Underspecified Semantics for Dependency Grammars ⋆

Alexander Dikovsky

LINA-FRE CNRS 2729, Université de Nantes, 2, rue de la Houssinière BP 92208 F 44322 Nantes cedex 3 France
Alexandre.Dikovsky@univ-nantes.fr

Abstract. We link generative dependency grammars meeting natural modularity requirements with underspecified semantics of *Discourse Plans* intended to *account for exactly those meaning components that grammars of languages mark for*. We complete this link with a natural compilation of the modular dependency grammars into strongly equivalent efficiently analysed categorial dependency grammars.

1 Introduction

Dependency grammars describe syntax in terms of dependencies: binary relations between words (see [1]). They generate dependencies explicitly, in contrast with grammars (HPSG, TAG, Type Logical Grammars etc.) calculating dependencies from other structures. Till recently, there was a common opinion that dependency grammars lack formal theory, in particular formal semantics. In recent papers, dependency grammars were defined as calculi of syntactic types: *Categorial Dependency Grammars (CD-grammars)* of [2, 3] and as generative *Dependency Structure Grammars (DS-grammars)* [4]. These grammars are weakly more powerful than CF-grammars, incomparable with mildly CS-grammars (in particular, with multi-component TAG) and expressive enough to represent discontinuous dependencies. The CD-grammars have an efficient polynomial time parsing algorithm [3] and are learnable from positive examples in rigid case [5].

Below we complete the DS-grammars meeting natural modularity requirements with underspecified formal semantics of *Discourse Plans* of [6] recently elaborated in [7], extendable to a higher order logical semantics. We demonstrate a natural morphism from the Discourse Plans to modular DS-grammars and show that these grammars are compiled into strongly equivalent CD-grammars.

2 Discourse Plans Semantics

2.1 Discourse Plans

DISCOURSE PLANS represent the course of event conceptualization by the speaker. They are *functional* in the following sense. Plans are *terms*, i. e. tree-like functional compositional structures. *All* elements (*semantemes*) of these structures,

⋆ This work was partially supported by the grants 04-01-00565 and 05-01-01006-a of the Russian Academy of Sciences.

with the only exception of special empty primitives (zero values, empty lists) have arguments and are functions (proper constants are set off in bold).

Semantic types and compositionality. Semantemes have functional types, defined from primitive types, of which there are four *initial basic types*: **n** (nominators, i. e. 'things' in the most general sense), **s** (sententiators, i. e. 'actions', 'processes', 'events', 'facts', etc.), **q** (qualifiers, 'meanings qualifying nominators'), and **c** (circumscriptors, 'meanings qualifying sententiators and qualifiers'). Basic types are extended by their specific instances. Thus the complete set of basic types is partially ordered under the *specific / generic relation* $\prec$. For instance: $\mathbf{n_a} \prec \mathbf{n}$ (animated nominator, e. g. ⟪$\mathbf{hearer^{n_a}}$⟫, $\mathbf{n_{mass}} \prec \mathbf{n}$ (mass nominator, e. g. ⟪$\mathrm{TV}^{\mathbf{n_{mass}}}$⟫, ⟪$\mathrm{milk}^{\mathbf{n_{mass}}}$⟫), $\mathbf{q_{qnt-mass}} \prec \mathbf{q_{qnt}} \prec \mathbf{q}$ (qualifier of mass nominators, e. g. ⟪$\mathrm{much}^{\mathbf{q_{qnt-mass}}}$⟫), $\mathbf{s_{attr}} \prec \mathbf{s}$, $\mathbf{s_{percep}} \prec \mathbf{s_{eff}} \prec \mathbf{s}$ (situations of attribution, direct perception and effect, e. g. ⟪$\mathrm{be_{attr}}^{\mathbf{s_{attr}}}$⟫, ⟪$\mathrm{kiss}^{\mathbf{s_{eff}}}$⟫, ⟪$\mathrm{watch}^{\mathbf{s_{percep}}}$⟫).

Every basic type **u** has a corresponding *optional* version $\mathbf{u}^{(0)}$ signifying optional arguments of this type. In addition, circumscriptor and qualifier types have *iterative* versions $\mathbf{u}^{(\omega)}$ (zero or more objects of type **u**). Together, basic, optional and iterative types constitute the set of *primitive* types. Finally, the primitive types serve to form complex *functional types* $(\mathbf{u_1} \ldots \mathbf{u_k} \to \mathbf{v})$ ($\mathbf{u_1} \ldots \mathbf{u_k}$ and **v** being respectively the argument and value types). E.g. the situation ⟪pour out⟫ has the type $\left(\mathbf{c}^{(\omega)}\mathbf{n_a}\mathbf{n_{lqm}^{(0)}}\mathbf{n_{ctr}} \to \mathbf{s}\right)$, where $\mathbf{n_{lqm}^{(0)}} \prec \mathbf{n_{mass}} \prec \mathbf{n}$ is the optional liquid matter nominator and $\mathbf{n_{ctr}} \prec \mathbf{n}$ is the container nominator type.

