continuum represent co-hyponyms of a higher categorical node the nature of which, however, is seldom described in more detail:



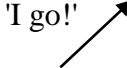
The term 'valence' (Tesnière 1959) is often used in this context, but its ontological status remains (in larger parts) obscure. The main question is whether the 'hypernymic' value has a distinct conceptual or structural ontology that is more than just a taxonomic descriptor, or whether this value exists only in its hyponymic instantiations. A further problem is to decide whether 'valence' is a semantic property of verbs that 'results' in syntactic features as soon as the valence at issue becomes satisfied, or whether verbs are subjected to a classification that starts from 'independent' syntactic classes such as (in)transitivity. In addition, if valence is a property of lexical verbs, it may be asked why there are so few basic valence patterns

⁵ The direction of the arrow indicates the target of property usurpation, see Schulze 2000.

observable from a typological point of view. If syntax is the source of the corresponding verb classification, the question arises whether we have to deal with syntactic building blocks (such as 'subject', 'object', 'verbs' etc.) or whether the 'parts' of a basic syntactic class are emergent properties of the whole. Finally, we have to bear in mind that the notion of 'obligatoriness' that is crucial for determining (in)transitive properties is all but clear. The standard hypothesis is here to claim that a constituent is obligatory if it must be expressed lexically (or pronominally etc.) in order to achieve a meaningful clause. For instance, the German phrase

- (10) *Ich geh-e in die Stadt*
 I:NOM go:PRES-1SG:PRES in ART:F:SG:ACC town
 'I go into town'

is usually interpreted as consisting of one obligatory (*ich*) and one facultative phrase (*in die Stadt*), describing the verb *gehen* ('to go') as an intransitive (monovalent) verb. However, the following dialog illustrates that the second phrase expresses a component that is processed or conceptually present even if it is lexically empty:

- (11) S1: *Ich gehe [Ø]!*
 'I go!'
 S2: *Wohin?*
 'Where (to)?'
- 

Wh-questions are usually regarded as specific anaphoric elements that refer to a referential or relational 'dummy' (see Schulze 2007). Example (9) illustrates that the clause *ich gehe* entails a covert informational segment or 'inferential option' that can be referred to by the hearer. In other words: 'Obligatoriness' is a feature of conventionalization rather than of verbal semantics as such.

The questions addressed above should also be seen in the light of those observations made in the first part of this section: If we start from the hypothesis that in texts every verb-like element is conceptually accompanied by roughly two noun-like elements, we may assume that grammatical intransitivity is a specific strategy based on the linguistic (not conceptual!) reduction of overt valence marking. Admittedly, this hypothesis goes against assumptions according to which the ontology of language is intimately related to one of the basic genres or to the 'pre-genre' (Swales 1990) of language use (see Hopper 2003), namely conversation (Bakhtin 1968). In fact, conversation seems to be marked for a relatively low degree of referential explicitness, or, as Hopper (2003) has put it:

"[T]he rarity of cardinal transitivity in conversation poses questions about the sources of transitive marking. Our recent findings seem to suggest (1) that grammatical marking is divorced from usage, and (2) that transitivity is relevant not for a language as a whole but only for certain genres."

However, Hopper's assumptions can be questioned for the following reason: First, the reduction of the ontology of language to conversation presupposes that conversation itself is the primary source for language evolution. Still, if we refer to e.g. Radical Constructivism, we may also argue that conversation reflects the construction of a collective hypothesis concerning the social experience of articulation-based 'expressions' of perception and experience (Schulze 1998, 2007). Accordingly, the 'transitivity' scale would be grounded in strategies to articulate experience, not in conversation. Second, Hopper's arguments are strongly related to linguistic utterances. As has been said above, linguistic utterances may

considerably differ from the underlying 'cognitive events' that are 'encoded' by the utterance: They may include a number of referential entities the processing of which is conditioned by context and cotext and which are retrieved through inference. In other words: The strong tendency towards intransitive structures in conversation is based on specific strategies to reduce the amount of linguistic 'material' (which is quite in accordance with Zipf's Law of least [mental] effort (Zipf 1949)).

In the following sections, I want to make use of the points addressed so far to set up a model of transitivity that is grounded in the framework of Radical Experientialism and Cognitive Typology (Schulze 1998, 2001, 2007, 2009, in press). Section 2 is a very condensed presentation of the relevant features of Radical Experientialism. In Section 3, I will propose a strongly deductive model of Cognitive Transitivity together with indications of how Cognitive Transitivity becomes 'parameterized' in articulated language. Section 4 summarizes some of the most central claims.

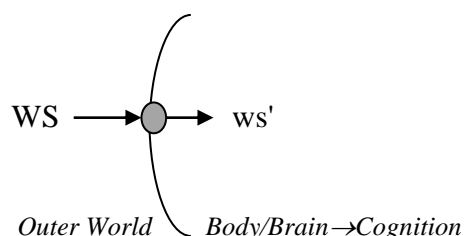
2. Radical Experientialism

Radical Experientialism (RadEx) is a model of linguistic knowledge and linguistic behavior that is grounded (among others) in Radical Constructivism, Gestalt Theory, non-Objectivistic approaches to Cognition (see Lakoff 1987), and Holistic Cognitive Linguistics. As to Radical Constructivism it suffices to quote the famous definition by Ernst von Glasersfeld:

What is radical constructivism? It is an unconventional approach to the problem of knowledge and knowing. It starts from the assumption that knowledge, no matter how it is defined, is in the heads of persons, and that the thinking subject has no alternative but to construct what he or she knows on the basis of his or her own experience. What we make of experience constitutes the only world we consciously live in. It can be sorted into many kinds, such as things, self, others, and so on. But all kinds of experience are essentially subjective, and though I may find reasons to believe that my experience may not be unlike yours, I have no way of knowing that it is the same. The experience and interpretation of language are no exception. (Glaserfeld 1996:1)

Radical Experientialism is heavily marked for reductionism (Schulze 2010).⁶ Accordingly, it assumes that linguistic knowledge and linguistic 'systems' are ultimately motivated by the complex interaction of perception and experience and by very 'simple' (radical), self-mirroring procedures to process them. By 'perception' is meant any kind of input ('world stimulus', WS) into the cognitive domain⁷ mediated and contorted by perceptual devices ($\rightarrow ws'$):

(12)



⁶ The scope of this paper does not allow for detailed discussion of why RadEx favors reductionisms. Apart from theoretical considerations, reductionistic approaches also have a presentational advantage: They escape from overloading graphical schemas with specific, non-reduced information. See Schulze (1998:413-426) for a brief ontology of formal representations in linguistics (also compare Engels (2007) who illustrates the emergence of the geometric representation of metaphysical concepts).

⁷ Note that RadEx defines 'Cognition' in a rather broad sense: Cognition is seen as the functional side of the neural substrate of the brain, no matter which subdomain of the brain is addressed.

The contortion of the input properties is primarily governed by the properties of the input device. This means that we see an 'object' differently from how we hear it (without having seen it or anything like it before). The same holds for the remaining senses (tactition, olfaction, gustation). In principle, all senses can theoretically account for language-oriented processing of word stimuli. However, the fact that language can be regarded as a strategy to link articulatory patterns to cognitive 'events'⁸ determines that audition figures among the most prominent input devices. The second relevant input device is given by vision. RadEx starts from the hypothesis that with human beings, vision represents the most prominent tool to interact with the Outer World (compare Holšánová 2008). The centrality of vision is also documented by the many instances which refer to it in order to symbolize knowledge.⁹ Here, it does not matter at what stage during the perception process a visual input becomes interpreted as being 'meaningful'. In RadEx, 'meaning' is the result of secondary processes that link a *ws'* to a memory segment and its symbolization (see below). It suffices to refer to the three-stage model described by Hollingworth & Henderson (1999:380):¹⁰

Current computational theories of visual perception tend to break down the perception of meaningful stimuli into three functional stages. First, primitive visual features (e.g., surfaces and edges) are extracted from retinal information. Second, these features are used to construct a description of the structure of a stimulus. Third, the constructed description is matched against stored descriptions.

