

Discourse, Themes and Word Order in Head-Driven Phrase Structure Grammar

Patrick M. Murphy

The University of North Carolina at Chapel Hill

murphy@email.unc.edu

1 Introduction

The framework of Head-Driven Phrase Structure Grammar (HPSG) boasts many elegant analyses of syntactic and semantic phenomena. However, little emphasis to date has been placed on syntactic structures when influenced by semantic or pragmatic considerations, such as that of *theme* and *rheme* in languages like Russian. Clearly, Russian grammar follows other principles, but the role of contextual information is crucial to determining word order and intonational patterns; these are facts which are hard to reconcile with generative frameworks which have no account of pragmatics. This paper seeks to lay the foundation for an analysis of pragmatic concerns in HPSG, building from analyses by Pollard and Sag (1987, 1994), and concentrating on the ‘simplest’ examples of literary Russian.

In a language like Russian, with a high degree of case marking, it is unsurprising to find very free word order. However, there is an important constraint on word order; Russian divides of sentences into two parts: one of old information, and one of new information. Such an analysis, that of the functional sentence perspective, labels these parts *topic* and *comment* or *theme* and *rheme*. Their purpose, roughly, is that the theme indicates the subject of the utterance, and the rheme conveys new information essential to the speaker (Krylova and Khavronina 1988:11). Just what part of the sentence qualifies for these roles may depend on any number of different factors, including such elusive phenomena as the emotional state of the speaker and intonational patterns involved; herein lies the problem of a generative analysis of the facts.

2 The Russian Facts

Without complicating an analysis of Russian by considering such influences, however, my approach will instead start by examining ‘regular’ word order, or the speech patterns involved in statements which lack emotional force. In such cases, *i.e.*, non-emotive speech, the theme precedes the rheme. This is accompanied by a ‘narrative’ intonation; the theme is given a slightly rising pitch, and the rheme has a falling pitch at the end of the sentence (Krylova and Khavronina 1988:12). These sorts of sentences are considered ‘stylistically neutral’.

- (1) Lev Tolstoj – avtor romana “Vojna i mir.”
Lev Tolstoj author novel-GEN “War and Peace.”

‘Lev Tolstoj was (is) the author of the novel *War and Peace*.’

- (2) Avtor romana “Vojna i mir” – Lev Tolstoj.
author novel-GEN “War and Peace” Lev Tolstoj

‘Lev Tolstoj was (is) the author of the novel *War and Peace*.’

To determine which of the above utterances is most appropriate, one must examine the context in which they occur. Example (3) should elicit the response (1); likewise, (4) should produce (2):

- (3) Kto Lev Tolstoj?
who Lev Tolstoj

‘Who was Lev Tolstoj?’

- (4) Kto napisal roman “Vojna i mir”?
who wrote novel-ACC “War and Peace”?

‘Who wrote the novel *War and Peace*?’

Crucially, there are times when it is unclear how a sentence should be divided into theme and rheme, and this must be supplied by context. That is, context is the determining factor for sentence division. Considering the following sentence:

- (5) Letom my sobiraems’a poexat’ na Volgu.
summer-INST we planning-1P to-go to Volga-ACC.

‘This summer we are planning to go to the Volga.’

It is unclear whether the theme-rheme division should be *Letom my—sobiraems’ a poexat’ na Volgu* or *Letom my sobiraems’ a poexat’—na Volgu*. The first division is appropriate to answer questions like *čto vy budete delat’ letom?* ‘What will you do this summer?’ or *Kakie u vas plany na leto?* ‘What kind of plans do you have for the summer?’ but the second agrees with the question *Kuda vy cobiraetes’ poexat letom?* ‘Where do you plan to go this summer?’. It is interesting to note that these divisions occur without respect to constituent structures like VP, etc.; the theme or rheme may be comprised of subject and predicate, subject and object, predicate and object, and so on.

There are also utterances which contain only ‘new’ information, and thus cannot be divided into a theme and rheme. Such sentences include:

- (6) Šol dožd’.
went rain-NOM
 ‘It rained.’

- (7) Neslyšno proletela kakaja—to neizvestnaja ptitsa.
quietly flew-in some unknown bird-NOM
 ‘Some unknown bird quietly flew in.’

These sentences are traditionally considered to have a ‘zero theme,’ whereas the ‘indivisible utterances’ themselves make up the rheme. Word order is fixed in these examples, however, and conforms to the following pattern: predicate—subject. A further interesting fact is that verbs in such indivisible utterances are invariably intransitive (Krylova and Khavronina 1988:26).

The catch is that the rules above govern formal or written Russian but not that of ‘emotive’ or colloquial spoken Russian. In fact, about the only criteria used to identify emotive Russian is that it breaks one or more of the rules outlined above; *e.g.*, theme and rheme are inverted, order of elements with a zero theme is changed, etc. Still, there remains an intonational pattern which marks the boundary between theme and rheme, so an account whereby the ‘theme intonation’ could be assigned to some elements would be contribute significantly towards accounting for the data.

