SUBJECT

in the Meta-Informative Centering Theory

André WŁODARCZYK [1] and Hélène WŁODARCZYK[2]
Centre de Linguistique Théorique et Appliquée
Université Paris – Sorbonne
{Andre & Helene Wlodarczyk} <fdc@paris4.sorbonne.fr>

1. Theoretical framework

The problem of subject will be dealt with in the framework of our meta-informative centering theory[3] which enables us to build models of how *information* is linearised and expressed in human languages. This theory is based on several assumptions. (a) Linguistic messages are always partial with respect to reference: when communicating something about a situation in the world, speakers choose a point of view about it and communicate explicitly only this point of view. The whole situation must be completed by the hearer on the basis of his knowledge both of that of the speech-act situation and that of the world in general. (b) Two different levels must be distinguished: *informative* and *meta-informative*. As a matter of fact, what is usually called by linguists *informative structure* concerns rather the way information itself is conveyed through linguistic messages. By *information*, we thus mean the semantic relational content of utterances: this concerns situation frames[4], their participants and spatio-temporal anchoring. *Meta-information* concerns the way *information* is ordered[5]: to achieve the ordering of non-linear representations as texts (sequences of linguistic utterances), the speaker must select a *centre of attention* (henceforth CA) and "predicate" about it. Hence, in our centering theory, the term "predicate" is reserved for the meta-informative level and is used in its initial Aristotelian sense. Moreover, our model of language makes use of the definition of sign (i.e. the distinction between *form* and *content* originating in structural linguistics (Saussure and Hjemlslev). However, we consider that form as well as content are layered: (a) *form* has two levels: phonology (with prosody) and morphology (b) *content* comprises both semantics and pragmatics. Directly expressed

---

[1] Professor at Charles de Gaulle University (Lille 3) <andre.wlodarczyk@ univ-lille3.fr >.

[2] Professor at Paris Sorbonne University <helene.wlodarczyk@paris4.sorbonne.fr>.

[3] For an outline of the meta-informative centering theory cf. Wlodarczyk A. (2004, 2005) and Wlodarczyk A. & H (2006).

[4] Cf. Wlodarczyk A. 2003 c.

[5] Cf. "le module représentatif" and "le module informatif' in Gutiérrez-Ordóñez S. (2006).

contents are known as semantics and concern information about situations, whereas indirectly expressed contents, known as pragmatics, concern relations between speech-act participants, centres of attention and everything which introduces new information into discourse. As a matter of fact, information contained in utterances is often referred to solely by meta-informative pointers which are used to distinguish between the speaker's CAs. CAs appear in two different types of expressions. *Simple utterances* contain a Subject and either additionally or optionally an Object. *Extended utterances* consist in adding to the CAs of simple utterances (Subject and/or Object) one or two more CAs: Topic and/or Focus.

TYPE OF EXPRESSION	CENTRE OF ATTENTION	
	Global	*Local*
1.1. Simple UTTERANCE	SUBJECT	OBJECT
1.2. Extended UTTERANCE	TOPIC	FOCUS
2. TEXT / DIALOG	GENERAL THEME	PARTICULAR THEME

Table 1 Pivots of discourse

The difference between the CAs of simple utterances (Subject and Object) and those of extended utterances (Topic and Focus) consist in their meta-informative status (given or new). CAs of simple utterances are of the same meta-informative status as the predicate whereas CAs of extended utterances are in contrast with the rest of the utterance (the 'rest' is called *comment* for the Topic and *background* for the Focus). The meta-informative status of a Topic is *given* and that of its Comment is *new*. On the contrary, the status of a Focus is *new* whereas that of its Background is *given*.

Two different types of *new* and *given* meta-information should be distinguished[6]: (1) speech-act oriented (*anaphoric* as "known by hearsay" and *cataphoric* as "unheard of") and (2) information-oriented (situation types and their occurrences)[7]. From a cognitive point of view, semantic situations (information) are **typed** as *generic*, *general*, *habitual* or *potential*, and their **occurrences** can be defined as their dual counterparts : *specific*, *particular*, *occasional* or *actual* respectively (table 2).

Thus the first *given-new* opposition is **speech-oriented** (**s**-*given* or **s**-*new*) because it corresponds directly to meta-informative contents such as the conceptual pair *anaphora* and *cataphora*. The second one is **information- oriented** (**i**-*given* or **i**-*new*) because it refers to what is motivated by *information* (semantic situation types and their occurrences[8]).

---

[6] Cf. Shatunovskiy I. (2006).

[7] This view separates anaphoric/cataphoric information from that of information-oriented statuses (generic, general etc.). For detailed discussion see Wlodarczyk A. & H. (2006).

[8] The distinction *semantic situation type* vs. *semantic situation occurrence* can be compared to what is known as *continuant* and *fluent* entities in philosophical ontology, or *persistent* and *transient* information in information science. Needless to say that the *generic/specific* distinction recalls the Universal and Existential logical quantifiers.

Situation Types	Situation Occurrences
Generic > General	Specific > Particular
Habitual	Occasional
Potential	Actual

Table 2. Common knowledge-based informative motivation of aboutness

Let us now quote some examples of simple and extended utterances where the Subject of the simple utterance has been topicalised (in #3) or focalised (in #4). But, of course, topicalisation and focalisation may concern constituents of the simple utterance other than the Subject.

Simple utterance with a *given* Subject and a *given* Predicate:

#1 Satellites turn around the Earth.

Simple utterance with a *new* Subject and a *new* Predicate:

#2 A new satellite was launched yesterday.

Extended utterance with a Topic (*given* CA) and a *new* Comment:

#3As concerns satellite X 05, it was destroyed yesterday by a meteorite.

Extended utterance with a Focu*s (new* CA) and a *given* Background:

#4 It is satellite X 05 that was destroyed yesterday by a meteorite.

In this paper, we will concentrate on the problem of the global CA of simple utterances, namely the Subject. Further reference about the meta-informative centering theory can be found in our previous works.

2. Subject as Centre of Attention

Most linguistic theories consider Subject and Object as purely syntactic or semantico-syntactic notions. Our theory, however, defines these notions on the level of meta-information. The vast literature devoted to defining Subject reveals the difficulty involved in finding appropriate criteria. Despite these attempts, no satisfactory definition has yet been found which would fit typologically different languages, be they accusative, ergative, etc. It seems that this difficulty arises from the fact that most theories of Subject lack a separately defined meta-informative level and use composite "semantico-syntactic" devices[9].

We will show that a theory of Subject must thoroughly sort out what is inextricably mixed up in the syntax of human languages, namely the fact that syntactic rules reflect

---

[9] It is unnecessary to give a long list of authors, but one of the best specialists of ergativity R.M.W. Dixon entitled the first section 1.1. of his important book on *Ergativity* (1994). "S, A and O ; the Universal syntactic-semantic primitives". S means *intransitive subject*, A *transitive subject* and O *transitive object.* These symbols are most commonly used nowadays. The letter A suggests that the semantic notion "agent" is hidden under the term of "transitive subject". On the other hand, Manning 1996 stresses the need to distinguish between two different levels, that of grammatical relations and that of argument structure, but at the same time he writes "I am thinking of argument-structure as a syntactic representation (as in Bresnan and Zaenen 1990), while some work in LFG […] has suggested that a-structure is a purely semantic representation." (p. 35).

semantic and pragmatic aspects of relations. For example, such partly syntactic and partly semantic terms as *actors* (or even *actants* in French, Tesnière L. 1959) or the very use of the term *subject* by some linguists as syntactic, or the so-called "grammatical subject", and by some others as semantic, so-called "deep" or "semantic" or "logical" subject (a sort of active, volitional, animate or human subject, also even named "agent") lead to confusion when describing and comparing facts from different languages.

In generative grammar, the subject is the NP which dominates the whole tree structure of sentences. For this reason it has a global character, e.g. the sentence *Peter loves Mary* can be represented by the tree structure: S (NP, VP(V, NP)). In our framework, Subject is defined exclusively on the meta-informative level as that segment of a simple utterance which is centred as a *global* CA. A segment of an utterance is said to be centred if it is highlighted by meta-informative markers (these may be either prosodic, morphological or syntactic). The different means used in languages to mark CAs are treated as functors belonging to infons[10] with a distinguished relator bearing the label <about>. Due to the linear nature of linguistic messages which convey information about non-linear situations, the speaker must necessarily identify a noun (or noun phrase) to create a Subject for intransitive utterances or must select a Subject CA of a transitive utterance from among the participants of the semantic situation. Subject, in this sense, is understood, in agreement with traditional grammar, as "what is spoken about". Choosing one of the participants of the situation entails distinguishing it, making it hierarchically more important than the others. This choice is represented by the dominant place of the Subject NP in the uppermost left node of a constituency tree[11], in which Object NPs are in lower nodes than Subject NPs because Objects are daughters of VPs. In our terms, therefore, in constituency structure Objects are local CAs.