Compositionality is restricted by the condition that a subplan with the value of type $\mathbf{t_1}$ may be substituted in another plan of type $(\mathbf{u_1}\bar{\mathbf{u}} \to \mathbf{v})$ in the place of the argument of type $\mathbf{u_1}$, to obtain a composite plan of type $(\bar{\mathbf{u}} \to \mathbf{v})$, only if $\mathbf{u_1} = \mathbf{t_2}$ or $\mathbf{u_1} = \mathbf{t_2^{(0)}}$, and either $\mathbf{t_1} = \mathbf{t_2}$ or $\mathbf{t_1}$ is a case of $\mathbf{t_2}$ (denoted $\mathbf{t_1} \preceq \mathbf{t_2}$). In classical lexicographical terms, this means that the class of semantic compatibility of $\mathbf{u_1}$ is a superclass of that of $\mathbf{t_1}$, or that $\mathbf{u_1}$ is optional $\mathbf{t_1}$.

Roles, actants, circumstantials. We distinguish traditionally two basic classes of semantemes: *situations* and *monads*. A situation is a function, with one or more arguments, called *actants* and identified by their *(semantic) roles* (different arguments of the same situation have different roles) [1]. Arguments without roles are called *circumstantials*. They are grouped into a single list. So there is one generic circumstantial argument, whose iterative type is determined by the semanteme's value type: $\mathbf{q}^{(\omega)}$ for nominator-value semantemes and $\mathbf{c}^{(\omega)}$ for semantemes of other value types. Semantemes without actants are called *monads*.

Most verbs express situations: *John*$_{\text{SBJ}}$ <u>*gave*</u> *the letter*$_{\text{OBJ}}$ *to Fred*$_{\text{DST}}$; Fr. *Il*$_{\text{SBJ}}$ *me*$_{\text{OBJ}}$ <u>*dérange*</u> *par son chant*$_{\text{INS}}$ /He disturbs me by his singing/. By contrast, most names, adjectives and adverbs do not express situations, so that their meanings are represented by monads. A situation is specified by its *profile*, which includes the situation functor, the argument types, the situation value type and the role of each actant. E. g., the situations *give* and *déranger* have the profiles:

[1] In this short paper, we do not touch at the fundamental problem of adequate choice of roles (see [7] for details and discussion) and use quite intuitively some conventional roles, such as SBJ, OBJ, DST (destination), RCP (recipient), INS (instrument), etc.

$\langle\!\langle$give $\left(\text{S}\textsc{bj}^{\mathbf{n_a}}, \text{O}\textsc{bj}^{\mathbf{n}}, \text{R}\textsc{cp}^{\mathbf{n_a}}\right)^{\mathbf{s_{caus-mov}}}\rangle\!\rangle$ and $\langle\!\langle$déranger $\left(\text{S}\textsc{bj}^{\mathbf{n}}, \text{O}\textsc{bj}^{\mathbf{n_a}}, \text{I}\textsc{ns}^{\mathbf{n}^{(0)}}\right)^{\mathbf{s_{eff}}}\rangle\!\rangle$. Given such a profile, one can reconstruct the complete functional type of the situation. For instance, the complete type of ⟪give⟫ is $\left(\mathbf{c}^{(\omega)}\mathbf{n_a}\mathbf{n}\mathbf{n_a} \rightarrow \mathbf{s_{caus-mov}}\right)$ ($\mathbf{s_{caus-mov}} \prec \mathbf{s_{eff}}$ is the type of 'caused movement' situations; $\mathbf{c}^{(\omega)}$ stands for a possibly empty list of circumscriptors: circumstantial arguments of ⟪give⟫).

Communicative ranks. Two main instruments of event conceptualization are (1) choosing semantemes from the dictionary to designate an event and reference its participants and (2) assigning the speaker's point of view on the relative salience of participants (COMMUNICATIVE STRUCTURE).

For each situation in Discourse Plan, a communicative structure is assigned to its *communicative group*: the situation's functor and *all* arguments (not exclusively actants). Each communicative group element is assigned a *communicative rank*. We distinguish two THEMATIC ranks: *topic continuation* and *implied* and three RHEMATIC ranks: *focus*, *background* and *periphery*.

Topic continuation (denoted T). Assigned to the member of the communicative group to which the new (rhematic) information will be relativized. Corresponds to a referent initialized or previously evoked in the discourse context, which becomes, in the current plan point, the main entity under consideration. E. g.: *Those girls, they*$_{\mathsf{T}}$ *giggle when they see me*.

Implied (O). Assigned to a member of a communicative group if its referent is extremely salient due to deixis, anaphora, etc. (often elided in the surface form). E. g.: *Remember Mark?* (= *Do you*$_{\mathsf{O}}$ *remember Mark?*).

Focus ($\odot$). Assigned to the member of a group, which conveys the new (as opposed to given or presupposed) information.

Background ($\oplus$). Assigned to those members of a group, which convey other pertinent information (and so cannot be dismissed). E. g.: *I*$_{\mathsf{T}}$ *hit the stick*$_{\oplus}$ *against the fence*$_{\odot}$.

Periphery ($\ominus$). Assigned to the members of a group, which become non-salient and should be dismissed. E. g. in the answer to *Do you sell your car? – Sorry, it's already been bought*, the figure of the buyer obtains the rank $\ominus$.