According to RadEx, a visual stimulus is contorted from the very beginnings of the processing stages, based on a limited number of schematic procedures (see below) and gradually loading the resulting *ws'* with a 'content'. Or, to put it into other words: A *ws'*, that is the state of cognition immediately 'after' its enervation by a (here) visual input, is processed (among other things) with the help of highly abstract schemas, gestalt models, and experience-based 'images'. A further contortion is given if the primary input effect (*ws'*) is coupled with an expressive mechanism that is based on motorics: As has been said above, RadEx subscribes to the assumption that language results from the conventionalized coupling of categorized and schematized cognitive events with articulatory models grounded in the motorics of respiratory perturbation.¹¹ If a *ws'* is processed in junction with a corresponding expression model, the properties of this model may immediately influence the gestalt of *ws'*. Note that here, RadEx does not refer to a Whorfian model of language and cognition: It is not a particular language that contorts *ws'* but the basic properties of language as such. Among them, the most relevant one is the dimension of linearization: The fact that human beings can 'perturb' their respiration

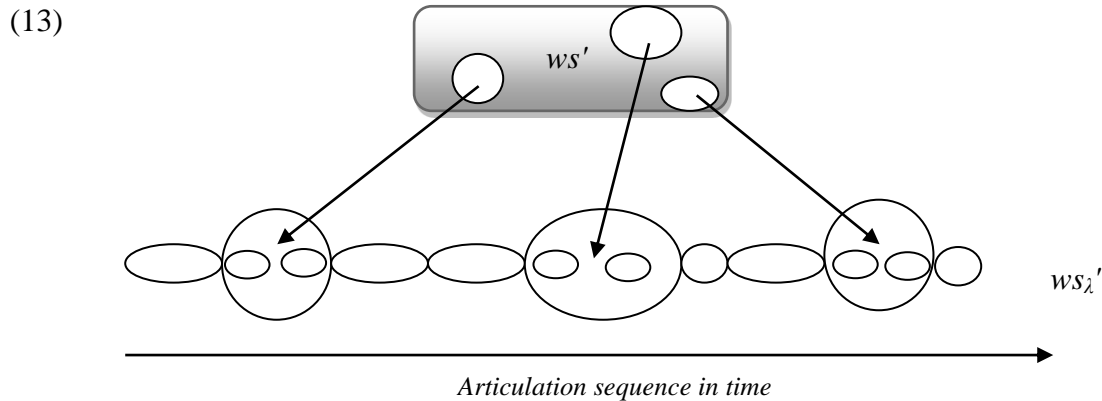
⁸ I use the term 'cognitive event' to denote any kind of cognitive activity related to the processing of a world stimulus. Note that in RadEx, a world stimulus can also be cognition-internal: In this case, a certain state of cognition serves as the stimulus for further cognitive activities, such as cogitation (intra-individual communication).

⁹ Compare Indo-European **uoidh-* 'to see' > Lat. *videre* 'to see', Greek εἶδομαι 'to be seen, appear', but οἶδα 'to know' (lit. 'having seen'), OHG *wizzan* 'to know'. See Evans & Wilkins (1998, 2000) for the correlation audition-knowledge and language.

¹⁰ The article is one of the many commentaries on Pylyshyn 1999 in the same journal (*Behavioral and Brain Sciences*).

¹¹ According to this definition, language is respiratory in nature. This does not mean this is the sole means to couple cognitive events with an expression model. Alternatively, human beings can make use of e.g. gesture motorics (resulting in natural sign languages) or (less conventionalized) mimic (facial) motorics etc. As for the articulatory domain, RadEx adopts the claim by Jürgens (2000:1): "Neuroanatomically, the step from genetically determined controlled vocal patterns is associated with the emergence of a direct connection between the motor cortex and the laryngeal motoneurons, a connection lacking in subhuman primates." As Jürgens points out, this connection is strongly related to learning. Accordingly, we cannot claim that there is a primary, iconic relation between the mirroring of WS properties (e.g. actions), the processing of their virtual 'images' in terms of motorics and the corresponding motorization types of articulation (except for sound symbolic features). In this sense, RadEx only partly adopts the 'neural exploitation hypothesis' (see Gallese & Lakoff 2005, Gallese 2007).

only in sequences requires that ws' become aligned according to one sequence *after* the other, be it on the phonetic, morphological, or syntactic level. Linearization may in parts be iconic with respect to the gestalt of the input, clustering those segments of the input image that are processed 'together'.¹² However, the basic principle of linearization is preserved for two reasons: First, the clusters are arranged one after the other; and second, many such clusters show an inherent linear order especially if they are reflected by compounds:



The general attitude of cognition towards an incoming World Stimulus is called *attention flow* in RadEx (see Schulze 2004 for details). In (13), the attention flow is schematized according to a language-based expression model (λ)¹³, resulting in a contorted variant of ws' , namely ws'_λ . Another schematic effect becomes relevant if we include the attitude of cognition towards the effect the processing of ws'_λ may have on other cognitions (*information flow*). Here, cognition is attentive towards a ws' to the extent it is stimulated by communicative frames, roles, scripts, all of which are grounded again in experience (the resulting schema is called *attention information flow* (AIF) in RadEx).

The non-Objectivistic layout of RadEx means that properties of 'real world' events do not figure as primary descriptors for linguistic 'products'. In other words, it is not the 'real world' that is reflected in or symbolized by linguistic expressions, but only their cognitive 'images' already contorted by the attention flow as well as by primary schemas etc. (this is what RadEx calls *diairesis*).¹⁴ In this sense, it does not matter, for instance, whether an 'event' in the real world is embedded into a causal chain or not: Causality is defined as a secondary constructional schema grounded in the conceptual enrichment of basic schemas such as *change, motion, and inference (reification in gestalt theory)*.¹⁵ In addition, we can incorporate the notion of Force Dynamics (Talmy 2000; see Schulze 1998 for a similar approach), however in a much broader sense than proposed by L. Talmy (also compare Cheng 1997). This point will be elaborated in section 3. The non-Objectivistic approach thus suggests that in fact everything that is conceptually present in language is the result of ws' processing, not of the real world stimulus as such. This hypothesis also includes the assumption that the basic

¹² Linearization is already present in visual perception (eye movement). Most likely, parameters of eye movement already pre-shape or influence linguistic linearization (see fn. 27) .

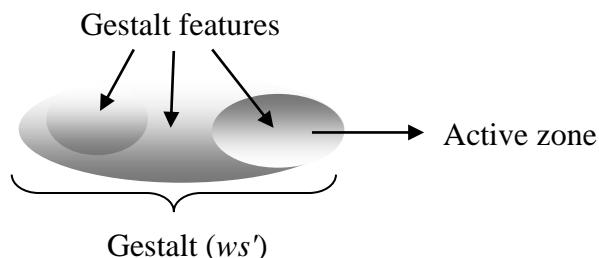
¹³ RadEx hypothesizes that the attention towards a WS ($> ws'$) is linguistically indexed if it is loaded with a language-based expressive ($>$ communicative) presetting.

¹⁴ This aspect is nicely expressed in Bertolt Brecht's apophthegm: "What do you do," Mr. K. was asked, "if you love someone?" - "I make a sketch of the person," said Mr. K., "and make sure that one comes to resemble the other." - "Which? The sketch?" - "No," said Mr. K., "the person." (German version: Bertolt Brecht, *Wenn Herr K. einen Menschen liebte*. In: *Geschichten vom Herrn Keuner*. Frankfurt/Main 1971:33 (stb 16)).