2.1. Subject and Object

Otto Jespersen (1924) already pointed out the similarity between Subject and Object calling them "first rank elements of the sentence". He argued that this similarity could be corroborated by the history of such languages as English or Danish, in which certain impersonal sentences with an Object were historically transformed into Subject-Verb sentences.

In our theory, both Subject and Object belong to the meta-informative level. The difference between them is that the Object is always a local CA, whereas the Subject is the global CA. We shall show further that a topicalised Object can fill a very similar meta-informative function in an extended utterance like the Subject in a non-extended utterance. Indeed, both Subject (as a meta-informative CA) and Topic (as a meta-meta-informative CA) are global CAs, respectively in a simple and in an extended utterance[12]. We shall see below that this meta-informative view of Subject and Object also makes it possible to explain the structure of ergative languages.

---

[10] Infons are borrowed from situation semantics (Barwise and Perry) and used in Wlodarczyk A. 2001 and to be published.

[11] Although our approach is not set in the generative framework, we accept the representation of the Subject as the uppermost NP of the tree representing the constituency relations in a sentence.

[12] Needless to say, any extended utterance is by definition an extension of a simple utterance.

2.2.Meta-informative Concord between Subject and Predicate

Simple (non extended) utterances differ from extended utterances in that the meta-informative status of the Subject is congruent with that of the Predicate. Accordingly, restrictions on the Subject NP and restrictions on the Predicate VP must have the same meta-informative statuses. As was pointed out in the introduction to this paper, two kinds of meta-informative statuses must be distinguished: *given*/*new* by speech-act (i.e.: **s**-*given*/**s**-*new*) or *given*/*new* by information (i.e.: **i**-*given*/**i**-*new*). Restrictions on the Subject are expressed by determiners (articles, adjective pronouns such as deictic, possessive, definite or indefinite quantifiers etc.,). Restrictions on the Predicate are expressed by verbal modifiers (of tense, aspect, mood etc.) of the form of inflectional morphemes or markers such as adverbs, particles, auxiliary verbs, etc. In a simple utterance there must be a categorial and intercategorial compatibility between the Subject determiners and the Predicate modifiers. Determiners and modifiers concern entities and situation frames indicated by the Subjects and the Predicates respectively, the semantic contents of each of them may be of both types: *qualia* and/or *quanta*; i.e.: the same linguistic expression may comprise alternately or simultaneously both *quanta* and *qualia*. However, although *quanta* and *qualia* only *motivate* the *given*/*new* statuses of utterances, their informative (semantic) meanings cannot contradict the meta-informative concord between Subject and Predicate.

#5 Every student has a car.

The above example suggests that the compatibility relations between *qualia* ('every' – quantifier) and *quanta* ('a' – indefinite determiner) has really a meta-informative character (rather than 'purely' informative). The meta-informative status of #5 is *given* which is motivated by the universally quantified Subject (*'every student'*) and by the *general* meaning of the Predicate (*'has a car'*) because, in general students, have no more than one car.

2.3.Subject and Semantic Roles

CAs and semantic roles belong to different levels, i.e. pragmatics (meta-information) and semantics (information) respectively. In the process of encoding (synthesis) and decoding (analysis) of utterances, the speaker (or the hearer) identifies the centres of attention and at the same time matches semantic roles with the entities which take part in the given semantic situation. In other words, among entities participating in a situation the speaker (or the hearer) identifies one of them and pushes it to the rank of the global CA (Subject). However, as for the Subject and the direct and indirect Objects at least, the speaker does not assign any precise semantic role to the participant identified in this way. It is therefore clear that the Subject (like the other CAs, i.e. Objects) may be mapped onto many different semantic roles.

In this approach, semantic roles are defined in a different way than, for example, in Fillmore's deep case theory (Fillmore Ch 1968) or even in the recent framework of FrameNet (Fillmore et al. 2002). Instead of using such concepts as *agentive*, *objective*, *experiencer*, etc., Wlodarczyk A. (to be published) distinguishes between various abstract ontological levels. On the highest level, he postulates generalised agents which are defined by the following ontological features: (1) control features (autonomy): *goal* and *feedback*; (2) emotive features (character): *desire* and *intention*, (3) epistemic features

(reason): *belief* and *cognition*, (4) communication features (language faculty): *verbal* and *visual*.

In accordance with this ontology, natural and artificial entities are suited to fill semantic roles of *agents* and *figures* respectively. Wlodarczyk A. further proposes to distinguish between three classes of abstract semantic roles (or macro-roles) which are ACTIVE, MEDIAN or PASSIVE. Let us also add that the ACTIVE and PASSIVE roles can be compared to Dowty's (1991) "proto-agent" and "proto-patient" respectively.

We shall mention here in passing that a deeper informative (semantic) analysis enables us to understand that the same NP in an utterance in fact fills several roles at the same time in a given situation[13]. For example:

#6 Peter bought a car for his daughter.

The semantic relation expressed by the verb *buy* in #6 can be analysed as composed of the two following component situations *(encapsulated procedures)* which underlie two modules of the given semantic utterance:

Schema (valence)	x buys y from w for z, where x = Peter, y = car, z = daughter, w = seller

Proc$_1$: `buy-sell`	`"w owns y"`
	`"x gives money to w"`
	`"w gives y to x"`
	`"x gets y from w"`
	`"x owns y"`

Proc$_2$: `offer`	`"x owns y"`
	`"x offers y to z"`
	`"z is a beneficiary of y"`
	`"z owns y"`

In the situation named by the English verb *to buy* in #6, both procedures may be represented declaratively as follows:

1. For the procedure Proc$_1$: `buy-sell`

- *Peter* fills the active role (agent : initiator >…> buyer)

- the *car* fills the passive role (figure : terminator >…> sold/bought object)

2. For the procedure Proc$_2$: `offer`

- *Peter* fills the active role (agent : initiator >…> donor)

- *daughter* fills the passive role (agent : terminator> … > beneficiary >…> receiver)

It is clear, then, that even in the declarative representation, *Peter,* the Subject of #6, is mapped twice onto the agent's active role.

Hereafter we will quote several examples of different mapping possibilities between CAs and semantic roles in English.

#7a. Mary treats Peter with aspirin.

 (Active verb + [Subject ‖ Active agent] + [Object ‖ Passive agent] + [Indirect Object ‖ Median figure])

#7b. Peter is treated with aspirin by Mary.

---

[13] cf. Wlodarczyk A., to be published.

(Passive verb + [Subject ‖ Passive agent] + [Indirect Object ‖ Median figure] + [Object ‖ Active agent]).

#8a. As for Mary, she treats Peter with aspirin.

(Active verb + [Topic ‖ Subject ‖ Active agent] + [Object ‖ Passive agent] + [Indirect Object ‖ Median figure])

#8b. As for Peter, he is treated with aspirin by Mary.

(Passive verb + [Topic ‖ Subject ‖ Passive agent] + [Object ‖ Active agent])

#9a. As for Mary, it is Peter whom she treats.

(Active verb + [Topic ‖ Subject ‖ Active agent] + [Focus ‖ Object ‖ Passive agent])

#9b. As for Mary, it is with aspirin that she treats Peter.

(Active verb + [Topic ‖ Subject ‖ Active agent] + [Focus ‖ Indirect Object ‖ Median figure] + [Object ‖ Passive agent])

#9c. As for Peter, it is Mary who treats him with aspirin.

(Active verb + [Topic ‖ Object ‖ Passive agent] + [Focus ‖ Subject ‖ Active agent] + [Indirect Object ‖ Median figure])

#10a. As for Peter, it is Mary by whom he is treated .

(Passive verb + [Topic ‖ Subject ‖ Passive agent] + [Focus ‖ Object ‖ Active agent])

#10b. As for Mary, it is Peter who is treated by her.

(Passive verb + [Topic ‖ Object ‖ Active agent] + [Focus ‖ Subject ‖ Passive agent])

Although most current syntactic theories separate subjects from semantic roles, many authors point out that — in Indo-European languages — the position of the Subject is often chosen for the most active (human or at least animate) participant of semantic situations[14]. As a matter of fact, the following universal rule may be established: the Subject is in a kind of manifest *solidarity* relationship with one of the Active, Median or Passive types of semantic roles, but this solidarity depends on the type of language (ergative or accusative languages). In accusative languages the Subject of direct utterances (i.e. in the basic diathesis) is prototypically *active* and in ergative languages, the Subject of direct utterances is prototypically *passive* (possibly *median*).

2.4. Explicit Subject

Subject is always explicitly expressed in languages in which verbs present a personal inflection, i.e. languages in which verbs mark the matching of the Subject[15] (global CA) of the utterance with the participants of the speech act (speaker, hearer or entity spoken about). In utterance 11, the Subject is neither the speaker nor the addressee

---

[14] Ikegami Y. [1981, 1991] points out the difference between Japanese and Indo-European languages as concerns what he calls "the degree of agentivity".