Diatheses and diathetic shifts. SEMANTIC DIATHESES [2] are discourse plan elements specifying, for a situation, the intended change in types and communicative ranks with respect to the dictionary profile and the prototypical rank assignment [3]. Every semantic diathesis is implemented by the corresponding *diathetic shift*, i.e. the transformation of the situation profile licensed by this diathesis (a definition of possible diathetic shifts can be found in [6]). Here, we limit ourselves to several examples.

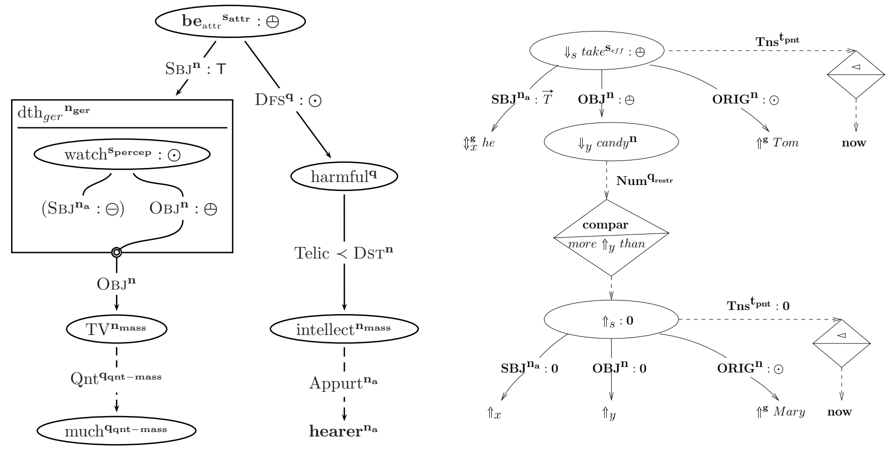
Example: English passivization occurs when the OBJ actant of an effect situation is promoted to the rank of topic. $\text{dth}_{f-passive}$ (PAGT ↜ SBJ$_{\odot}$, SBJ ↜ OBJ$_{\mathsf{T}}$)

[2] Diatheses used in discourse plans are close to those of [8].

[3] The *prototypical* assignment is as follows: thematic ranks are assigned only to SBJ; $\odot$ is assigned to the rightmost argument (in the order of plan points) which is not otherwise assigned a thematic rank (or to the situation functor, if no such argument exists); other arguments are assigned $\oplus$.

is one of passivization diatheses in English. If we take the situation which, under prototypical rank assignment, has the form *Eve*$_{\text{SBJ}:\top}$ *gave*$_{\oplus}$ *the apple*$_{\text{OBJ}:\oplus}$ *to Adam*$_{\text{RCP}:\odot}$, the diathetic shift for f-passive transforms it to *The apple*$_{\text{SBJ}:\top}$ *was given*$_{\oplus}$ *to Adam*$_{\text{RCP}:\oplus}$ *by Eve*$_{\text{PAGT}:\odot}$.

Example: The gerund in English is brought about by the need to demote the subject actant and to change the type of the affected situation to $\mathbf{n_{ger}} \prec \mathbf{n}$. The plan in Fig. 1(a) shows the gerund diathesis $\langle\!\langle \text{dth}_{ger}\,(\varnothing \leftharpoondown \text{SBJ}_{\ominus})\,^{\mathbf{n_{ger}}} \rangle\!\rangle$ being applied to the situation $\langle\!\langle \text{watch}\,(\text{SBJ}^{\mathbf{n_a}}, \text{OBJ}^{\mathbf{n}})\,^{\mathbf{s_{percep}}} \rangle\!\rangle$.



(*a*) *Watching much TV is harmful for your intellect* (*b*) *He took more candies from Tom than from Mary*

Fig. 1. Examples of discourse plans

(1) Ellipses signify functional semantemes. (2) The box represents a diathesis (communicative ranks assignment) and the corresponding diathetic shift (the deleted actants, which are bracketed, the connections to the outside of the box, the projection of roles defined by the situation onto those defined by the diathetic shift). (3) Solid lines link actants to their situations and are labeled with roles. (4) Dashed lines link circumstantials to their governor semantemes and are labeled with names of attributes. (5) Diamonds represent embedded functions applying to attributes (e.g., $\lhd$: pointwise and interval time precedence).

Dynamic context. Realizable plans. In discourse plans, one can use embedded functions accessing two contexts: invariable long term *global context* and dynamic short term *local context* which is updated in the course of planning. In the plan in Fig. 1(a) we see four context access functions: $\Uparrow^{\mathbf{g}}$ *Tom* and $\Updownarrow^{\mathbf{g}}_{x}$ *he* access the referents of *Tom* and *he* in the global context (the latter also introduces local reference x), $\Downarrow_{s}\text{take}^{\mathbf{s_{eff}}}$ adds to the local context local reference s to the function $\text{take}^{(\mathbf{c}^{(\omega)}\mathbf{n_a}\mathbf{n}\mathbf{n}^{(0)} \to \mathbf{s_{eff}})}$ dereferenced by $\Uparrow_{s}$. The local context is a bounded resource memory: adding to it a new reference may cause deletion of some other reference. Exact reference scoping rules can be found in [6].

In order for a discourse plan to be realizable, it is necessary that all local references satisfy scoping rules and the rank assignment for every situation in the plan must match the definition of a diathesis in the dictionary and must be implementable by one of the shifts licensed by this diathesis.