3 Towards a Theory of Discourse

In my initial analysis of the theme/rheme phenomena, I will assume that what is important is a constraint on the ordering of the phonological realization of elements, depending on their relevance to material previously expressed. In other words, I will assume that the syntactic constraints on

sentences in Russian must be met independently of the theme/rheme distinction. We then predict that improper separation of a sentence into theme and rheme, switching their positions, etc., will be ‘bad’ in a different sort of way than utterances that violate purely syntactic rules (much like a sentence can still be grammatical and not make sense or be infelicitous). Well-formed Russian sentences must still obey the principles of Russian syntax, akin to those like the ID Principle and Head Feature Principle of Pollard and Sag (1994). Phrases must still be headed, SLASH features still require a Filler-Gap construction, etc.; pragmatic concerns will dictate the order of constituents in an otherwise grammatical sentence. Following Pollard and Sag, I will assume that there are schemata dictating the linear precedence of constituents.

One area which needs to be explored is what constitutes ‘old information.’ A first attempt at quantification might attempt to link the theme with the content of the preceding sentence, but this seems generally false. According to Pollard and Sag (1994), the content of a sentence, and possibly that of a discourse, should be the combination of the content of its daughter words/phrases/sentences, by the Principle of Contextual Consistency:

(8) PRINCIPLE OF CONTEXTUAL CONSISTENCY:

The CONTEXT|BACKGROUND value of a given phrase is the union of the CONTEXT|BACKGROUND values of the daughters.

(Pollard and Sag 1994:333)

Pollard and Sag themselves point out that this principle will not hold; conditionals and propositional verbs may affect presupposition. In the following examples, only the first set have as a presupposition (c):

- (9) a. Pat regrets that Terry is dead.
b. Kim fears that Pat regrets that Terry is dead.
c. Terry is dead.

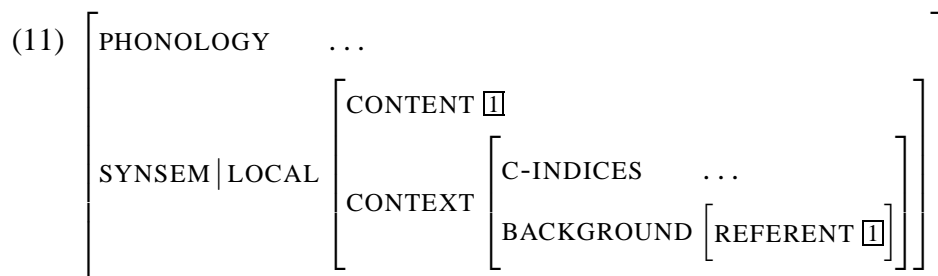
- (10) a. Pat regrets that Terry is dead.
b. If Terry is dead, then Pat regrets that Terry is dead.
c. Terry is dead.

(Pollard and Sag 1994:333)

Of course, it is interesting to note that while this may be the case, it seems to have little to do with the theme/rheme distinction. Presupposition does not seem to be paramount; whether or not

the theme has been discussed before does. Intuitively, the theme should be something relating to a previous utterance, but it also is certainly not the same thing as a previous utterance, as sentences (2) and (4) demonstrate; here words like *avtor* ‘author’ show up and are ‘old news’ even though not previously uttered.

Our goal is to ensure that the theme contains some information previously introduced in the course of the discourse. We cannot accomplish this by making *theme* a subsort of *discourse*, or something along those lines, as the theme will then not only inherit the relevant information and proper features from the discourse, but a good bit of irrelevant information as well. This notion will also have to be refined, but for our first formulation let us presume that the CONTEXT|BACKGROUND of sort *theme* must be token-identical with some element(s) of the discourse, and that this will insure that the theme’s BACKGROUND is some subset of that of the discourse as a whole. That is, a sign will be chosen from the discourse, and if we assign it the index $\boxed{1}$, then the theme will tentatively look something like this:



Unfortunately, the current formalizations will not allow us to maintain our generalizations about themes and the discourse. Consequently, we will abandon this approach in favor of one along the lines of file change semantics per Heim (1982).

To develop a more specific theory of discourse in HPSG, we will need to introduce some new machinery. The first steps are relatively simple: we need to change the sort hierarchy in such a way as to facilitate later principles. We will expand the partition of *sign* to include *sentences*, *themes*, and *rhemes* as subsorts, as we will want to talk in these terms hereafter. Our motivation for making *theme* and *rheme* separate subsorts is so that a sign can be both a *theme* and a *word*, a *rheme* and a *sentence* (as in the case of ‘zero-theme’ utterances), and so on, inheriting the relevant features from both types. Defining *sentence* in HPSG terms is relatively simple: a *sentence* is a VP with an empty subcategorization frame and a finite verb.