[15] In some languages verbs also have personal affixes matching not only the Subject but also Objects with the speech act participants, e.g. in Basque " there is agreement between the case-markers of the NPs and the personal affixes in these three cases [the absolute, the ergative and the dative case], Brettschneider G. "Typological Characteristics of Basque" in Plank 1979 p. 372.

but a person spoken about (*student*), therefore the verb is in the third person form.

#11a L'étudiant écoute.

#11b Student słucha.

#11c The student is listening.

In Martinet's approach (Martinet A. 1962,1967), Subject is the obligatory NP that cannot be deleted from the utterance without causing agrammaticality. In our view, this criterion is difficult to apply because of the frequent phenomenon of ellipsis that may affect any constituent of an utterance in a coherent text.

For this reason, we prefer to put forward the morphological factor. Subject is explicit when it is marked by a morpheme in the verbal ending, regardless of which semantic role it refers to, that is to say that Subject may be explicit not only in Indo-European accusative languages but also in ergative languages as soon as a language of this group exhibits verbal morphemes pointing at the participants of the speech-act.

However, in languages where the personal inflection of the verb is little distinctive as English or French, the subject must be made explicit by the use of the personal pronoun:

*#12a **I/ you/ we/ you/ they** come, **he** comes.*

*#12b **Je/tu** viens, **il** vient, **nous venons, vous venez, ils viennent.***

In languages in which verbs have personal morphemes, the presence of the Subject in the utterance may be marked only by the personal form of the verb. This is the case in many Indo-European languages. In Polish[16], where the personal morpheme of the verb is distinctive in all modes and tenses, the first and second person pronoun in the nominative form do not appear in the utterance unless the speaker intends to put an emphasis on them, i.e. treat them not as a Subject but either as topic or focus.

*#13 – Można sie było tego spodziewać. **Ty** (2nd person pronoun as Topic) zawsze na siebie bierzesz winę.* (Maria Nurowska, *Panny i wdowy*, Niezależna Oficyna Wydawnicza, Warszawa,1992, p.117)

(It was easy to foresee. You always take the fault on yourself.)

#14 – Coś mi tu nie pasuje w tym pokoju. Popatrz jak brzydko wiszą firany.

To nie ja (1st person pronoun as Focus) *je tak powiesiłam tylko Julia. (M.Musierowicz, Pulpecja, Signum, Kraków, 1993, p.133)*

(– There is something I don't like in this room. Look at the curtains, how they hang. – It isn't I who hung them but Julia.)

The same usage occurs in Italian where the personal pronoun is used only when it is necessary to put an emphasis on it.

#15 Pago … (I pay…)

*#16 **Io** pago … (It's me who pays …)*

---

[16] In this respect, Polish differs from Russian, where personal morphemes have disappeared in the past tense, thus obliging the speaker to use Subject personal pronouns. Compare Polish : *widziałem, widziałeś, widział* and Russian: *ja videl, ty videl, on videl* (I, you, he saw).

As a matter of fact, like Italian, many languages possess two different series of stressed and unstressed forms of personal pronouns in all cases of the declension. Even in English or French where declension has entirely disappeared in nouns, there still exist different forms for Subject or Object pronouns: *I/me, he/him* etc. Moreover, stressed forms sometimes differ from unstressed one, e.g. in French *moi/je, toi/tu* etc. Stressed forms are used to topicalise or focalise the pronoun. In languages like Italian or Polish1[st] and 2nd person unstressed nominative pronouns are simply omitted. The possibility to omit the Subject personal pronoun in languages with personal morphemes in verbs is considered as a universal device by Laskowski R. (Encyklopedia 1999 p. 414). This omission can be observed in typologically different languages and concerns not only Subject pronouns but also Object pronouns in languages with multiple (Subject and Object) personal morphemes in verbs (e.g. Basque).

In this connection, it is worth adding that, concerning personal inflection, from a diachronical perspective, modern personal morphemes of Indo-European verbs are former personal pronouns that were agglutinated to the verbal radical[17].

2.5. Implicit Subject

In Japanese, there is no agreement between verb and Subject because Japanese verb has no personal inflection and consequently, Subject may remain implicit, that is to say, when Subject is not expressed it must be inferred from the context or from the situation in which the speech act takes place. When only the Object is present in the utterance, it implies (as local CA) that the Subject (global CA) must be inferred from the context or situation.

#17 Sakana o tabemasu.

  [I/you/he/she/we/you/ they] eat fish.

#18 Nihon de wa sakana o tabemasu

In Japan, [people/we] eat fish

In the situation *to eat* (utterances #17, 18), *fish* as a Passive agent implies that human being(s)[18] are involved in the situation as Active agent. But it is possible to express the Subject explicitly (utterance #19).

#19 Nihon de wa <u>minna</u> sakana o tabemasu

In Japan, <u>everybody</u> eats fish.

In Japanese, the verb has no personal inflection, no number and gender categories, there is therefore no agreement between the verb and the subject as concerns these categories. However, the honorific morpheme of the verb is chosen in agreement with the subject[19], be it only implicit or explicitly expressed in the utterance.

---

[17] In a more recent time, in the history of Slavic languages, Polish reinforced the plural first person morpheme of the verb by adding *–y* to the *-m* inherited from Proto-slavic, e.g; Russian *idëm* (we go), pol. *idziemy* (we go). The morpheme *–my* in modern standard Polish is thus homonymous with the pronoun *my* (we).

[18] In Japanese the distinction between singular and plural is rarely marked morphologically.

[19] Wlodarczyk A. 1996 shows that verbal appreciative honorifics agree with the Subject (not with the semantic agent).

#20 (Suzuki-sensei ga) irasshaimashita.

(Prof. Suzuki) has arrived.

Hereafter is a short dialog in Japanese with its English translation: none of the following Japanese utterances has a subject and there is no personal form of the verb which would indicate the Subject, but this must be added in English translation (we put them in square brackets).

#21

A — Konnichi wa ! (Good morning/afternoon)

B — Yaa! (Hello!)

A — Doko ni iku no ? (Where are [you] going ?)

B — Hashi no shita ni sakano o tsuri ni iku n'da. ([I am] just going over there for fishing under the bridge.)

2.6. Implicit Subject in extended utterances

Japanese, as a language with implicit Subject, frequently uses utterances with a Topic. In the framework of the theory of Centres of Attention this can be explained by the fact that both Subject and Topic are global CAs, respectively in non-extended and extended utterances. When the global CA of the non-extended utterance (the Subject) remains implicit, the utterance is often extended and entails its own global CA (the Topic). Below are a few examples of Japanese utterances with a Topic which corresponds at the same time to the Subject or Object. In English, the explicit Subject has to be reintroduced at least in the form of an anaphoric pronoun (it appears in the following translations in square brackets).

#22 Sono sensô wa san-nen-kan tsuzuita.

(Lit. As concerns this war, [it] lasted three years. This war lasted three years.)

#23 Nihon wa kono nijû-nen no aida ni ookiku kawatta.

(Lit. As concerns Japan [it] has changed a lot during the last twenty years. Japan has changed a lot during the last twenty years.)

#24 Wakai mono wa samui to sake bakari nonde iru yo.

(Lit. As concerns those young people, when it is cold, [they] only drink alcohol.)

#25 Tanaka-san kara wa mô henji wo moraimashita.

(As concerns Mr Tanaka, [I] have already received his answer.)

#26 Kono hon wa kinô katta bakari desu.

(As concerns this book, [I] bought it only yesterday.)

Japanese utterances where the implicit Subject has the same referent as the explicit Topic can often be translated into English by a simple utterance with a Subject, (utterances #22-24), instead of the literal translation with a Topic. Otherwise, the Topic remains present in the English translation alongside the Subject (utterances #25-26). But in the final analysis, the choice of an utterance with or without a Topic in the translation from Japanese into a language with explicit Subject depends on the text coherence (on the place of a given utterance within the whole text), and this cannot be demonstrated on the basis of utterances taken out of context. Further on, translation with or without a

Topic depends also on stylistic conventions, which vary from language to language.

On the other hand, it is necessary to stress the fact that in Japanese, a widespread opinion among linguists notwithstanding, the particle *wa* is not a marker of Topic exclusively. Consequently, in Japanese discourse, topic-comment utterances are not as frequent as occurrences of *wa*. As a matter of fact, the particle *wa* is also used to mark the Subject of *given* meta-informative status matching a predicate of the same meta-informative status (cf. Wlodarczyk A. 2005). In utterance #27 below, both *sora (sky)* and *aoi (is blue)* are *given* by the situation-type and the whole situation spoken about has the *general* informative meaning.

> #27 *Sora wa aoi desu. (The sky is blue.)*

Similarly, it is worth emphasizing that the particle *ga* is neither an exclusive marker of Subject ; very often this particle is used to mark the Focus of an utterance[20]. Let us also add that in Japanese, Subject need not always be marked by a particle as for example *minna (everybody)* in utterance #19 above.