2.2 Relation with complete semantics

Realizable plans are transformed by a semantical morphism *sem* into (meaning) λ-terms. *sem* is relativized to a signature interpretation I which interprets semantemes and embedded signature functors by 1st order λ-terms and diatheses by 2d order λ-terms. In this semantics, actants and circumstantials are translated differently:

$$sem^I\big[K(\pi_1^{\mathbf{u}}\!:\!R \mid \bar{\pi})\big] = \big(\big(K^I\ sem^I[\pi_1]\big)\ \ sem^I[\bar{\pi}]\big)\,,$$

if $\mathrm{K}^{(\mathbf{u_1}\bar{\mathbf{u}}\to\mathbf{v})}$ is either a dictionary semanteme or a semantical derivative $(S_{dth}K_0)$ of a situation K_0 through diathetic shift S_{dth} of some licensed diathesis dth of K_0, π_1 is the subplan of the plan $K(\pi_1^{\mathbf{u}} : R \mid \bar{\pi})$ for the actant of K identified by the role R and $\mathbf{u} \prec \mathbf{u}_1$. Much as the semantics of an actant is determined by its role, the semantics of a circumstantial is determined by an ATTRIBUTE. The difference is that the attributes are specified only by plans, but not by the dictionary profiles.

$$sem^I\Big[K\Big(\big[\pi_1^{\mathbf{u}}\!:\!A \mid \bar{L}\big]^{\mathbf{u}_1^{(\omega)}} \mid \bar{\pi}\Big)\Big] = \big(\big(A^I\ sem^I[\pi_1]\big)\ \ sem^I\big[K^I(\bar{L} \mid \bar{\pi})\big]\big)\,,$$

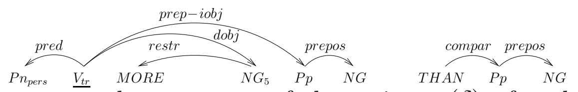
if $\pi_1^{\mathbf{u}}$ is the first circumstantial in the circumstantial list $\big[\pi_1^{\mathbf{u}} \mid \bar{L}\big]^{\mathbf{u}_1^{(\omega)}}$ of $\mathrm{K}^{\left(\mathbf{u}_1^{(\omega)}\bar{\mathbf{u}}\to\mathbf{v}\right)}$ in the plan $K\Big(\big[\pi_1^{\mathbf{u}} \mid \bar{L}\big]^{\mathbf{u}_1^{(\omega)}} \mid \bar{\pi}\Big)$ and $\mathbf{u} \prec \mathbf{u}_1$.

Example: Appurtenance. The plan in Fig. 1(a) conceptualizes the noun phrase *your intellect* using the monad $\left\langle\!\left\langle \mathrm{intellect}^{(\mathbf{q}^{(\omega)}\to\mathbf{n_{mass}})}\right\rangle\!\right\rangle$ whose *appurtenance* attribute *Appurt* applies to $\langle\!\langle \mathbf{hearer^{n_a}} \rangle\!\rangle$. The meaning of this attribute is defined through the meaning of the preposition $\langle\!\langle \mathrm{of} \rangle\!\rangle$ whose logical type is $(\mathbf{e} \to \mathbf{t}) \to ((\mathbf{e} \to \mathbf{t}) \to (\mathbf{e} \to \mathbf{t}))$. Extending underspecified types to logical, we obtain the traditional semantics :

$sem^{log}[intellect([\mathbf{hearer}\!:\!Appurt])] = \Big(\Big(\mathrm{of}^{log}\ \mathbf{hearer}^{(\mathbf{e}\to\mathbf{t})}\Big)\ \ \mathrm{intellect}^{(\mathbf{e}\to\mathbf{t})}\Big)\,.$

3 Dependency Structure Grammars

If it were not for distant dependencies, DSG would be CF generative dependency grammars. To work with discontinuities, DSG use *generalized dependency structures (gDS)*: linear ordered graphs with labeled arcs (dependencies) and nodes (words or nonterminals), in which some maximal connected component is selected as *head component* and some node in this component is selected as *gDS head*. One-component gDS is a *dependency tree*. For instance, in two-component gDS below the underlined node is head.

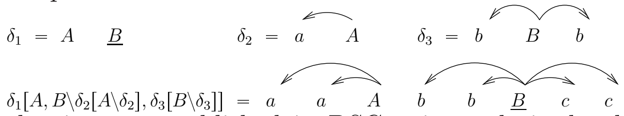


Each gDS δ can be seen as the structure of the string $w(\delta)$ of node labels in natural order.

DSG use the following composition $\delta[\alpha\backslash\delta_1]$ (simultaneous composition: $\delta[\alpha_1, \ldots, \alpha_n\backslash\delta_1, \ldots, \delta_n]$) of gDS. $\delta[\alpha\backslash\delta_1]$ results from gDS δ by substituting gDS δ_1 for node α of δ so that should be preserved:

(i) the order of nodes in δ, δ_1: $w(\delta[\alpha\backslash\delta_1]) = x\ w(\delta_1)\ y$, if $w(\delta) = x\ label(\alpha)\ y$,

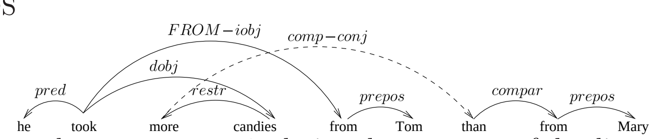
(ii) the dependencies in δ, δ_1: the head of δ_1 inherits dependencies of α in δ [4], as in the following example:



Discontinuous dependencies are established in DSG using polarized valencies and potentials.