¹⁵ The present framework assumes that 'causality' is not a 'basic' human concept. This hypothesis is corroborated by the fact that lexical expressions of causality concepts are usually derived via metaphorization or represent more recent borrowings based on source terms such as Latin *causa*, Arabic *sabāb* etc.

pattern of *ws'* diairesis is marked for a *gestalt* that is interpreted as an 'event' (in the broadest sense). Accordingly, a WS is not mirrored in terms of individual 'building blocks' that combine to higher level structures. Rather, gestalt properties are secondarily isolated in terms of generic indexes that highlight what one may call an 'active zone' of the gestalt (taking on a term coined by R. Langacker (2000)). The drawing in (12) attempts to illustrate this aspect:

(14)



In the framework of RadEx, an active zone is that segment (property) of an *event image* that is typically or conventionally present and processed with respect to the event image as well as in related images.¹⁶ Children learn to construe such salient properties as more or less independent entities that are marked for 'object permanence' (Piaget 1954) and 'relational permanence' (Schulze 2001). Object permanence can be understood as the typical way of appealing to a memory segment (experience) that has become 'stable' in time (or: that can be activated by imagination). Resulting conceptual segments thus gain *referential* properties. Relational permanence is less discrete: Usually, the relation between two or more referential segments present in an event image becomes stabilized ('permanent') in combination with a hypothesis about the presence of typical referential entities. For instance, the concept <SWIM> includes knowledge about a 'swimmer' and some kind of liquid substance in which the act of swimming takes place. In this sense, relational permanence is much closer to the original event image than isolated gestalt properties construed as referential entities.¹⁷

As has been said above, the isolation and reification of gestalt properties is a secondary process that links a *ws'* to memory segments (both individual experience and convention). In other terms: Segments isolated from an event image are cognitive artifacts, *not* primary entities that would constitute an event image. Still, language offers especially lexical tools to manipulate the attention flow towards a WS by offering standard options for the reification of gestalt properties: If people have learned to use a typical articulation template in co-activation with a given concept, it is rather likely that they will 'parse' an adequate event image for just this pattern. In this recursive sense, a secondary cognitive artifact may become a primary experiential tool for the diairesis of a given WS.

3. Cognitive Transitivity

What has been said so far mainly concerns some basic aspects of the RadEx framework. Nevertheless, the arguments brought forth lay the ground to approach the questions addressed in the first section of this paper. Below, I summarize the main points:

¹⁶ The RadEx term 'event image' has nothing in common with Langacker's "canonical event model" that is defined as follows: "[T]he canonical event model represents the normal observation of a prototypical action" (Langacker 1991:285). The non-Objectivistic perspective taken by RadEx suggests that 'action' is by itself a cognitive construction. It cannot be 'observed' but only construed and projected upon an event image. The event image itself is defined as the final output of the contortion processes related to a given *ws'*. In this sense, 'event' is a cover term that refers to the construction of both states *and* dynamic processes.

¹⁷ Note that the use of 'verbs' as nominals (verbal nouns etc.) turns relational permanence into object permanence.

1. Parameters of vision pre-structure the gestalt of event images.
2. Event images (*ws'*) constitute the basis for cognitive processing, *not* events.
3. Event images are contorted (among other things) by memory, primary schematic processes, social knowledge, and (if loaded with a language-based expression model) linguistic features.
4. Event images are *gestalts* that are profiled for active zones and that are processed according to features of object permanence and relational permanence.
5. Event images are linguistically contorted because of the underlying expression model (dimension reduction, compare Oberschelp 2007), resulting in linearization).

Starting from the first hypothesis we can claim that the basic diairetic parameter in vision pre-shapes the structure of the *ws'* gestalt. It is a well-known observation that, with human beings, vision is dominated by binocular disparity (Qian 1997).¹⁸ As one of the results, three-dimensional vision becomes organized with the help of the *figure-ground* schema (itself an emergent property that is also present e.g. in hearing, see Blauert 1996).¹⁹ The overall relevance of this schema is described by Kurt Goldstein as follows:

Any excitation in the nervous system has the character of a figure/ground process. Any performance invariably shows this figure/ground character (...). Figure and background can be discriminated as readily in speaking, thinking, feeling, etc. (Goldstein 1963:12-13)

The figure-ground schema (F/G) conditions that any *ws'* is processed with respect to a more central, salient, and confined structure the borders of which set it apart from its 'background'. In fact, we have to deal with a mutual relationship: No figure without ground and no ground without figure. Or, as Rosalind Krauss has put it:

(...) a sense that painting's meaning was to be found in the simultaneous separation and intactness of figure and ground, in the gestalt's operation as the concordance between absolute difference (figure versus ground) and complete simultaneity (no figure without ground). (Krauss 1994:216)

The F/G schema permits us to interpret the gestalt structure of event images, especially if they are loaded with a language-based expression model (linearization). Accordingly, the gestalt of any event image - itself being the basic mode to mirror (and contort) an incoming WS - is processed by isolating a figure domain embedded into a corresponding ground. The mutual, vector-like relation between these two gestalt properties is construed as a 'relator', by itself an inferential property that only shows up in 'changes' (both positive and negative) with respect to the position of F and G or in changes of F or G properties. It should be noted that in RadEx, the F/G vector (in short: F→G) is not necessarily bound to spatial organization that isolates a ground as being 'behind' a figure. Rather, ground is defined as that gestalt segment of an event image that somehow 'surrounds' figure or that emerges from the isolation of figure. In other words: Even though F→G is grounded in spatial vision, it has highly 'abstract' (or: radical) properties that are blended with spatial segments. This includes (among others):

- | | | |
|------|--|--|
| (15) | Figure
<i>Smaller</i>
<i>With boundaries</i> | Ground
<i>Larger</i>
<i>Without boundaries</i> |
|------|--|--|

¹⁸ Qiang defines binocular disparity as follows: "We perceive the world in three-dimensions even though the input to our visual system, the images projected our two retinas, has only two spatial dimensions. How is this accomplished? It is well known that the visual system can infer the third dimension, depth, from a variety of visual cues in the retinal images. One such cue is binocular disparity, the positional difference between the two retinal projections of a given point in space." (Qian 1997:359).

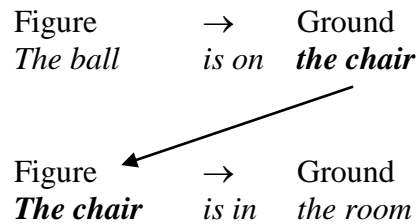
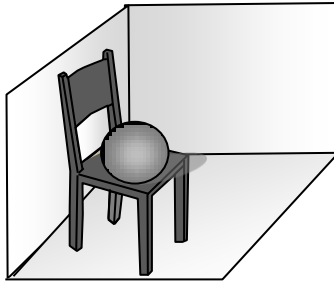
¹⁹ The reader may recall the well-known dictum coined by the biophysicist Georg von Békésy: "The purpose of the ears is to point the eyes".

More accessible
More salient
More mobile

Less accessible
Less salient
Less mobile

Cognition is thus more attentive towards fixing the figure portion of an event image. As a matter of fact, a ground can only be primarily accessed if it is turned into a figure embedded into a 'new' ground. For instance:

(16)



This observation has two consequences: First, the ground domain may be further subcategorized, resulting in multiple grounding.²⁰ The ground that serves to identify a figure is called *primary ground* that includes features of figure as soon as it is linked to another (secondary) ground. The link between a primary and a secondary ground plays the same role (*relator*) as the link between figure and ground:

(17) F → G₁ → G₂

Second, the preference to 'parse' an event image (*ws'*) for a figure domain can result in the neglect of the ground domain (termed *masking* in RadEx) or in its inclusion in the relator domain (*incorporation*, see below).