3. Diathesis and Centres of Attention

3.1. Semantic Relations and Logic

Semantic situations and their participants expressed in human languages are frequently modelled and represented as logical relations and their arguments. This is in fact the most common way linguists use predicate logic to model semantic relations, e.g. *Peter loves Mary* is often represented as *loves (Peter, Mary)*, and more generally a verb with two NPs is represented by p(x, y) or simply by x R y. But we must avoid transferring logical laws as they are used in logic into linguistic description. This danger arises because human languages express simultaneously in the same utterance elements from different levels (in our case, the levels of information and meta-information). Yet the above logical notation does not represent operations from two different levels, and we must strongly caution against following logicians when they distinguish between relations and their inverse: a R b and b R^{-1} a. We must be careful to identify accurately different levels of linguistic discourse in order to show that another kind of information – namely meta-information – accompanies pure semantic relations.

In reality, diverse ways of expressing semantic relations in human languages (in different sequential orders) do not always match inverse relations in logic because different word orders may reflect only different points of view that can be chosen by the speaker for presenting situations and their participants in an utterance. In other words, the difference between two linguistic expressions may have only a meta-informative and not informative (semantic) character. Human languages have several lexical or grammatical devices to present situations and their participants, depending on which participant the speaker intends to present as a global or local CA.

Languages possess different lexical verbs or verbal locutions making it possible to present the same situation from different points of view. For instance, the same real

---

[20] For a more detailed discussion of the problem of Focus in Japanese, see Wlodarczyk A. 2005.

world situation of 'selling-buying' may be expressed in an utterance by using either the verb *to buy* or *to sell*. The former enables the speaker to present the buyer as the global CA, the latter sets the seller in the position of the global CA. Some examples are given in semantic studies (Apresjan Ju. 1974), but they are mostly understood as "inverse relations". From our point of view, not all related lexical utterances that are treated in semantic studies as expressing inverse relations do actually correspond to inverse logical relations. Indeed, some reflect only a different point of view. Therefore, such utterances should be considered as pragmatic or meta-informative (and not semantic or informative) conversions. This problem deserves to be thoroughly analysed in a separate paper, but we shall concentrate here solely on grammatical devices.

The grammatical device used to present the same situation from different points of view (i.e. choosing this or that participant of the situation as global CA) is known as diathesis. Under diathesis we understand not only the opposition of active and passive utterance but also utterances without an explicit Subject called "impersonal utterances" in traditional grammars. As regards the active (or unmarked) voice in accusative languages, it is normally used to introduce the active agent of a situation as the Subject of the utterance. When the text coherence and the meta-informative presentation of *given* and *new* information make it necessary to put the passive agent in the position of the Subject, accusative languages turn to verbs in the passive voice. On the contrary, ergative languages use a non-marked voice to present as Subject the passive participant and use what is generally regarded as the "anti-passive" voice in order to build utterances in which the Subject is the active agent of the situation.

Thus, in the meta-informative centering theory, diathesis is not mainly a problem of syntax and semantics, but it is most of all a device belonging to the pragmatic level allowing speakers to establish a mapping relation between the global CA of the utterance and the active or passive agent of the semantic situation.

3.2. Passive Utterances

In accusative languages, passive utterances allow us to consider the passive agent as global CA of an utterance, thus putting the active agent in the position of an indirect Object *(by whom)* or leaving it unmentioned.

#28a Active voice *His father offered John a book for his birthday.*

#29a(1) *A book was offered to John for his birthday (by his father).*

#29a(2)*John was offered a book for his birthday (by his father).*

In English, when a verb has two Objects in the active voice (e.g. *#28a*), the Subject of the passive voice may correspond both to the direct (cf. #29a(1)) and to the indirect Object (cf. #29a(2)). In other languages, as French and Polish, only the direct Object of an active utterance may become the Subject of the passive utterance.

#28b Fr. *Son père a offert un livre à Jean pour son anniversaire.*

#29b Fr. *Le livre a été offert à Jean pour son anniversaire (par son père).*

#28c Pol. *Ojciec podarował Jankowi książkę na urodziny.*

#29c Pol. *Książka została podarowana Jankowi na urodziny (przez ojca).*

12

In Polish, moreover, the direct object may have the form of different cases (not only that of the accusative), e.g.: genitive or instrumental.

#30 Pol. Rodzice wymagali posłuszeństwa od dzieci.

(Lit. Parents demanded obedience from children. The parents demanded obedience of their children.)

#31 Pol. Posłuszeństwo było wymagane od dzieci (przez rodziców).

(Lit. Obedience was demanded from children by parents. Obedience of their children was demanded by their parents.)

As we shall see below, the active agent of the "impersonal" utterance is not really unmentioned, but is expressed as a generic (unspecified) member either of the class of human or non-human entities.

In current linguistic theories the difficulty in explaining the different diatheses (which make it possible to linearise semantic situations in various ways) is due to the fact that diathesis is defined only as expressing the mapping between syntactic positions and semantic roles, and does not include the pragmatic meta-informative level of utterances. As a matter of fact, without taking into account the pragmatic level, passive utterances are merely arbitrary transformations of active ones. In traditional grammars, morphological and syntactic devices of the passive voice were described thoroughly, but such grammars could not explain how and why speakers use passive utterances.

3.3. Anonymous Subject

It is a well known fact that languages with explicit Subject may use utterances which are called *impersonal*. As one of the co-authors was working on a formal model of impersonal sentences in Polish and Russian (Wlodarczyk H. 1994) she was led to put forward the hypothesis that such utterances in fact contain a hidden Subject in order to explain different types of impersonal sentences that were distinguished by native speakers. We proposed to call this hidden Subject *anonymous* because it is in a way comparable to an anonymous variable in computer science[21]. In languages whose verbal inflection displays distinctive personal morphemes (e.g. Slavonic languages, but also some Romance languages), those morphemes are the only mark of the presence of this Anonymous Subject (from now on AS) in the utterance.

#32a Pol. Ø pada.

#32b Lat. Ø pluit

In languages (like French or English) whose verbs do not have distinctive personal morphemes, the presence of an AS is marked by an impersonal pronoun. The pronoun referring to an AS is sometimes called "dummy subject" in syntactic theories: this term

---

[21] For more arguments in favour of the hypothesis of Anonymous Subject, especially in Russian and Polish, see Włodarczyk H. 1994, 1996. The hypothesis of AS was elaborated and tested in a computer program written in Prolog. This program contains rules generating and analysing different kinds of « impersonal » utterances in Polish and Russian, these rules being homologous to those generating utterances with a Subject (oral communication at the University of Warsaw, prof. Swidzinski's seminar, may 2002).

captures the fact that something must fill in the place of the Subject left open by a verb in languages with personal inflection[22].

> *#32c Fr. Il pleut.*

> *#32d It is raining.*

Although German verbs have distinctive personal morphemes, a neutral pronoun must be used to mark the AS.

> *#32e Ger. Es regnet.*

The hypothesis that in languages with explicit Subject, *impersonal* utterances entail an AS is all the more corroborated by the existence of two distinct semantic types of AS: one refers semantically to the set of non-human entities (-hum), the other to human beings (+hum)[23]. This distinction can be observed in many European languages: English, French, German, Russian, Polish etc. As a matter of fact, alongside the -hum AS (like those in #32a-e above) that are used to refer to non-human entities as, for instance, atmospheric phenomena, there exists another type of AS identifying human beings. Furthermore, within the category of human AS, languages are endowed with different devices enabling speakers to distinguish between a human AS considered as the generic human being (utterances #33a-b) and a slightly more specified human being belonging to a socially identified group (a company, the government, etc. as in #36 below). In the former case the speaker may identify himself[24] as belonging to this generic human person and frequently he uses a pragmatic device, i.e. the second person verbal form and pronoun, in order to involve not only himself but also the addressee (cf. #34a-b).

> *#33a One never knows what can happen.*

> *#34a You never know what can happen.*

> *#33b Pol. Nigdy się nie wie, co może się zdażyć.*

> *#34b Pol. Nigdy nie wiesz, co może się zdażyć.*

On the other hand, the speaker may insist on the fact that he himself does not belong to the mentioned group (cf. Polish utterances #35-36 below).

> *#35 Pol. Zatrzymano go na drodze.*

> *(Lit. [someone] arrested him on the road . 'He was arrested on the road'.)*

> *#36 Zatrzymali go na drodze.*

> *(Lit [They] arrested him on the road.)*

The utterance with a +hum AS in some languages (e.g. Polish #35) can often be

---

[22] I. Mel'cuk 1988 puts forward the concept of "zero subject".