Definition 1 *For a finite set* $\mathbf{C}$ *of dependency types, four polarities are introduced: left and right positive:* $\nearrow, \nwarrow$ *and left and right negative:* $\swarrow, \searrow$ *. For instance, the polarities* $\nearrow$ *and* $\searrow$ *mean* outgoing *(respectively,* incoming*)* from left to right. *For each polarity* v*, there is the unique "dual" polarity* $\breve{v}$*:* $\breve{\nearrow} = \searrow$, $\breve{\nwarrow} = \swarrow$, $\breve{\swarrow} = \nwarrow$, $\breve{\searrow} = \nearrow$. *A* polarized valency *is an expression* (vC)*, in which* v *is one of the four polarities and* $C \in \mathbf{C}$. $\mathbf{V}$ *is the set of all valencies.*

For instance, in gDS



the long distance dependency *comp–conj* relating the two parts of the discontinuous comparative construction *more..than* can be defined using dual valencies $\nearrow comp-conj$ and $\searrow comp-conj$.

Definition 2 *Valency strings* $\Gamma \in \mathcal{P} =_{df} \mathbf{V}^*$ *are called* potentials. *Let* $\Gamma = \Gamma_1(vC)\Gamma_2(\breve{v}C)\Gamma_3$ *and* $\Gamma' = \Gamma_1\Gamma_2\Gamma_3$ *be two potentials such that* $(vC) = (\nearrow A)$, $(\breve{v}C) = (\searrow A)$ *or* $(vC) = (\swarrow A), (\breve{v}C) = (\nwarrow A)$. (vC) *is* **first available (FA)** *for* $(\breve{v}C)$ *in* Γ *and both are* neutralized *in* Γ' *(denoted* $\Gamma \twoheadrightarrow_{FA} \Gamma'$*) if* Γ_2 *has no occurrences of* (vC) *and* $(\breve{v}C)$. *This reduction of potentials* $\twoheadrightarrow_{FA}$ *is terminal and confluent. So each potential* Γ *has a unique FA-normal form denoted* $[\Gamma]_{FA}$. *The* product $\odot$ *of potentials defined by:* $\Gamma_1 \odot \Gamma_2 =_{df} [\Gamma_1\Gamma_2]_{FA}$ *is clearly associative. So we obtain the monoid of potentials* $\mathbf{P} = (\mathcal{P}, \odot)$ *with the unit* ε.

Definition 3 *A* **Dependency Structure Grammar (DSG)** G *has the* rules $r = (A \to \delta)$ *with* $A \in N$ *and a gDS* δ*, on which* potential assignments $[\Gamma^L](r, X, \omega)[\Gamma^R]$ *may be defined,* (r, X, ω) *being an occurrence of* X *in* δ *and* Γ^L, Γ^R *being left and right potentials assigned to this occurrence* [5].

Derivation trees of G *result from derivation trees* T *of cf-grammar* $\{A \to w(\delta) \mid A \to \delta \in G\}$ *by defining potentials* $\pi(T, n)$ *of nodes* n *:*
1. $\pi(T, n) = \varepsilon$ *for every terminal node* n*;*
2. $\pi(T, n) = \Gamma_1 \odot \ldots \odot \Gamma_k$*, for every node* n *with sons* $n_1, ..., n_k$ *derived by rule* $r = (A \to \delta)$*, in which* $w(\delta) = X_1...X_k$ *and* $\Gamma_i = \Gamma_i^L \odot \pi(T, n_i) \odot \Gamma_i^R$*, where* $[\Gamma_i^L](r, X_i, \omega_i)[\Gamma_i^R]$ *are the rule potential assignments. A gDS is generated in the node* n *by the composition:* $gDS(T, n) = \delta[X_1 \ldots X_k \backslash gDS(T, n_1),$

[4] We use nonterminals $label(\alpha)$ in the place of α when no conflicts.
[5] For instance, $A \to [\searrow D_1]\underline{B}[\nearrow D_2]\ C$ denotes the rule $A \to \underline{B}\ C$ with assignment $[\searrow D_1]B[\nearrow D_2]$. We omit empty potentials.

$\ldots, gDS(T, n_k)]$. *Every pair of dual valencies neutralized at this step corresponds to a long distance dependency added to this gDS.*

A derivation tree T *is* complete *if the potential of its root* S *is neutral:* $\pi(T, S) = \varepsilon$. *We set* $G(D, w)$ *if there is a complete derivation tree* T *of* G *from the axiom* S, *such that* $D = gDS(T, S)$ *and* $w = w(D)$.

$\Delta(G) = \{D \mid \exists w \in W^+ \; G(D, w)\}$ *is the* gDS-language *generated by* G.

$L(G) = \{w \in W^+ \mid \exists D \; G(D, w)\}$ *is the* language *generated by* G.