The isolation of the F/G domains conditions that gestalt properties of a given event image (*ws'*) become associated with 'typical' memory segments. Originally, this association is deictic in nature: The gestalt property at issue is deictically related to an experiential feature, which again leads to a (primitive) event image: For instance, in (14) the event image <BALL ON CHAIR> is based on the deictic identification of the figure domain with the help of the qualification <IS BALL>:

(18) Figure → Ground
 [what/this] is ball(-like)

In this sense, the event image <BALL IN CHAIR> originally reads:²¹

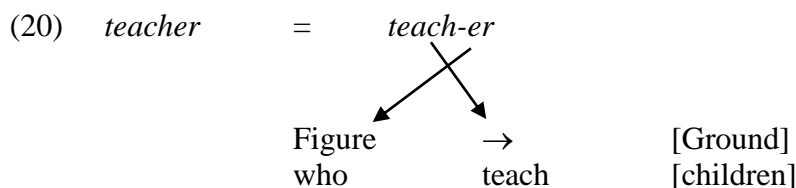
²⁰ The number of ground segments (or chunks) processable in short term memory is limited. Recent research suggests a mean memory capacity in adults of 3 to 5 chunks (Cowen 2001), thus revised the famous 'magical number 7 plus minus 2' (Miller 1956). This assumption goes together with the linguistic observation that in non-technical, non-literary texts, the number of overt referential segments rarely exceeds three/four (see Schulze 2004).

²¹ Note that this analysis does not aim at interpreting the expressions *ball* and *chair* in terms of a token-type relation that would label a token as Figure and a type as Ground. The grounding of tokens in a type is seen as a secondary experiential extension of the perception-based F/G-relation.

(19)

Figure			→	Ground		
Figure	→	Ground		Figure	→	Ground
[[what	is]	ball(-like)]	is on	[[what	is]	chair(-like)]

Note that the assumption of fractal structures like (17) is typical for the architecture of RadEx (see Schulze 2010, in press). Geometrically, we have to deal with the inversion of a so-called Pythagoras Tree (Lauwerier 1991), that is a plane fractal, here constructed from $F \rightarrow G$ gestalts. The deictic structure is usually condensed to referential units ("what is ball-like is a *ball*"), even though the event-based nature of referential expressions may be frequently preserved in the corresponding lexical expressions:



Example (20) illustrates that referential expressions are not necessarily represented by nouns or the like. In RadEx, nouniness is a special way of condensing (deictic) event images. Still, it becomes clear that the standard cognitive 'interpretation' of gestalt properties related to the figure and ground domains is marked for referentiality (*object permanence*). This domain is indicated by \mathfrak{R} in RadEx. The (basically inferential) relator often shows up as a verb-like segment, but again we have to bear in mind that *verbs* are a language-dependent category. The *relator* (\rightarrow) may have many other linguistic instantiations, such as prepositions, conjunctions, case marking (*relational echoes*, see Sallaberger & Schulze 2007) etc. In sum, the figure ground schema that applies when cognition becomes attentive towards a WS is coupled with a cognitive model that reads:

(21) $\mathfrak{R} \leftrightarrow^{22} \mathfrak{R}$

The formula reads: *Any event image is processed according to a model that links (at least) two referential segments with the help of a relator.* This formula is called 'Cognitive Transitivity' in RadEx. Contrary to other approaches, RadEx does not treat Cognitive Transitivity as a subclass of event images that would be determined by the 'semantics' of the verbal expression at issue. Accordingly, Cognitive Transitivity does not depend on the presence of specific semantic features associated with the referents such as 'Proto-Agent' or 'Proto-Patient' (Dowty 1991), role archetypes as proposed by Langacker (1991),²³ or agency and causation (Turner

²² I use the symbol \leftrightarrow to denote bidirectional transitivity. As soon as asymmetric properties become relevant, the symbol is turned into \rightarrow or \leftarrow .

²³ "The archetype agent is a person who volitionally initiates physical activity resulting, through physical contact, in the transfer of energy to an external object. Its polar opposite is an archetypal patient, an inanimate object that absorbs the energy transmitted via externally initiated physical contact and thereby undergoes an internal change of state" (Langacker 1991:285). Note the infelicitous use of Outer World terms (such as 'person', 'object') in the context of cognitive event images. The same holds for Croft's definition of transitivity: "[T]he initiator is an agent exercising his/her volition, and the endpoint undergoes a complete, even irreversible, change of state. The conceptual explanation for the prototypical character of this situation type is that this is the most clearly individuated situation type (...). An agent acting from his/her own volition has no salient antecedent cause, and a patient that ends in a state, especially an irreversible state, has the least likelihood of bringing about subsequent events" (Croft 2000:60). Taylor (2002:415-428) at least recognizes considerable degrees of syntactic variation within transitivity (but note Taylor 1998:187: "The transitive prototype involves an agent (encoded by the

1996). According to RadEx, such semantic or conceptual features are neither archetypes nor prototypical, but emergent properties that stem (among others) from the interaction of the $\mathfrak{R} \rightarrow \mathfrak{R}$ model with the $F \rightarrow G$ schema and its dynamics, see below. In addition, RadEx assumes that every linguistically intransitive structure is cognitively transitive. A key argument stems from the observation that in quite a number of languages, the referent typically called the 'object' is encoded in just the way a spatial referent is encoded in intransitive clauses. Randomly taken examples are (Classical Arabic and Latin):

(22) *dahaba* *s-sūq-a*
 go:PERF:3SG:M DEF-market-ACC
 'He went to the market.' (Haywood & Nahmad 1965:392)

kāna *l-bustān-u* *kabīr-a-n*
 be:PERF:3SG:M DEF-garden-NOM large-ACC-INDEF
 'The garden was large.' (Haywood & Nahmad 1965:105)²⁴

fataḥa *l-walad-u* *l-bāb-a*
 open:PERF:3SG:M DEF-boy-NOM DEF-door-ACC
 'The boy opened the door.' (Haywood & Nahmad 1965:99)

(23) *cum* *autem* *ven-iss-et* *domu-m*
 when thus come-PLU-3SG house-ACC
 'When he had thus come into the house' (Matthew 9:28)

One might argue that - as for the Latin example - the spatial expression is encoded as a referent in O-behavior ('object'). However, this hypothesis goes against the assumption that the semantics of 'functional' case forms are metaphorically derived from spatial semantics (see Schulze 2009 among many others). In this sense, the term *domum* has retained the invariant component of the semantics of the accusative.

Analogically, we can reinterpret case-marked prepositional clauses as cognitively transitive, compare (German):²⁵

(24) *Ich* *ging=in* *d-en* *Garten*
 I:NOM go=into:PAST:1SG ART-SG:M:ACC garden
 'I went into the garden.'

Ich *sah* *d-en* *Hund.*
 I:NOM see:PAST:1SG ART-SG:M:ACC dog
 'I saw the dog.'

(25) *Ich* *stand=auf* *d-er* *Wiese*
 I:NOM stand=on:PAST:1SG ART-SG:F:DAT meadow
 'I stood on the meadow.'

subject nominal), which intentionally acts on a patient (the direct object nominal) so as to effect a change-of-state in the patient.").

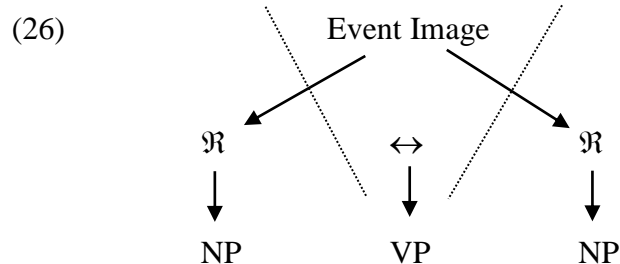
²⁴ The fact that Classical Arabic does not clearly distinguish between nouns and adjectives allows to read the sentences as follows: 'the garden was (< *became towards) a large one.'

²⁵ Note that in the so-called Kanak variety of German (language of youngsters of the second and third immigration generation), the intransitive pattern of motion verbs is even closer to that of transitives, compare *isch geh bahnhof* 'I go [to the] station', *isch mach dich messer wenn du nicht kino kommst* 'I tie you into knots if you do not go (with me) to the cinema' (note the Turkish-based use of *machen* 'to do' (= Turkish *etmek*) as a light verb incorporating the concept *messer* 'knife'.