[23] As concerns Polish, cf. Laskowski R. : „Podczas gdy konstrukcje nieosobowe na *się i -no / -to : chodzi się, wykonano, milczano, bito,* implikują istnienie osobowego wykonawcy czynności lub nosiciela stanu czy procesu, chociaż go bliżej nie określają, konstrukcje nieosobowe typu *zasypało (drogę)* wykluczają osobowego sprawcę czynności (agensa)." (Laskowski, Gramatyka T.2, 1984: 147) in English: "Whereas such [Polish] impersonal utterances with forms *się and -no / -to (e.g. chodzi się, wykonano, milczano, bito)* implicate the existence of a personal agent of the action (or that of a person involved in a state or process) without determining this person closer, other impersonal utterances of the type *zasypało (drogę)* exclude a personal agent of the action (an *agens*)".

[24] It is perhaps worth mentioning that in English the masculine *he* is the unmarked gender form for the generic human being, and it is therefore not necessary to use the unpronounceable form *s/he.*

rendered in English or French by a passive utterance in which the participant playing the active role remains implicit. Moreover, as was already pointed out for Polish, languages have very often several different expressions referring either to a generic human being or to a more limited group of persons: the choice between them is determined by fine semantic nuances, pragmatic conditions or style[25].

Only when taking into account the semantic distinction between +hum and –hum AS, can we understand the communicative purposes for which impersonal utterances are used. Such utterances allow us to identify the global CA very generally, merely as a participant (of the situation) belonging to a class of entities (+hum or –hum) without selecting a particular element of one of this classes. As is the case with an explicit Subject, the semantic role of the participant referred to as AS has to be recognised by the hearer when interpreting the utterance (Wlodarczyk H. 1996). This role is not explicitly expressed in the utterance.

In languages with explicit Subject, AS is a linguistic sign with both a form (morphonological zero, or invariant form as *man, es* in German, *on, il* in French) and a content (it refers to a generic either human or non human entity). But ASs differ from indefinite pronouns (*someone, something* etc.) in that they can occur only in the position of the Subject of an utterance. The presence of an AS in an utterance (even when it has a zero morphonological form) is manifested by agreement constraints on the verb. For instance, when automatically generating impersonal utterances in Russian or Polish we must take into account these constraints in order to get acceptable utterances (Wlodarczyk H. 1994).

Among other examples, the concept of AS makes it possible to explain French utterances in which the active participant is demoted to the position of the Object (local CA) whereas the position of the Subject is occupied by an AS. Such utterances enable the speaker to introduce an active participant as a local CA (i.e. as an Object !) even of an *intransitive* verb.

#37 Fr. Il vient des étrangers.

(Lit. There comes foreigners. Foreigners are coming.)

#38 Il arrive de drôles de choses.

(Lit. There happens strange things. Strange things are happening.)

As we shall see below, due to the frequent topicalisation of the Object of impersonal utterances in languages with explicit Subject, such topicalised Object is sometimes interpreted as a Subject in an oblique case or as an explicit Object with *subject-like* or *subjectal* properties. In our view, we treat as Subjects only those NPs which are in the unmarked case form (nominative[26] or absolutive).

To return to the problem of AS, let us stress that there is in fact no significant

---

[25] About Polish, Russian and French, see Wlodarczyk H. 1996. We distinguish several types of +hum and –hum SA in Russian and Polish, for a detailed discussion see Wlodarczyk H. 1994.

[26] The only exception we admit for Russian and Polish concerns genitive NPs (with a partitive meaning), in fact this genitive is an 'adnumeral' form of the nominative or accusative cases with numerals or with the negation. For details see Wlodarczyk 1992, 1993, an 1996. The concept of 'adnumeral' form is introduced by Menantaud H. 1999.

difference between ASs that have a zero morphonological form and those which have the form of an indefinite pronoun appearing only in the nominative case. On the contrary, the semantic distinction between human and non-human anonymous subjects is cross-linguistically attested regardless of their form (zero or not). Necessarily, however, impersonal utterances with AS occur only in languages where Subject is explicit. In such languages, utterances with AS must not be confused with utterances in which, due to text coherence, the Subject undergoes ellipsis (utterance #39b).

#39a – What is Bill doing ?

#39b – Watching TV.

It is worth adding that in languages with implicit Subject (as Japanese) the concept of Anonymous Subject is simply absurd.

4. Isomorphism between the Subject and the Topic

At this stage, it is important to show the isomorphism between centres of attention of simple and extended utterances, as it may create problems of interpretation, in particular the confusion between Subject and Topic[27].

This isomorphism concerns on the one hand Subject as global CA of a simple utterance and Topic as global CA of an extended utterance[28], and on the other hand, both Object and Focus as local CAs, but respectively in simple and extended utterances.

It is sometimes all the more difficult to interpret an NP as Subject or Topic as there exists in human languages weak Topics that are very close to Subjects, those two types of CAs being very often distinguished only by intonation (stress and pause), which may be a matter of interpretation as regards written language[29] where suprasegmental phonology remains unspecified.

However, when markers of topicalisation appear in the utterance as special morphemes, the confusion between Subject and Topic is thus impossible, as in the following examples from different languages   but which all have the same meaning.

#40a Eng As concerns Mark's father, he is a doctor.

#40b Fr. Le père de Marc, il est médecin.

#40c Pol. Ojciec Marka to jest lekarzem.

#40 d Jap. Taro no otosan wa isha desu.

4.1. Subject of active utterances with topicalised Object

The isomorphism between Subject and Topic often leads linguists to interpret a

---

[27] In some languages, it is difficult to decide whether a NP should be interpreted as Subject or Topic. According to Palmer R. F. (1994;p. 201 sq), this concerns "most of the languages of the Philippines", especially in Tagalog where the particle *ang* may "topicalise" an NP playing the role of agent, patient, beneficiary or location. We claim that the concept of CA and the distinction between extended and simple utterances could be useful for explaining these problems in such languages.

[28] The close relation between Subject and Topic has been highlighted by many scholars. See Hagège C.1978, 1981, Martin R. 1985, Lambrecht K. 1974 etc. .

[29] For more details on *weak* and *strong* Topics, cf. Wlodarczyk H. 1999.

topicalised Object as a Subject, especially in Indo-European languages when a sentence has no explicit Subject. This concerns the impersonal utterances when the presence of an AS is not taken into account: some linguists have argued that when there is no explicit Subject in the nominative form another NP in an oblique case may be considered as the Subject or "subject-like" NP[30].

Compare utterances #41 (with a Subject) and #42 (with an AS and a topicalised object hereafter).

#41 Pol. Snieg zasypał drogi. (Lit. Snow covered the roads.)

#42 Pol. Śniegiem zasypało drogi. (Lit. By snow 'it' covered the roads. The roads have been covered by snow.)

In #42 *śniegiem* is not a subject but a topicalised object, and the Subject has the form zero of a non-human AS (rendered by 'it' in English)[31]. Notice that it is also possible to topicalise the direct object *drogi* (roads) which, semantically, plays a passive role and therefore is not considered to be a « subjectlike » NP.

#43 Pol. Drogi zasypało śniegiem. (Lit. The roads 'it' covered by snow. The roads have been covered by snow.)

Likewise, in Japanese, in an utterance without an explicit subject the distinction must be made between a topicalised Subject (ex. #44) and a topicalised (direct or indirect) Object (#45). In #45 the topicalised NP *Watakushi ni wa* contains the 'dative' particle *ni* and must not be considered a topicalised Subject although it refers to the same person (the speaker) as the topicalised NP in ex. #44.

#44 Jap. Watakushi wa wakaranai. (I do not understand.)

#45 Jap. Watakushi ni wa wakaranai. (Lit. To me it is not understandable. It is not comprehensible to me.)

4.2. Subject of passive utterances with topicalised Object

In accusative languages, both the Subject of a passive utterance and the topicalised Object of an active utterance (with an explicit Subject or with an AS) refer to the passive role of the semantic situation expressed by the verb. The difference lies in that the topicalised Object is the global CA of an extended utterance, whereas the Subject is the global CA of a simple utterance. The choice of one or the other utterance depends on the text coherence.

In French, this problem is mentioned by Gaatone (1998 p. 239), who quoted the following utterance #46a, in which the Object ("cet effort") is topicalised.

#46a Fr. L'effort pour coordonner les données tactiles et visuelles, reste l'un des plus prodigieux de ma vie […]. Cet effort, chaque enfant en bas âge l'accomplit. (The effort to coordinate tactile and visual data remains one of the most tremendous efforts in my life […] Such an effort, each infant has to accomplish it.)

---

[30] Farrell P. 2005 uses terms as "quasi-subjects" and "oblique subjects" p.98 sq.

[31] Kazimierz Polański (2003) defends the same point of view, considering that the use of NPs in oblique cases appearing in Polish impersonal utterances should not be treated as Subjects: "…traktowanie form *Janka [mdli], Jankowi [smutno][…]* jako podmioty nie wydaje się uzasadnione." (treating forms like *Janka [mdli], Jankowi [smutno][…]* as subjects seems unjustified .)

In #46a the Object NP "cet effort" is topicalised but Gaatone stressed that it could just as well be the Subject of a passive utterance (as in #46b).