For instance, the following four-rule DSG:

$$G_1: \quad S \rightarrow \; a[\swarrow a] \; \underline{S} \quad \| \quad A \quad c \qquad A \rightarrow \; [\nwarrow a]b \quad A \quad c \quad \| \quad [\nwarrow a]b$$

generates the language $L(G_1) = \{a^n b^n c^n \| n > 0\}$.

4 Modular DSG

As expressive as they are (see [4]), the DSG are only simplified models of grammars usable in real linguistic applications. They don't support features, and propose no link with semantics. The former is easy to fix. We will suppose that terminals and nonterminals are decorated with finite-valued feature structures. We will extend terminals to *proto-terminals* (denoted $\{X\}$): classes of words with unifiable feature structures. We will also extend the rules by *constraints*: conditions of applicability in terms of feature values, and we will admit iterated dependencies (denoted $d*$). In order to link these generalized DSG (*gDSG*) with the discourse plans, we choose their subclass that we dub *modular DSG*.

Definition 4 Modular DSG (MDSG) *is a gDSG decomposable into a union of subgrammars (*modules*)* $G = \bigcup\limits_{(M)} G_{(M)}$ *with disjoint sets of nonterminals (the nonterminal* (M) *serves as the axiom of* $G_{(M)}$*), satisfying conditions:*
$(\mathbf{m_1})$ *if several rules of* $G_{(M)}$ *share the same left hand nonterminal:* $A \rightarrow \delta_1/c_1 \,\|\, \ldots \,\|\, \delta_k/c_k$, *then the constraints* c_i *are one-to-one exclusive;*
$(\mathbf{m_2})$ *terminals of* $G_{(M)}$ *are proto-terminals of* G *or axioms of other modules;*
$(\mathbf{m_3})$ *in every module, no recursion is possible through the heads of right-hand gDS of its rules;*
$(\mathbf{m_4})$ *in every rule* $A \rightarrow \delta$, δ *has at most one (possibly iterated) dependency.*

As it is clear from definition 4, due to the modularity constraints, each complete derivation of a MDSG G is a composition of derivations in modules $G_{(M)}$ (cf. example in Fig. 2(a)). These module derivations start with module axiom (M) and have in the leaves either proto-terminals of G or other modules' axioms. We call such derivations of $G_{(M)}$ *m-terminal*. m-terminal derivations are *minimal* if each iterative dependency is iterated at most once. Clearly, there are finitely many minimal derivations in each MDSG. More importantly, each minimal derivation D in a module $G_{(M)}$ is uniquely identified by the constraint σ on features, which is a composition of constraints of rules applied in D. We will call *d-form* a minimal m-terminal derivation decorated by its feature constraint: $D_\sigma =_{df} D \circ \sigma$ (see the verb group module d-form in Fig. 2(b)).

Example of English verb group module:

$$(VG) \to VG_1 \;/dist = 0 \;\;\|$$

$$\underline{VG_1} \quad (THAN-compar)$$
$$/dist = comp-conj$$

$$VG_1 \to (PreCG) \overset{pre-circ*}{\longleftarrow} VG_2$$

$$VG_2 \to VG_3 \overset{post-circ*}{\longrightarrow} (PostCG)$$

$$VG_3 \to (NG) \overset{pred}{\longleftarrow} VG_4$$
$$/SBJ.cr = \top, SBJ.type = \mathbf{n_a}, \boxed{\text{SBJ=pred.subd}} \;\;\|$$

$$VG_4 \;/SBJ.cr = \circ$$

$$VG_4 \to VG_5 \;/asp = nperf \;\;\|$$

$$AuxG \overset{aux}{\longrightarrow} (PerfG) \;/asp = perf$$

$$VG_5 \to VG_6 \overset{inf-obj}{\longrightarrow} (V_{inf}G)$$
$$/PRED.type = \mathbf{s}_{mph}, \boxed{\text{OBJ=infobj.subd}} \;\;\|$$

$$VG_6 \overset{sent-obj}{\longrightarrow} (Obj-cl)$$
$$/PRED.type \neq \mathbf{s}_{mph}, \boxed{\text{OBJ=sentobj.subd}} \;\;\|$$

$$VG_6 \;/\boxed{\text{OBJ=dobj.subd}} \;\vee\; OBJ.cr = \circ$$

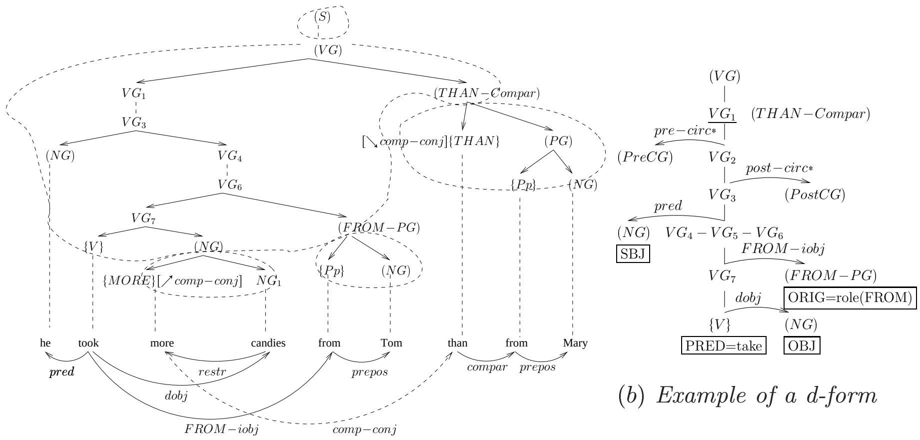
$$VG_6 \to VG_7 \overset{P-iobj}{\longrightarrow} (P-PG)$$
$$/role(P).type = \mathbf{n}, role(P).cr = \odot, \boxed{\text{role(P)=Piobj.subd}} \;\;|$$