<i>Ich</i>	<i>half</i>		<i>d-er</i>	<i>Frau</i>
I:NOM	help:PAST:1SG		ART-SG:F:DAT	woman
'I helped the woman.'				

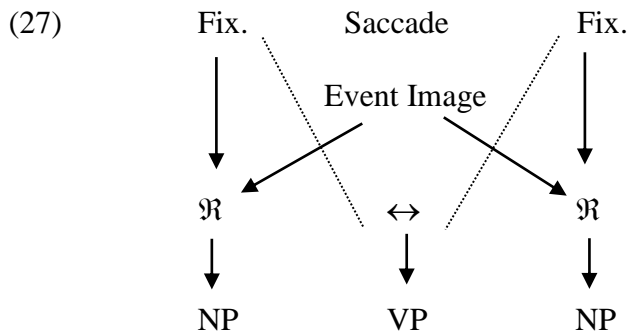
It goes without saying that the intransitive clauses given above have a different behavioral potential than their transitive counterparts (they are less liable to morphosyntactic passivization, for instance). In addition, they lack the metaphorical dimension typically present with referents in A-function (see below). However, this does not argue against the assumption that linguistically intransitive clauses are transitive from a cognitive point of view. In the light of Radical Experientialism, intransitivization results from secondary processes that are related to the structure of the attention information flow (AIF, see Schulze 2004).

Before turning to this point it is important to recall that the relator that links a figure and a ground is inferential in nature: The process of e.g. *reading* or *hitting* cannot be imagined without invoking at least rudimentary referential concepts. The fact that we can nevertheless observe the lexical expression of such relators (e.g. in terms of *verbs*) leads to the assumption that such lexical expressions do not reflect the relator as such, but the whole event image. Accordingly, *verbs* (better: verb phrases) are the meronymic expression of (linguistically speaking) clauses.²⁶



I assume that there is an iconic relation between the sequencing of fixation and saccades in visual perception and cognitive transitivity: Fixation periods are highly informative, whereas no information is processed during periods of saccades (eye movement), see Fulton 2000. Fixation lays the ground for object recognition and, once entrenched, for object permanence, whereas saccades set cognition into a state of 'blindness'. Cognitive blindness (or: cognitive saccades) can be regarded as that state of cognition that allows it to draw inferences from given referents pinpointed during fixation. The ensemble of a fixation-saccade-fixation sequence is construed in terms of a common gestalt that evolves into the matrix of event images. In this sense, cognitive verbs are cognitive saccades, and referents are cognitive fixations. The scheme in (26) can be thus augmented as follows:

²⁶ In RadEx, the terms Noun Phrase (NP) and Verb Phrase (VP) are used to denote the linguistic representation of those gestalt properties of an event image that are canonically isolated during the process of diairesis. Their immediate conceptual correlates are Referent (\mathfrak{R}) and Relator (\rightarrow).



The schemes in (26) and (27) also help to explain what is meant by verbal valence (as addressed in section 1): The 'valence' of a verb emerges from the structure of the event image it represents, not from the verb itself.²⁷ This assumption is related to models that link 'valence' and frames/scenes (Fillmore) or to the hypothesis set up by L. Tesnière according to which "le noeud verbal [...] exprime tout un petit drame. Comme un drame en effet, il comporte obligatoirement un procès et le plus souvent des acteurs et des circonstances" (Tesnière 1959:102). The meronymic nature of the verb phrase can best be illustrated if we refer to verbal nouns, compare:

(28) *Flying is dying.*²⁸

Here, the two event images "someone is=flying somewhere" and "someone gets=into death" are represented solely by the verbal segment. The condensed expression of event images has strong referential properties. It can hence be used to identify appropriate segments of more complex gestalts as it is true for (28). The corresponding part of the Pythagoras tree reads as follows:

(29)

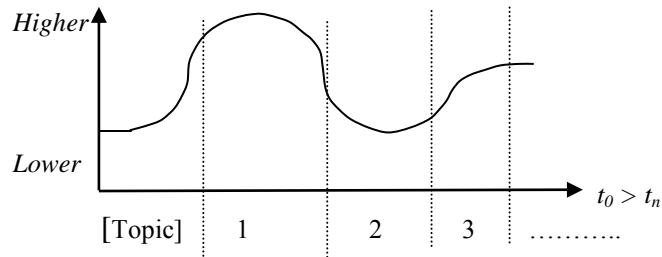
	R		↔	R		
R	↔	R		R	↔	R
X	flies=to	[Y]		X	gets=into	death

As has been said above, the gestalt of the event image (schematized according to the $R \rightarrow R$ vector) is always construed with the help of the figure ground schema. In fact, both vectors, namely $R \leftrightarrow R$ and $F \rightarrow G$ are dependent on each other. But whereas $F \rightarrow G$ is grounded in the architecture of the perceptual system, $R \leftrightarrow R$ is strongly related to memory. To put it differently: $F \rightarrow G$ is grounded in perception, $R \leftrightarrow R$ is grounded knowledge. The interaction of the two schemas conditions a number of emergent properties the quality of which depends on the way the attention information flow is articulated. Above, it has been said that a *ws'* can be loaded with a language-based expression model that provokes the linearization of $F \rightarrow G$ (and thus $R \leftrightarrow R$) expressions. Linearization, however, has a important effect on the individual segments: The degree of attention varies in such linear structures:

²⁷ A more accurate version of this definition has to take into account the categorization of the event image itself that is mainly grounded in convention. This categorization may include the stereotypical masking of one of the referents.

²⁸ Retrieved from http://www.treehugger.com/files/2006/03/george_monbiot.php (1.8.08).

(30)



Accordingly, the first chunk in a linear sequence (that may be preceded by a field that takes up a topic chunk) gain rather high attention followed by a chunk of lower attention.²⁹ The third chunk usually is slightly higher in attention than the second one, but often lower than the initial chunk. If we apply this scheme to the $F \rightarrow G$ vector we can assume that it is marked for a basic asymmetry that results from the degree of attention correlated with each chunk. In a standard interpretation, F would be marked for a high value, whereas \rightarrow and G are marked for lower values. Note that the correlation of $F \rightarrow G$ and this type of attention flow is highly conventionalized and language-dependent. In addition, it may be manipulated with the help of diathetic processes such as foregrounding (passivization, antipassivization) etc. Nevertheless, (30) allows the hypothesis that the $F \rightarrow G$ vector entails a syntactic value that tends to highlight the figure domain and to shadow the ground domain.

On the other hand, the $\mathfrak{R} \leftrightarrow \mathfrak{R}$ vector tends to be marked for conceptual, memory-based values that are ultimately derived from actional patterns of human behavior. These patterns are grounded in what has been termed the Perception Action Cycle (PAC):³⁰

[...] directed behaviors of animals comprise continuous cyclic relations between the detection of information and the performatory and exploratory activities that serve, in significant part, to facilitate that detection and which, in turn, are guided and shaped by it (Swenson & Turvey 1991:319)

This cyclic pattern can also be paraphrased as follows: The environment is perceived in accordance with the motion ($>$ behavior) of an active organism *in* it. Individuals move in the world in order to perceive and perceive in order to move (see Vernadsky 1929). The 'object' in the Outer World that helps to *inform* (or, phylogenetically speaking: *feed*) the individual is judged via perception according to the 'question' whether the effect compensates for the effort to 'reach' it. This vital behavioral pattern results in another schema that is based on 'force': The individual interprets its energetic (or: informational) state as 'force' and, whereas an 'object' in the Outer World is related to this feature in accordance with the individual's experience with former representatives of this 'object'. The default is a high force value for the individual and an α -value³¹ for the other 'object'. The resulting vector is $F_o \rightarrow \alpha F_o$. In case the 'object' is thought to have antagonistic force³² (termed *counterforce* (cFo) in RadEx), the individual may be stimulated to bring in stronger force or to respect the cFo feature of the object. The grading of F_o (actor/agonist) and αF_o (perceived object/antagonist) leads to important types of pragmatic and semantic variation, especially if expressed linguistically. One prominent type is the splitting of O (e.g. honorific *pars pro toto* (e.g. the emergence of the Slavic O-split based

²⁹ I assume that there is an iconic relation between the attention flow and the sequencing of fixation and saccades in visual perception, see Fulton 2000...