#46b Fr. Cet effort est accompli par chaque enfant en bas âge. (Lit. This effort is accomplished by each infant.)

Both utterances #46a-b fit in the text coherence, but the choice between them makes it possible to put more or less emphasis on the NP "this effort". The topicalised NP is emphasized more than the Subject of the passive, which is probably the reason why it was used by the author of the original version of this utterance.

In Polish, as in other languages often characterised as "free" word order language[32] because of the declension marks in nouns, the topicalisation of Object is very frequent, although the passive voice may also be used to treat a NP referring to a passive role as Subject. Compare utterances #47a-c.

#47a Pol. Książka została podarowana Jankowi na urodziny (przez ojca). (The book was offered to Janek for his birthday (by his father).)

#47b Pol. Książkę to[33] ojciec podarował Jankowi na urodziny. (Lit. The book father offered to Janek for his birthday.)

#47c Pol. Książkę podarowano Jankowi na urodziny. (Lit. The book [someone] offered to Janek for his birthday.)

The utterance #47c with AS is very close to the passive utterance #47a in that it makes it possible not to mention at all in the utterance the NP referring to the active role (e.g. *ojciec : father*).

Stylistically, the topicalisation of an Object is frequent in written Polish because it does not require the presence of an anaphoric pronoun in the comment part of the utterance as it does in French, a language in which the topicalisation of Objects is frequent in spoken but not in written language[34].

Moreover, in many languages, it is possible to topicalise the Subject of a passive utterance, e.g. in Polish #47d.

#47d Pol. Książka to została podarowana Janowi na urodziny (przez ojca). (Lit. As concerns the book, it was offered to Janek for his birthday (by his father).)

In French the topicalisation of the Subject of a passive utterance is characteristic of the spoken language (as is the topicalisation of Subject or Object NPs in active utterances).

#48 Fr. Le livre, il a été offert à Janek pour son anniversaire. (Lit. The book, it was offered to Janek for his birthday.)

Such a topicalisation is often used to express a contrast between the topicalised NP

---

[32] Polish scholars insist that in Polish (and other Slavonic languages) word order is "free" but not "arbitrary": "Układ wyrazów w zdaniu w języku polskim jest *swobodny,* jakkolwiek niezupełnie *dowolny*." ("The word order in sentences is *free*, however it is not completely *arbitrary*) (Szober St. 1967). As for us, we claim that each different order has a different meta-informative interpretation.

[33] In Polish, the use of the particle *to* after a topicalised NP is characteristic of the spoken language, but it is not obligatory and is not used in written literary style.

[34] As concerns French/Polish contrast see Wlodarczyk H. 2004.

and another NP previously mentioned in the text, this could be the case of "the car" in utterances #49a-b, contrasting with "the book" in utterances #47a-c and #48.

> #49a Fr. La voiture, Janek l'a eue pour Noël. (Lit. The car Janek got it for Christmas.)
>
> #49b Pol. Samochód Janek dostał na gwiazdkę.

Translators must be aware of the different stylistic properties of Subject and Topic, if they want to reflect not only the semantic content of the original text but also its pragmatic force.

5. Subject in ergative languages

The definition of Subject as global CA makes it possible to give a new explanation of verb valence in ergative languages. Obviously, ergativity is a vast problem and extensive research has already been done about it[35]. We will propose, however, a general framework in which such a crucial concept as Subject can be defined as CA on a level different from that of agentivity. The problem of Subject in ergative languages seems to derive from the projection of the solidarity relationship (*solidarity* for short) observed in different types of languages between Subject and either an active, passive or median role. Most descriptions of ergative languages make use of *semantico-syntactic* concepts named respectively S (standing for the Subject of intransitive sentences), A (Subject of transitive sentences) and O (Object of transitive sentences). As an alternate solution, we claim that in order to explain the diversity of these languages, we must first make a clear distinction between *information* (semantics) and *syntax* (which can be seen as the interface between form and content).

First, let us recall briefly what is commonly accepted about Subject in both accusative and ergative languages. It is currently stated that in most ergative languages : (1) the noun-phrase in the Absolutive Case is obligatory (called also « first participant », cf. Tchekhoff Claude (1978, p. 58-66)), (2) the Absolutive Case has no marker, (3) the Absolutive Case indicates (a) the Patient of transitive sentences and (b) the Subject of intransitive sentences. As far as accusative languages are concerned, most linguists consider that (1) the Nominative Case has no marker, (2) it corresponds either to (a) the Agent of transitive utterances or (b) the Subject of intransitive utterances ; (3) in most cases, in the active voice, the Subject corresponds to the Agent.

When we use the meta-informative theory in order to explain the phenomenon of ergativity, we get[36]:

(1) Assuming that the prototypical diathesis of accusative languages is the active one, we claim that the prototypical diathesis of ergative languages is the passive one. Our

---

[35] A few books and papers on ergativity are quoted in this paper. Let us stress here the specificity of Japanese, which is not an ergative language but differs from Indo-European accusative languages in that it has an implicit Subject. For typological characterisation of Japanese both in contrast with accusative and ergative languages cf. Tsunoda T. 1991.

[36] Below we quote examples from the Tonga language (Polynesia) following Tchekhoff Claude (1978). ABS stands for absolutive case, ERG for ergative.

claim is based on the fact established above that in accusative languages (a) there is a solidarity relation between Subject and the Agent which plays an *active* semantic role and that (b) the prototypical diathesis is *active*. We propose the reverse perspective for ergative languages: (a) there is a solidarity relation between Subject and the Agent which plays a *passive* semantic role and that (b) the prototypical diathesis is *passive*.

(2) In *intransitive* utterances the Subject is expressed by the Absolutive Case

Here we shall use the following notations: *CA :: Case* is to be read as "the CA marked by the Case." and *A ‖ B* is to be read as "A is the same as B".

#50 Na'e alu 'a Tevita ki Fisi.

   *PAST go SBJ :: **ABS**-David to Fiji.*

(David went to Fiji).

(3) In *transitive* utterances :

(3.1) In presence of the Ergative phrase referring to the Active Role, the Subject (expressed by the Absolutive Case) points to the Passive Role.

#51 Na'e tamate'i 'a Kolaitate 'e Tevita

*PAST kill-PERF SBJ :: **ABS**-Goliath OBJ :: ERG-David.*

(*Subject ‖* Passive role "Goliath", kill(*Object ‖* Active role = "David")

(David killed Goliath).

#52 'Oku ui 'a e pepe 'e he fa'é.

*PRESENT call SBJ :: **ABS** DEF baby OBJ :: **ERG** DEF mother*

(*Subject ‖* Passive role "baby", call(*Object ‖* Active role = "mother")

(The mother calls the baby.)

(2.2) In the absence of the Ergative phrase indicating the Active semantic Role, the role expressed by the Subject (marked by the Absolutive Case) is ambiguous : it can be either the Active or the Passive role, e.g. in Tonga, in transitive utterances with a single explicit CA, the Absolutive case may correspond not only to the Active but also to the Passive role, hence the following ambiguity arises:

#53 'oku ui 'a e fa'e.

*PRESENT call SBJ :: **ABS** DEF mother.*

[1] *(Subject ‖ Passive role "mother", call (Object ‖ Active role = Ø)*

([someone] calls the mother or *the mother is called)*

[2] *(Subject ‖ Passive role = Ø, call (Object ‖ Active role ="mother")*

(The mother calls [someone] or *someone is called by the mother.)*

In ergative languages, in sentences containing two CAs (a Subject and an Object) there is no ambiguity. In sentences like #53 above containing only one CA, however, a certain ambiguity appears, i.e. as the Absolutive case is bound to the global centre of attention (Subject), the first meta-informative interpretation that spontaneously comes to the listener's mind is (Subject, (Pred(Verb, Ø *Object*))) and the second is (Ø Subject, (Pred (Verb, Object))).

Because the Subject (marked by the absolutive case) is obligatory in Tonga,

utterances with only one ergative NP (example given #54) are reported to be unacceptable in that language. This fact seems to confirm that there is a solidarity relation between the Subject and the passive role in Tonga.

#54 *'oku ui 'e he fa'e.

*PRESENT call   **ERG** DEF   mother.*

*Ø call the mother.

After defining *pragmatic centres of attention* and *semantic situation roles* as belonging to different levels in our multi-layered model of predication, we claim that the main difference between accusative and ergative languages can be represented as in Table 6 below:

Accusative languages	Ergative languages
(1) CA *Subject* ‖ Active agent	(1a) CA *Subject* ‖ Passive agent
(2) CA *Object* ‖ Passive agent	(2) CA *Object* ‖ Active agent

Table 3: Prototypical solidarity between CAs and semantic roles in accusative and ergative languages

In our terms, in ergative languages the global CA is in the absolute case and we call it the Subject. Let us mention that several authors who dealt with ergativity problems, Nicolai Trubetzkoy and Charles Fillmore, to mention the most famous of them, also consider that the NP in the absolutive case corresponds to the Subject of the verb[37]. More recently some linguists also support this point of view, although they prefer not to use the term Subject but only speak of "subjectal" properties of the NP in the absolutive case: "In syntactically ergative languages, the properties of absolutive NPs are the properties of grammatical subjects." (Manning 1996 p. 41). According to our framework, Subject in ergative languages is characterised by a solidarity with the passive roles (usually called Patient in most syntactico-semantic theories). However, in certain ergative languages, this solidarity concerns even the Subject and the Median agent role. In split-intransitive systems (cf. Dixon 1994) the Subject may correspond to the Active agent with some semantic classes of verbs and to the Passive agent with others.