$$VG_7 \;/PRED.act < 3 \;\vee\; role(P).cr = \circ$$

$$VG_7 \to \{V\} \overset{diobj}{\longrightarrow} (NG)$$
$$/\boxed{\text{prof(PRED)=prof(V)}}, \boxed{\text{OBJ=dobj.subd}}, OBJ.type = \mathbf{n},$$
$$(OBJ.cr = \circ \;\vee\; (role(P).cr \neq \circ = OBJ.cr) \;\vee$$
$$(OBJ.cr = \oplus \;\wedge\; role(P).cr = \odot)) \;\;\|$$

$$\{V\} \;/\boxed{\text{prof(PRED)=prof(V)}},$$
$$(PRED.act < 2 \;\vee\; OBJ.cr = \circ)$$

LEGEND: For a role R, $R.cr$ and $R.type$ denote the communicative rank and the type of the actant with the role R; for a preposition P, $role(P)$ is the role of the oblique actant represented by the corresponding prepositional phrase (e.g., $role(FROM) = ORIG$); $PRED$ is the situation functor's role, $PRED.act$ is the number of its actants, $prof(PRED)$ is the profile of the situation; for a dependency d, $d.subd$ is the subordinate; the boxed equations define a morphism from derivations onto discourse plans (e.g., $OBJ = dobj.subd$ maps direct objects onto actants with the role OBJ).



(b) *Example of a d-form*

(a) *Complete derivation in MDSG*

Fig. 2. Modular derivations

Linking Discourse Plans to MDSG. Semantical interpretation Φ of a MDSG is defined through its d-forms. As it is explained above (see the legend to the verb group module), the morphism Φ is defined through equations of the form $Role = dependency.subd$. A simplified rule of top-down construction of $\Phi(D)$ for a modular complete derivation D is as follows. 1. Determine the topmost m-terminal subderivation D_0 of D and its module $G_{(M)}$. 2. Define the d-form F of which D_0 is an instance. 3. Find the semanteme equation $PRED = Key$ in

F [6]. 4. For each actant A of Key with $Key.cr \notin \{\ominus, \mathsf{O}\}$, find the corresponding object dependency d in F and set $\Phi((d.subd)) = A$. 5. For each instance c_0 of an iterated dependency $c*$ in D_0, determine from the feature constraint the corresponding circumstantial C and the attribute a, set $\Phi((c.subd)) = C$ and decorate this argument with a in the plan. □
E.g., the d-form in Fig. 2(b) of the second topmost subderivation in Fig. 2(a) defines the topmost situation *take* and its arguments in the plan in Fig. 1(b).

5 Compiling modular DSG into categorial DG

Categorial dependency grammars (CDG) are simply related with the classical categorial grammars. They assign to words "curried" variants of first order types: $[l_1\backslash \ldots \backslash m/ \ldots /r_1]$, whose left ($l_i$), right ($r_i$) and main ($m$) subtypes can be elementary or polarized. When a word w is assigned an elementary type l (denoted $w : l$) and a phrase p has type $[l\backslash\alpha]$ ($p : [l\backslash\alpha]$), then $l[l\backslash\alpha] \models \alpha$ ($wp : \alpha$) and dependency l is established from the head of p to w. When two distant phrases have dual main and argument valency types, e.g. $p_1 : (\swarrow d)$ and $p_2 : [(\nwarrow d)\backslash\alpha]$, then $(\swarrow d)..[(\nwarrow d)\backslash\alpha] \models \alpha$ ($p_1..p_2 : \alpha$) and distant dependency d is established from the head of p_2 to the head of p_1 if the two dual types are in a sense "closest". To express immediate neighborship, valency types can be decorated by two adjacency modalities: $\flat$ (host) and $\#$ (anchored). $p_1 : \#(\swarrow d)$ means that $p_1 : (\swarrow d)$ must be anchored immediately on the left of a phrase host for left $(\swarrow d)$. $p_2 : [\flat(\swarrow d)\backslash\alpha]$ means that p_2 is host for left $(\swarrow d)$. So $\#(\swarrow d)[\flat(\swarrow d)\backslash\alpha] \models (\swarrow d)\alpha$ and $p_1p_2 : (\swarrow d)\alpha$. For instance, $[\flat(\swarrow clit{-}dobj)\backslash subj\backslash S/aux]$ is one of possible types of auxiliary verbs in French, which defines them as host words for a cliticized direct object.

A calculus of dependency types in CDG and an efficient polynomial time parsing algorithm for it can be found in [3]. CDG can be transformed in *strongly* equivalent DSG and DSG can be transformed in *weakly* equivalent CDG [4]. As it turns out, MDSG are equivalent to DSG and CDG in the following sense.