³⁰ There are numerous ways of defining and describing the Perception Action Cycle. Here, I take the view point of ecological psychology.

³¹ An α -value is given, if the value is irrelevant in the context of a physical or mental action.

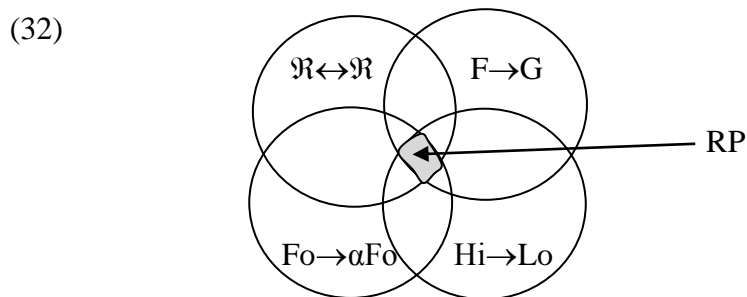
³² See Talmy 2000 for his use of the terms agonist and antagonist in Force Dynamics. The underlying terminology stems from the structural analysis of narratives, starting with Propp 1928, also compare Beaugrande & Colby 1979 and Wildgen 1990.

on the use of the genitive-partitive), differences in directional marking (e.g. the Spanish opposition Accusative vs. Dative/Lative), another one the splitting of A (actional vs. potential vs. affected, etc., compare the Udi pattern A:ERG (actional) vs. A:POSS potential vs. A:DAT affected, see Schulze [forthcoming]). In addition, modal features like 'limited control' (*finally managed to*) may emerge (as in Salish languages). Further examples are discussed in Schulze 1998.

The three vectors mentioned in this brief presentation of the RadEx approach are structurally coupled³³ with the AIF feature 'higher-lower'. The basic correlation is:

(31)	Experience	\mathfrak{R}	\leftrightarrow	\mathfrak{R}	Semantic
	Perception	F	\rightarrow	G	Syntactic
	PAC(?)	Fo	\rightarrow	$\alpha\text{Fo} \sim \text{cFo}$	Semantic/Pragmatic
	Attention Flow	Higher	Lower	Pragmatic

As a result, relational primitives (RP) emerge (see Schulze 2000) that combine pragmatic, semantic, and syntactic features:³⁴



The relational primitives (S, LOC, A, O, IO, IA, AO) emerge at the 'interface' of the four relevant schemes that again copy their basic properties onto the primitives. The most unmarked type of processing a *ws'* is characterized by the following transitive relation:

$$(33) \quad \mathfrak{R}_{F/\text{Fo}} \quad \rightarrow \quad \mathfrak{R}_{G/\alpha\text{Fo}}$$

The type of relational primitive emerging from this interaction depends on various factors. The strong (cognitively) localistic orientation of RadEx suggests that the $F \rightarrow G$ vector is loaded with the RP correlation $S \rightarrow \text{LOC}$, which would read: $S \rightarrow \text{LOC}$ is a linguistic schema of event images that relates a F-referent (F) to a G-referent (both stative and dynamic). The metaphorization of this schema starts from the overall hypothesis that what is perceptually salient is *before* the non-salient segment. The well-known metaphorization path *space > time > cause* determines that F is loaded with Fo-properties resulting in the relational primitive A (or IA). The LOC-domain is analogically metaphorized to O (or IO).

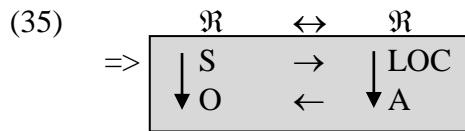
(34)

\Rightarrow	\mathfrak{R} \leftrightarrow \mathfrak{R} \downarrow S \rightarrow \downarrow LOC \downarrow A \rightarrow \downarrow O
---------------	---

³³ Structural coupling was first described comprehensively by Humberto Maturana: "I have called the dynamics of congruent structural changes that take place spontaneously between systems in recurrent (in fact recursive) interactions, as well as the coherent structural dynamics that result, *structural coupling*" (Maturana 2002:16-17).

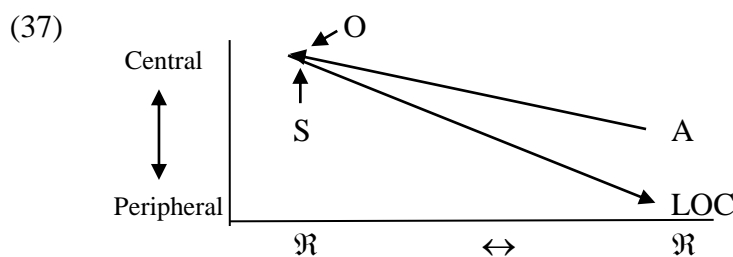
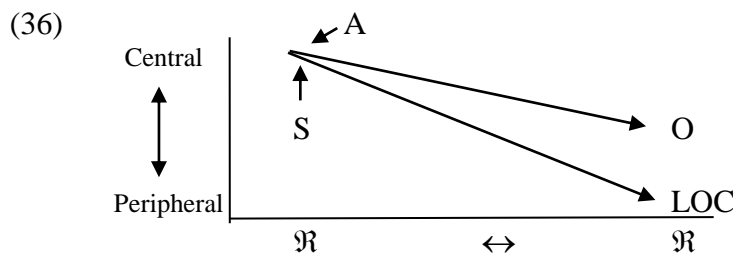
³⁴ See fn. 3.

This pattern is marked for a perspective that interprets the causal vector $A \rightarrow O$ according to the linear sequence 'no cause (*central*) without effect (*peripheral*)'. Therefore, A becomes associated with S, leading to the standard accusative pattern $S=A;O$ (A-centered). The revised perspective is taken in an ergative behavioral pattern:



Here, the effect domain is more central. The scheme thus reads: 'no effect (*central*) without cause (*peripheral*)'. As a result, O becomes associated with S ($S=O;A$ or 'O-centered').

It is clear that the two endpoints of the scale of the accusative ergative continuum ($S=A;O$ and $S=O;A$) depend on their structural coupling with the source domain $S \rightarrow LOC$. This pattern differs from $A \rightarrow O$ especially with respect to the degree to which the ground domain is expressed. In $F \rightarrow LOC$, ground (LOC) is schematically associated with the periphery. By 'peripheral' is meant that a referent gains lesser cognitive attention than the central one. In $A \rightarrow O$, the secondary domain (O in $S=A;O$ and A in $S=O;A$) is less peripheral due to the fact that the agonist/antagonist 'role' becomes apparent especially if its counterpart is overtly marked, too. (33) illustrates the O/LOC gradient for an accusative pattern, (34) the A/LOC gradient for an ergative pattern.