But it should be stressed that the relation between Subject and semantic roles is only a matter of *solidarity* : neither does Subject in accusative languages always refer to the Active roles nor does it always refer to the Passive roles in ergative languages. Moreover, probably in all languages, the function of diathesis is to make it possible to change this mapping. Thus, the notion of Subject as global CA is common to all languages and its use is not bound to any one semantic role. The choice of the participant, which is frequently treated as Subject (main Centre of Attention), is merely a matter of point of view, and is imposed on human languages by the sequential nature of morphonological units. The most probable choice depends on the language type[38].

---

[37] It is quite logical thus that Charles Fillmore (1968) uses the term "nominative" instead of "absolutive".

[38] The preferred solidarity of Subject with one type of role (be it active, passive or median) should

6. Conclusion

Only when defining Subject on the level of meta-information (pragmatics) as the global centre of attention of the utterance is it possible to give a universal definition valid for all languages. The identification of a participant that is treated as global CA is the first stage (*signification*) in the communication process but the *interpretation* of the Subject as a participant playing a semantic role is a more complex operation that takes place at a second stage. As a matter of fact, this interpretation becomes even more complex when we introduce metaphor and metonymy, as they enable some figures to play agent-like roles.

Definitions of Subject which take into account only the level of syntax are inadequate for all types of languages: for instance, the definition of Subject as obligatory alongside the predicate for the grammaticality of the utterance (Martinet 1962,1967) or the view of the Subject as non constitutive of the utterance but accommodating the predicate (Kuryłowicz 1949, 1964) are not suitable for such languages as Japanese where Subject is not obligatory because verbs have no personal conjugation and consequently have no morphological agreement with the Subject. On the other hand, the meta-informative theory of Subject and Object as CAs of simple utterances makes it possible to explain ergative languages without binding Object or Subject to one semantic role only.

Thus, the proposed definition of Subject as global CA of an utterance and the solidarity between CAs and semantic roles enables us to distinguish more appropriately between different types of languages:

(1) *accusative* languages in which (explicit or implicit) Subject most frequently corresponds to the Agent playing the Active Role in the semantic situation expressed by the protypically active diathesis,

(2) *ergative* languages where Subject most often corresponds to the Agent playing the Passive Role expressed by the protypically passive diathesis,

(3) *ergative* languages in which Subject matches most frequently the Median agent role and

(4) *ergative* languages with *split-intransitive Subject* (cf. Dixon 1994) discriminate between two sorts of Subjects: one whose morphology is like that of Subjects (prototypically matching the Passive Agent Roles) of transitive utterances and another one whose morphology is like that of Objects (prototypically matching the Active Agent Roles) of transitive utterances.

Moreover, as concerns the description of one language, the meta-informative definition of Subject as CA explains different diatheses (including active, passive, "anti-passive" and "impersonal") as offering the possibility to change the global CA of an utterance depending on the point of view on participants that is chosen by the speaker who seeks to preserve the coherence of the text or discourse.

---

not be construed as an indication of the degree of evolution of a given language. There is no reason to believe that the solidarity between Subject and Active role that we find in accusative languages should be used as argument to suggest different levels of cultural development.

7. References

APRESJAN Ju. D., 1980, Semantyka leksykalna. Synonimiczne środki języka (Leksićeskaja semantyka. Sinonimiceskije sredstva jazyka, Moskva 1974), Ossolineum, Wrocław.

BARWISE, J, and PERRY, J,1983 *Situations and attitudes* (fourth printing 1986), The MIT Press.

DIXON, R. M .W., 1994, *Ergativity,* Cambridge Studies in Linguistics, Cambridge University Press.

DOWTY, D.R., 1991, "Thematic proto-roles and argument selection", Language 67: 547-619.

Encyklopedia językoznawstwa ogólnego, wyd. drugie (1998) pod red. K. POLAŃSKIEGO, Ossolineum, Wrocław – Warszawa – Kraków.

FARRELL, P., 2005, *Grammatical Relations,* Oxford University press.

FILLMORE, Ch. J., 1968, "The Case for Case", in Emmon Bach and Robert T. Harms eds., *Universals in Linguistic Theory,* New York, Holt, Rinehart, Winston.

FILLMORE, Ch. J., BAKER, C. F. and SATO, H., 2002, The FrameNet Database and Software Tools. In Proceedings of the Third International Conference on Language Resources and Evaluation (LREC). Las Palmas. 1157-1160.

GAATONE, D., 1998, *Le Passif en français,* Champs linguistiques, Duculot, Paris, Bruxelles.

GUTIÉRREZ ORDÓNEZ S., 1997, *Temas, remas, focos, tópicos y commentarios,* Arco / Libros, Madrid.

 — 2002, *Forma y sentido en sintaxis,* Arco / Libros, Madrid.

— 2006, "Thématisation, Focalisation, Topicalisation" in Wlodarczyk A. & Wlodarczyk H; *La Focalisation dans les langues,* Coll. Sémantiques, L'Harmattan, Paris.

IKEGAMI, Y., 1981, *Suru to naru gengogaku* (Linguistics of doing and becoming), Taishukan, Tokyo.

— 1991 " 'Do-languages' and 'Become-language': Two Contrasting Types of Linguistic Representation", in *The Empire of Signs, Semiotic essays on japanese Culture,* John Banjamins, Amsterdam –Philadelphia

HAGÈGE C., 1978, « Du thème au thème en passant par le sujet : pour une théorie cyclique », *La linguistique,* 14, 2, 1978, pp.3-38.

— 1981, Bulletin de la Société de Linguistique de Paris, tome LXXVI, fasc. 1, pp. iv-v.

JESPERSEN, O, 1924, *The Philosophy of Grammar,* George Allen & Unwin Ltd, London..

KAROLAK S. et al., 1984, *Gramatyka współczesnego języka polskiego*, T.1 Składnia, P.W.N., Warszawa

KURYŁOWICZ J., 1949, "Le Problème du classement des cas", *Biuletyn Polskiego Towarzystwa Językoznawczego* 9.

1964, The Inflectional Categories of Indo-European, Heidelberg.

LAMBRECHT K.., 1994, *Information Structure and Sentence Form. topic, focus, and the mental representations of the discourse referents,* Cambridge University Press, Cambridge, p. 388.

LASKOWSKI R. et al., 1984, *Gramatyka współczesnego języka polskiego,* T.2 Morfologia P.W.N., Warszawa

L'HERMITTE R., 1979, "Un type de phrase impersonnelle en slave : les participes passifs à valeur prédicative", *Revue des Etudes Slaves* T. LI/21-2 (VIIIe Congrès des slavistes à Zagreb 1978), Paris.

MANNING, Ch. D., 1996, *Ergativity, Argument Structure and Grammatical relations,* CSLI Publications, Stanford.

MARTIN R., 1983, *Pour une logique du sens*, PUF

MARTINET, A., 1962, *A functional View of Language*, Clarendon, Oxford (1969, trad. fr. *Langue et fonction*, Denoël, Paris).

1967, *Eléments de linguistique générale,* Armand Colin, Paris.

MEL'CUK, I. A., 1988, *Dependency syntax: theory and practice,* State University of New York Press.

MENANTAUD H. "La négation comme catégorie grammaticale en polonais et en lituanien", CdLIN, 1999/1, 43-57.

PALMER F. R., 1994 (reprinted 1998), *Grammatical Roles and relations,* Cambridge University Press.

POLAŃSKI, K., 2003, „Na marginesie dyskusji o podmiocie", *Études linguistiques romano-slaves offertes à Stanisław Karolak*, Oficyna Wydawnicza Edukacja, Cracovie.

PLANK, Frans, ed., 1979, *Ergativity, Towards a Theory of Grammatical Relations,* Academic Press, London, New York, Toronto, Sydney, San Francisco.

SALONI Z., ŚWIDZIŃSKI M. (1985) *Składnia współczesnego języka polskiego*, Wyd. drugie zmienione, PWN, Warszawa.

SHATUNOVSKIY I., 2006 (to be published), "Two types of new – given and ways of introducing new objects into the world and discourse", to be published in Wlodarczyk A.& Wlodarczyk H; *La Focalisation dans les langues,* Coll. Sémantiques, L'Harmattan, Paris.

SZOBER S., 1967 (8th ed.), *Gramatyka języka polskiego,* Warszawa.