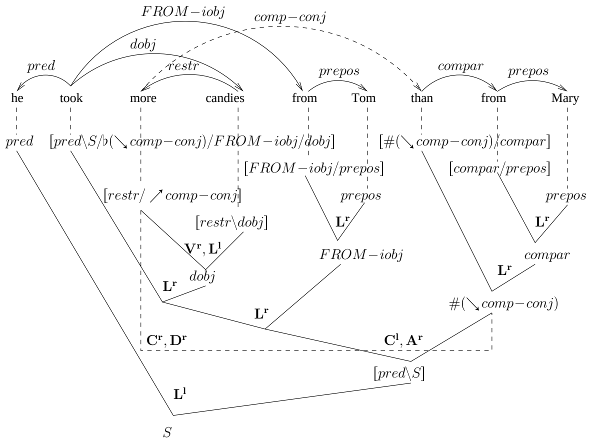
Theorem 1 *For each DSG, there is a weakly equivalent MDSG.*

Theorem 2 *Each MDSG G can be transformed into a CDG G' such that $\Delta(G') = \Delta(G)$ (and so $L(G') = L(G)$).*

Construction sketch. 1. Specialization. Resolve feature constraints to construct all minimal derivations D_σ. Let D_σ be such derivation in a module $G_{(M)}$. Its head leaf is a proto-terminal $\{W\}$. Assign to all word forms $w \in \{W\}$ the types $c_\sigma \underset{\mathrm{df}}{=} [\gamma_{l_m}\backslash \ldots \backslash\gamma_{l_1}\backslash H/\gamma_{r_1}/ \ldots /\gamma_{r_m}]$ constructed as follows. **2. Subcategorization.** Let $r_1, \ldots, r_m$ be the sequence of rules applied in D_σ and $\backslash v_1 \ldots v_k \underset{\mathrm{df}}{=} \backslash v_1 \ldots \backslash v_k$. Then $\gamma_{l_i} = \backslash\Gamma_{l_i}\backslash d_{l_i}$, if r_{l_i} defines left dependency d_{l_i} and assigns to the rule head h left potential Γ_{l_i} (similar for

[6] We suppose that F has this equation. It is not always the case: in the verb group module above, the second rule for VG_4 defining perfective verb forms presumes two modules: one for the auxiliary taking the subject argument and the other for the participle taking all complement arguments and determining the semanteme key.

right subtypes). Otherwise, $\gamma_{l_i} = \backslash\flat^l(v_{l_i})\backslash\Gamma_{l_i}$, if r_{l_i} defines no left dependency, assigns Γ_{l_i} to h and if a negative valency v_{l_i} is assigned to the head of a gDS in a minimal derivation in the module $G_{(L)}$ of the left neighbor (L) of h. **3. Main subtype definition.** $H \underset{\mathrm{df}}{=} d$, if (M) is subordinate through dependency d in some rule of G, $H \underset{\mathrm{df}}{=} \#^l(v)$, if isolated node (M) is assigned negative valency v at the step 2, $H \underset{\mathrm{df}}{=} S$ otherwise. □



Compiled CDG:
This proof uses types compiled, using the method of Theorem 2, from the rules of MDSG, whose derivation is shown in Fig. 2(a). In particular, the type of *take* is compiled from the minimal m-terminal derivation of this grammar, whose d-form is shown in Fig. 2(b).
Dependency type assignments:
he: $pred$
took: $[pred\backslash S/\flat(\searrow comp-conj)/FROM-iobj/dobj]$
more: $[restr/\nearrow comp-conj]$
candies: $[restr\backslash dobj]$
from: $[FROM-iobj/prepos]$, $[compar/prepos]$
Mary, Tom: $pred, prepos$
than: $[\#(\searrow comp-conj)/compar]$

Fig. 3. A proof in CDG compiled from a MDSG.

6 Conclusion

The Discourse Plans are intended to establish a formal link between dependency syntax and referential semantics on a minimal natural language specific cognitive base. The logical aspect of this link will be presented elsewhere.

References

1. Mel'čuk, I.: Dependency Syntax. SUNY Press, Albany, NY (1988)
2. Dikovsky, A.: Dependencies as categories. In: "Recent Advances in Dependency Grammars". COLING'04 Workshop. (2004) 90–97
3. Dekhtyar, M., Dikovsky, A.: Categorial dependency grammars. In: Proc. of Intern. Conf. on Categorial Grammars, Montpellier, France (2004) 76–91
4. Béchet, D., Dikovsky, A., Foret, A.: Dependency structure grammars. In: Proc. of the 5th Int. Conf. "Logical Aspects of Computational Linguistics". LNAI 3492 (2005) 18–34
5. Béchet, D., Dikovsky, A., Foret, A., Moreau, E.: On learning discontinuous dependencies from positive data. In: Proc. of the 9th Intern. Conf. "Formal Grammar 2004". (2004) 1–16
6. Dikovsky, A.: Linguistic meaning from the language acquisition perspective. In: Proc. of the 8th Intern. Conf. "Formal Grammar 2003". (2003) 63–76
7. Dikovsky, A., Smilga, B.: Semantic roles and diatheses for functional discourse plans. In: Proc. of the 2nd International Conference "Meaning-Text Theory", Moscow, Russia (2005) 98–109
8. Padučeva, E.V.: Diathesis and diathetic displacement. Russian Linguistics **26** (2002) 179–215