As a result, linguistically intransitive structures emerge that are characterized by the masking of the peripheral domain (LOC). The reasons for the intransitivization of the language-based expression of a *ws'* can be related to the above-mentioned fact that $S \rightarrow LOC$ schemas are close to the functional mode of the $F \rightarrow G$ vector (source domain): It tends to exclude a ground from being further processed by the attention flow. Such masking strategies also underlie ambitransitives and what commonly shows up as incorporation. In the latter case, the O domain is no longer isolated from the relator domain. It loses its referential profile and becomes an adverb-like segment of the relator:



Conversely, standard intransitive patterns can be profiled as transitive structures by reifying the event as an entity. Here, the relator is expressed in terms of a rather general, nearly generic concept (*decorporation*), e.g.:³⁵

$$(39) \quad \begin{array}{l} F \quad \rightarrow \quad G \\ S \quad \rightarrow \quad \text{LOC} \\ I \quad \text{swim} \quad \emptyset \end{array} \quad \Rightarrow \quad \begin{array}{l} F \quad \rightarrow \quad G_1 \quad \rightarrow \quad G_2 \\ A \quad \rightarrow \quad O \quad \rightarrow \quad \text{LOC} \\ I \quad \text{do} \quad \text{swim} \quad \emptyset \end{array}$$

As a matter of fact, the options for the language-based expression of *ws'* waver between the overt expression of peripheral segments and their total masking. It remains the task of language-specific research to tell this story.

4. Conclusions

In the first section of this very sketchy approach to Cognitive Transitivity, I have presented a selection of observations that argue in favor of the hypothesis that linguistic utterances are marked for a preference to relate a verbal concept to roughly two noun-like segments. The limitation of space does not allow further corroboration of this hypothesis with the help of comprehensive typological data (Schulze (in preparation) will try to compensate for this deficiency). Nevertheless, the observations laid the ground to approach the hypothesis from a deductive point of view that is strongly devoted to reductionism. Reductionism does not necessarily speak against holistic models of cognition and emergentism, as it is occasionally argued (see e.g. Scott 2004). If we refer to ontological reductionism, we may likewise assume that basic physiological properties of human beings as well as basic schematic procedures of cognition represent (both phylogenetically and ontogenetically) the input of a multicausal scenario the emergent properties of which result in a complex functional network. In this sense, the radical (or: simple) structures of the input are mirrored as again 'simple' structures on a higher, more complex level. This mirroring process comes close to what has been described as fractal structures enriched by emergent features.³⁶ If we accept this approach, we can hypothesize that the structuring of linguistic utterances is ultimately motivated by very simple pre-linguistic patterns of perception and experience. These patterns (schemas etc.) construe primary event images (*ws'*) that are further manipulated by language-based expression models. The gestalt-like 'inner' perception³⁷ of event images is grounded (among others) in the isolation of referential units (*object permanence*) coupled with each other through the event image itself and its meronymic conceptualization (*relator*). This relational structure ($\mathfrak{R} \leftrightarrow \mathfrak{R}$) directs the attention towards the event image together with the perceptual schema $F \rightarrow G$. The schema determines the basic asymmetry of the relational structure that is in addition contorted by linearization processes. Accordingly, RadEx assumes that any event image is cognitively processed in terms of transitive constructions. Cognitive Transitivity hence shows up as a cognitive universal that underlies any linguistic construction (symbolizing more or less complex event images). It should be stressed that Cognitive

³⁵ Note that this is a structural analysis only. Naturally, the construction has strongly grammaticalized resulting in an emphatic variant of the underlying form.

³⁶ In this paper, I have neglected the question of voluntary perception and expectation (leading to modal structures, compare Krauzlis (2005:124): "The selection process itself is guided by a variety of complex processes, including attention, perception, memory, and *expectation*" (my emphasis). See Metzinger 2004 on the question of intentionality in neurophilosophy.

³⁷ RadEx does not make a systematic difference between the perception of Outer World stimuli and the attentive state of cognition towards the event images that mirror them. In this sense, cognition *perceives* its event images, even though there is no special perceptual device present (as far as we know).

Transitivity is not a basic conceptual unit but an emergent schema. Accordingly, we cannot expect to describe prototypical features of Cognitive Transitivity. Rather, prototypical effects become relevant in the conventionalization of the language-based expression of event images. Here, the ball is in the camp of Language Typology and Cognitive Grammar. However, without reference to Cognitive Transitivity linguistic studies in (in)transitivity seem to be at risk of starting from descriptive and analytic units that may be adequate only from a heuristic point of view. As soon as we look for the cognitive motivation of basic clause structures and their universal properties, we realize that language is not *un système où tout se tient*, but *un système où tout dépend*.

References

- Anagnostopoulou, Elena 2003. *The Syntax of Ditransitives*. Berlin etc.: Mouton de Gruyter.
- Brassett, Cecilia & Philip Brassett & Meiyun Lu 2006. *The Tujia Language*. Munich: Lincom.
- Baddeley, Alan D. 2000. The episodic buffer: a new component of working memory? *Trends in Cognitive Science*, 4: 417-423.
- Baddeley, Alan D. & Graham J. Hitch 1974. Working memory. In: Gordon H. Bower (ed.). *The psychology of learning and motivation: Advances in research and theory*, Vol. 8, 47-89. New York: Academic Press.
- Bakhtin, Mikhail M. 1986. *Speech Genres and Other Late Essays*. University of Texas Press.
- Beaugrande, Robert de & Benjamin N. Colby 1979. Narrative Models of Action and Interaction. *Cognitive Science: A Multidisciplinary Journal* 3,1:43 - 66.
- Blauert, Jens 1996 [1983]. *Spatial Hearing. The Psychophysics of Human Sound Localization*. Revised edition. Cambridge, Mass.: MIT Press.
- Cheng, Patricia W. 1997. From Covariation to Causation: A Causal Power Theory. *Psychological Review* 104: 367-405.
- Cowan, Nelson 2001. The Magical Number 4 in Short-term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences* 24,1:87-185.
- Croft, William 2000. *Verbs: aspect and argument structure*. Partial Draft, August 2000.
- Dixon, Robert M.W. & Aleksandra Y. Aikhenvald 2000. Introduction. In R.M.W. Dixon and Aleksandra Y. Aikhenvald (eds.), *Changing valence. Case studies in transitivity*, 1-29. Cambridge: Cambridge University Press.
- Dowty, David 1991. Thematic proto-roles and argument selection. *Language* 67,3:547-619.
- Engels, David 2007. Geometrie und Philosophie. Zur Visualisierung metaphysischer Konzepte durch räumliche Darstellungen in der Pythagoreischen Philosophie. In: Dominik Groß & Stefanie Westermann (Hrsgg.), *Vom Bild zur Erkenntnis? Visualisierungskonzepte in den Wissenschaften*, 113-129. Kassel: kassel university press
- Evans, Nicholas & David Wilkins. 1998. *The knowing ear: an Australian test of universal claims about the semantic structure of sensory verbs and their extension into the domain of cognition*. Institut für Sprachwissenschaft, Universität zu Köln. Arbeitspapier No. 32 (Neue Folge).
- Evans, Nicholas & David Wilkins 2000. In the mind's ear: the semantic extensions of perception verbs in Australian languages. *Language* 76,3: 546-592.
- Foley, William 1991. *The Yimas language of New Guinea*. Stanford: Stanford UP.
- Fulton, James T. 2000. *Processes in Animal Vision* {online} {Corona Del Mar, CA. USA} Vision Concepts, Available on the Internet: URL:<http://www.4colorvision.com>.
- Gallese, Vittorio 2007. Mirror neurons and the social nature of language: The neural exploitation hypothesis. *Social Neuroscience* 2007: 1-17.
- Gallese, Vittorio & George Lakoff 2005. The brain's concepts: The role of the sensorimotor system in reason and language. *Cognitive Neuropsychology* 22: 455-479.
- Glaserfeld, Ernst von 1996. *Radical Constructivism: A Way of Knowing and Learning*. London: Falmer Press.
- Goldstein, Kurt 1963. *Human nature in the light of psychopathology*. New York: Schocken. (First edition 1940, Cambridge, Mass.: Harvard UP).
- Haywood, John A. & H. M. Nahmad 1965. *A new Arabic grammar of the written language*. London: Lund Humphries.
- Hollingsworth Andrew & John M. Henderson 1999. Vision and cognition: Drawing the line. *Behavioral and Brain Sciences* 22,3: 380-381.
- Holšánová, Jana 2008. *Discourse, Vision, and Cognition*. Amsterdam/Philadelphia: Benjamins (Human Cognitive Processing 23).