TCHEKHOFF C., 1978, *Aux Fondements de la syntaxe : l'ergatif ,* PUF, Paris, p. 202.

TESNIERE Lucien, 1959, *Eléments de syntaxe structurale,* Klincksieck, Paris.

TSUNODA T. 1991, *Sekai no gengo to Nihongo* (The Languages of the World and the Japanese), Kuroshio Pub. Co., Tokyo.

WIERZBICKA, A., 1986, "The Meaning of a Case: A Study of the Polish Dative", in Brecht and Levine, *Case in Slavic,* 1986, 386-426.

WLODARCZYK, A., 1996, *Politesse et personne, Le japonais face aux langues occidentales,* L'Harmattan, coll. Sémantiques, Paris.

— 2003a, "Les Homotopies du topique et du focus", *Ordre et distinction dans la langue et le discours*, Actes du colloque international de Metz (18,18 et 20 mars 1999, publiés par Bernard Combettes, Catherine Schnedecker, Anne Theissen, Paris, Honoré Champion éditeur, p.513-526.

— 2003b, "Linguistique dynamique: évolution du discours dans le temps", *Etudes linguistiques romano-slaves offertes à Stanisław Karolak*, Oficyna Wydawnicza Edukacja, Cracovie 2003, p. 497-510.

— 2003c, « Les cadres des situations sémantiques », *Etudes cognitives / Studia*

kognitywne V, SOW, Warszawa.

— 2004, "Centres d'intérêt et ordres communicatifs" in *Enoncer, l'ordre informatif dans les langues,* réd. par Cotte P., Dalmas M. et Wlodarczyk H., L'Harmattan, Paris.

— 2005, "From Japanese to General Linguistics, Starting with the *wa* and *ga* Particles", in *Paris Lectures on Japanese Linguistics,* invited for the New Millenium and edited by Wlodarczyk A., Kurosio Shuppan, Tokyo, p. 27-61.

— to be published, "Les rôles des situations sémantiques", *International Conference "Cognitive Studies, Semantics and Contrastive Linguistics,* Polish Academy of Sciences, Institute off Slavistics, Warsaw, April 4-5th 2005.

WLODARCZYK, A., & WLODARCZYK, H., 1998, "Graded Informative Content of Linguistic Messages", CD-ROM: *Proceedings of the 16th International Congress of Linguists.* Pergamon, Oxford, Paper N° 0431

— 2006, "Focus in the Meta-informative Theory of Utterances", in *La Focalisation dans les langues,* Wlodarczyk A. & H. ed., Coll. Sémantiques, L'Harmattan, Paris.

WŁODARCZYK, H., 1992, « Sujet syntaxique, formes casuelles et traits sémantiques en polonais contemporain », *Mélanges en l'honneur du Professeur Paul Garde,* Université d'Aix-en-Provence

— 1993, "Sentences without a Nominative NP in Polish", *Studies in Polish Inflectional Morphology and Syntax - Synchronic and Diachronic Problems ,* ed. R. Laskowski and G. Hentschel, Verlag Otto Sagner, München.

— 1994, « Les Phrases à sujet anonyme et verbe réfléchi en polonais et en russe », *Organisation sémantiue de la phrase dans les langues slaves,* Revue des Etudes Slaves T. LXVI fasc. 3, p. 669-680, Paris.

— 1996, « Les Traits sémantiques du sujet « anonyme »,en polonais, russe et français », *Semantyka a konfrontacja językowa,* Akademia Nauk, Instytut Slawistyki, Slawistyczny Ośrodek Wydawniczy, pp. 179-198, Warszawa.

— 1997, „Informacyjna jakość wypowiedzenia w przekładzie literackim", *Język polski w kraju i za granicą,* Materiały Międzynarodowej Konferencji Polonistów, Warszawa, 14-16 września 1995, t. 1 237-246, Uniwersytet Warszawski, Polonicum, Elipsa, Warszawa.

— 1998, "Wykładniki wartości informacyjnej wypowiedzenia w j. polskim i francuskim (aspekt, okresloność, modalność)", Congrès des Slavistes Cracovie 1998, *Revue des Etudes Slaves* T. LXX/1, p. 53-66, Paris.

— 1999, "Les marqueurs de la validation des énoncés en français et polonais", *Etudes cognitives / Studia kognitywne III*, SOW, PAN, Warszawa. pp. 135-162 (Erratum : dans cet article, il faut lire Σ pour Π et Π pour Σ).

— 2003, "L'Interprétation dynamique des centres d'intérêt dans les dialogues français et polonais", in *Études linguistiques romano-slaves offertes à Stanisław Karolak*, Oficyna Wydawnicza Edukacja, Cracovie pp.511-528.

2004, "Les Centres d'intérêt de l'énoncé en polonais et en français", in *Énoncer – l'ordre informatif dans les langues*, Cotte P., Dalmas M. & Wlodarczyk H. ed., Coll. Sémantiques, L'Harmattan, Paris, p.33-48.

Résumé

La théorie méta-informative (pragmatique) du centrage de la prédication distingue, dans l'énoncé

simple (c'est à dire ayant tout entier le statut *donné* ou *nouveau*), des constituants appelés *centres d'intérêt* (CI) : ce sont le *sujet*, défini comme CI global et *l'objet* comme CI local. Remarquons que, selon la même théorie, l'énoncé dit *étendu* est divisé en donné et nouveau et ses CI sont le *topique* (global) et le *focus* (local).

Le *sujet* est donc défini ici comme un constituant pragmatique (et non exclusivement syntaxique ni même syntactico-sémantique) ; le sujet parlant exprime au premier plan le participant de la situation sémantique qu'il a choisi comme centre d'intérêt principal de l'énoncé. Une telle conception *pragmatique* du problème de la prédication permet d'aboutir à une définition universelle du sujet. Bien que le sujet puisse remplir divers rôles sémantiques, il semble nécessaire d'introduire dans la théorie de l'énoncé le concept de *solidarité* entre le sujet et les *rôles passifs* (dans le cas des langues ergatives) ou le sujet et les *rôles actifs* (dans le cas des langues accusatives). La relation de solidarité n'exclut pas la possibilité de briser le lien qu'elle établit, car les langues naturelles disposent de moyens le permettant ; la diathèse est l'un des principaux. Ajoutons, pour être plus précis, que dans les langues ergatives – à l'inverse de ce qui se passe dans les langues accusatives — c'est la voix passive qui est *non marquée*.

Les énoncés appelés *impersonnels* ou *sans sujet* ont également été pris en compte. Nous y reconnaissons les traits caractéristiques des sujets dits « sous-entendus » que nous avons appelés *sujets anonymes*. Mais cela ne concerne pas les langues dans lesquelles, pour différentes raisons, la présence du sujet dans l'énoncé n'est pas obligatoire.

Streszczenie

Proponowana meta-informacyjna (pragmatyczna) teoria predykacji wyróżnia w wypowiedzeniu *prostym* (tj. w wypowiedzeniu mającym w całości *status datum* lub *novum*) składniki zwane *centrami uwagi* (CU), którymi są podmiot jako ogólne (globalne) CU oraz dopełnienie jako CU szczególne (lokalne). Zauważmy, że w tejże teorii wypowiedzenie zwane *rozszerzonym* jest podzielone na *datum* i *novum* a jego centrami uwagi są *topik* (CU ogólne) oraz *fokus* (CU szczególne).

Podmiot jest więc określony tutaj jako składnik pragmatyczny (nie zaś wyłącznie składniowy czy też syntaktyczno-semantyczny); mówiący wyraża pierwszo-planowo tego uczestnika sytuacji semantycznej, którego wybrał na główne centrum uwagi wypowiedzenia. Takie *pragmatyczne* ujęcie zagadnienia predykacji umożliwia zdefiniowanie podmiotu w sposób uniwersalny. Mimo iż podmiot może spełniać różne role semantyczne, koniecznym wydaje się wprowadzenie do nauki o języku pojęcia *solidarności* między podmiotem a *biernymi* (w przypadku j. ergatywnych) lub też *czynnymi* (w przypadku j. akuzatywnych) rolami semantycznymi. Relacja solidarności nie kłóci się z możliwością rozbicia tego związku, ponieważ języki naturalne dysponują środkami pozwalającymi na to. Najważniejszym takim środkiem jest diateza. Gwoli ścisłości, pragniemy zaznaczyć, że w j. ergatywnych — odwrotnie niż w j. akuzatywnych — strona bierna jest *nienacechowana*.

Uwzględnione zostały również tzw. wypowiedzenia *bezpodmiotowe* jako takie, w których rozpoznajemy cechy charakerystyczne dla podmiotów „domyślnych" bądź też tylko „zasygnalizowanych", a które nazwaliśmy *anonimowymi*. Nie dotyczy to jednak języków, w których z różnych przyczyn obecność podmiotu w wypowiedzeniu nie jest obowiązkowa.