- Hopper, Paul 2003. Transitivity: What a Difference Two Decades Make! PIONIER Workshop on Case, Valency and Transitivity, Nijmegen 2003. Panel paper abstract.
- Hopper, Paul & Sandra Thompson 1980. Transitivity in grammar and discourse. *Language* 56: 251-299.
- Hopper, Paul & Sandra Thompson (eds.) 1983. *Studies in Transitivity*. New York: Academic Press. (*Syntax and Semantics* vol. 15).
- Jürgens, Uwe 2000. A comparison of the neural systems underlying speech and non-speech vocal utterances. In: Bernhard H. Bichakjian & Tatiana Chernigovskaya, & Adam Kendon & Anke Möller, Anke (eds.), *Becoming Loquens. More Studies in Language Origins*, 1-13. Frankfurt am Main: Lang.
- Krauss, Rosalind E. 1994. *The Optical Unconscious*. Cambridge, Mass.: MIT Press.
- Krauzlis, Richard J. 2005. The Control of Voluntary Eye Movements: New Perspectives. *Neuroscientist* 11,2:124-137.
- Kulikov, Leonid & Andrej L. Malchukov & Peter de Swart (eds.) 2006. *Case, Valency and Transitivity*. Amsterdam/Philadelphia: Benjamins.
- Kuno, Susumo & Ken-ichi Takami 2004. *Functional Constraints in Grammar*. Amsterdam/Philadelphia: Benjamins.
- Lakoff, George 1987. *Women, Fire, and Dangerous Things*. Chicago, IL: University of Chicago Press.
- Langacker, Ronald 1987. *Foundations of Cognitive Grammar V. I: Theoretical Prerequisites*. Stanford: Stanford University Press.
- Langacker, Ronald 1991. *Foundations of Cognitive Grammar V. II: Descriptive Application*. Stanford: Stanford University Press.
- Langacker, Ronald 2000. *Grammar and Conceptualization*. Berlin: Mouton de Gruyter.
- Lauwerier, Hans 1991. *Fractals: Endlessly Repeated Geometric Figures*. Princeton, NJ: Princeton University Press.
- Lazard, Gilbert 2002. Transitivity revisited as an example of a more strict approach in typological research. *Folia linguistica* 36,3-4: 141-190.
- Maturana, Humberto 2002. Autopoiesis, Structural Coupling and Cognition: A history of these and other notions in the biology of cognition. *Cybernetics & Human Knowing* 9,3-4:5-34.
- Metzinger, Thomas 2004 [1993]. *Subjekt und Selbstmodell. Die Perspektivität phänomenalen Bewußtseins vor dem Hintergrund einer naturalistischen Theorie mentaler Repräsentation*. Paderborn: mentis.
- Miller, George A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81-89.
- Næss, Åshild 2007. *Prototypical Transitivity*. Amsterdam/Philadelphia: Benjamins.
- Oberschelp, Walter 2007. Bild und Wirklichkeit. In: Dominik Groß & Stefanie Westermann (Hrsgg.), *Vom Bild zur Erkenntnis? Visualisierungskonzepte in den Wissenschaften*, 29-46. Kassel: kassel university press.
- Piaget, Jean 1954. *The Construction of Reality in the Child*, translated by Margaret Cook. New York: Basic Books.
- Propp, Vladimir J. 1928. *Morfologija skazki*. Leningrad: Academia (Voprosy poetiki XII).
- Pylyshyn, Zenon 1999. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences* 22,3: 341-365.
- Qian, Ning 1997. Binocular Disparity Review and the Perception of Depth. *Neuron*, 18: 359-368.
- Sallaberger, Walter & Wolfgang Schulze 2007. Grammatische Relationen im Sumerischen. In: *Zeitschrift für Assyriologie* 2007,2: 163-214.
- Schulze, Wolfgang 1998. *Person Klasse, Kongruenz*, vol. 1,1-2. Munich: Lincom.
- Schulze, Wolfgang 2000. Towards a typology of the Accusative Ergative Continuum. *General Linguistic* 37:71-155.
- Schulze, Wolfgang 2001. *Selbstlernen und Selbstreflexion: Grundlagen einer Emergenz-Theorie der sprachlichen Interaktion auf der Basis der 'Grammatik von Szenen und Szenarien'*. Munich Working Papers in Cognitive Typology 1.
- Schulze, Wolfgang 2004. Pragmasyntax: Towards a cognitive typology of the Attention Information Flow in Udi narratives. In: Augusto Soares da Silva, Amadeu Torres, Miguel Gonçalves (eds.) *Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva*, 2 vols, 545-574. Coimbra: Almedina 2004.
- Schulze, Wolfgang 2007. Communication or Memory Mismatch? Towards a Cognitive Typology of Questions. In: Günter Radden, & Klaus-Michael Köpcke, & Thomas Berg & Peter Siemund (eds.), *Aspects of Meaning Construction*, 247-264. Amsterdam/ Philadelphia: Benjamins.
- Schulze, Wolfgang 2009. A new model of metaphorization: Case systems in East Caucasian. In: A. Barcelona, K.U. Panther, Günter Radden. L.L. Thorburg (eds.), *Metonymy and Metaphor in Grammar*, 147-175. Amsterdam & Philadelphia: Benjamins.
- Schulze, Wolfgang 2010. Einfalt (in) der Vielfalt: Reduktionismus in den Sprachwissenschaften und die Fraktale der Sprache. *Slowakische Zeitschrift für Germanistik* 2009, 2:11-41.
- Schulze, Wolfgang (in press). Sprache als kommunizierte Wahrnehmung. Ein Essay in sieben Teilen. In: Christine Jacquet-Pfau et al. (éd.), *Au mémoire de Jean-Marie Zemb*. Paris.

- Schulze, Wolfgang [forthcoming]. *A functional Grammar of Udi*.
- Schulze, Wolfgang (in preparation). *Aspects of Cognitive Typology*.
- Scott, Alwyn 2004. Reductionism Revisited. *Journal of Consciousness Studies* 11, 2:51–68.
- Swales, John 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- Swenson, Rod & Michael Turvey 1991. Thermodynamic Reasons for Perception-Action Cycles. *Ecological Psychology* 3(4):317-348.
- Talmy, Leonard 2000. *Toward a cognitive semantics*, vol. I: *Concept structuring systems*. Cambridge: MIT Press
- Taylor, John 1998. Syntactic Constructions as Prototype Categories. In: Michael Tomasello (ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, 177-202. Mahwah, NJ: Lawrence Erlbaum.
- Taylor, John 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Tesnière, Lucien 1959. *Elements de syntaxe structurale*. Paris: Klincksieck.
- Thompson, Sandra & Paul Hopper 2001. Transitivity and Frequency in Conversation. In: Joan Bybee and Paul Hopper (eds.), *Frequency Effects and Emergent Grammar*, 27-60. Amsterdam/Philadelphia: John Benjamins.
- Turner, Mark. 1996. *The Literary Mind: The Origins of Thought and Language*. Oxford: Oxford University Press.
- Vernadsky, Vladimir I. 1929. *La biosphère*. Paris: Alcan.
- Wildgen, Wolfgang 1990. Sketch of an Imaginistic Grammar for Oral Narratives. In: Karl Heinz Wagner & Wolfgang Wildgen (Hrsgg.), *Studien zur Grammatik und Sprachtheorie*, Vol. 2, 85-121. Bremen: Institut für Allgemeine und Angewandte Sprachwissenschaft, Universität Bremen.
- Wilks, Yorik 1978. Semantic primitives in language and vision. *Theoretical Issues In Natural Language Processing*. Proceedings of the 1978 workshop on Theoretical issues in natural language processing, 180-183. Urbana-Champaign, Illinois: The Association for Computational Linguistics.
- Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge, Mass.: Addison-Wesley Press.