THE SORT HIERARCHY (REVISIONS AND ADDITIONS)

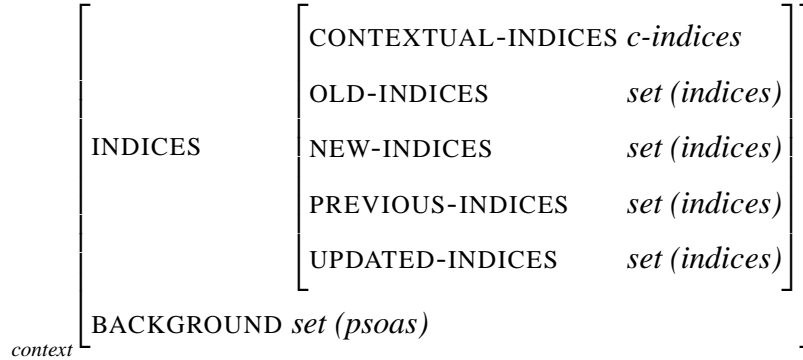
PARTITIONS:

Partition of *sign*: *word*, *phrase*, *sentence*, *theme*, *rheme*

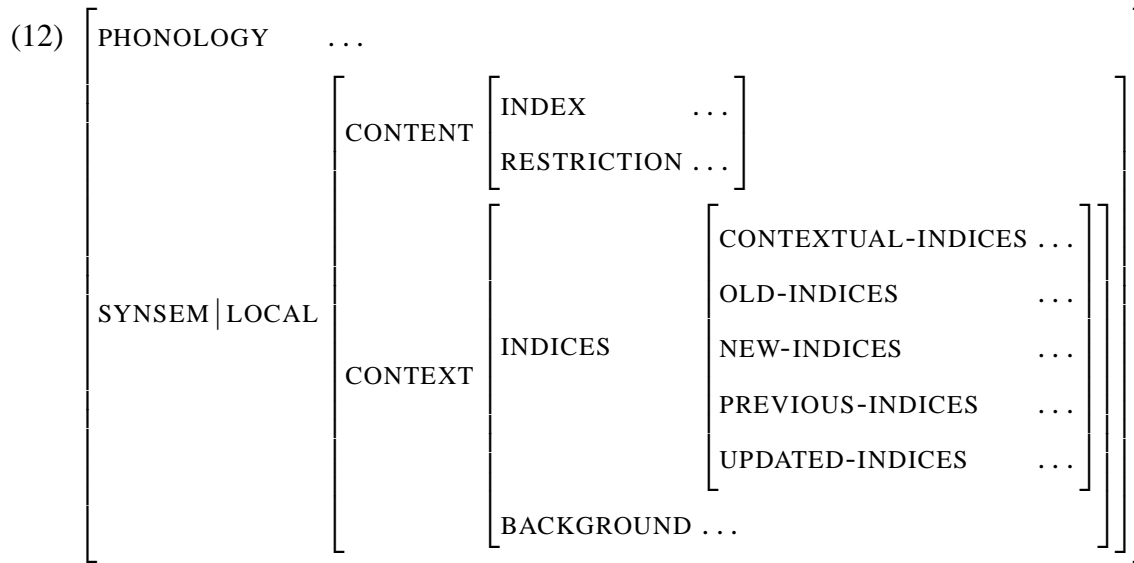
Partition of *discourse-sign*: *sentences*

Partition of *sentences*: *list-of-sentences*

FEATURE DECLARATIONS:



Changing the sort hierarchy as outlined above would make the relevant parts of an AVM look like that in (12):



Having added the necessary machinery to facilitate the encoding of new versus old information, we may now formalize our intuitive statement the theme must be old information as (13):

(13) **THE THEME PRINCIPLE:**

The $\text{SYNSEM|LOCAL|CONTENT|INDEX}$ value of a sign of sort *theme* must be token-identical to a member of the $\text{SYNSEM|LOCAL|CONTEXT|BACKGROUND|PREVIOUS-INDICES}$.

As yet, however, this principle has no teeth. We have yet to explain how file change is accomplished in HPSG terms, or how indices are handled in discourse. Our account should allow for the introduction of indices into the discourse that will be kept ‘on file’; rhemes will accomplish

such introduction, whereas themes will instead point to some index or indices already present within the discourse file. Towards that end, we will now introduce the Discourse Principle (14) and Conditions on Familiarity (15):

(14) **THE DISCOURSE PRINCIPLE:**

Let [sentences $\boxed{1}$] be of type *discourse-sign*. Then:

- (i.) For any S_i, S_{i+1} in $\boxed{1}$: the value of the SYNSEM|LOCAL|CONTEXT|INDICES| UPDATED-INDICES path for S_i is structure-shared with the value of the SYNSEM|LOCAL|CONTEXT|INDICES|PREVIOUS-INDICES path for S_{i+1} ;
- (ii.) For any S_i in $\boxed{1}$: the value of the SYNSEM|LOCAL|CONTEXT|INDICES| UPDATED-INDICES path for S_i is the union of the values of the SYNSEM|LOCAL|CONTEXT|INDICES| PREVIOUS-INDICES and SYNSEM|LOCAL|CONTEXT|INDICES| NEW-INDICES paths for S_i .

(15) **CONDITIONS ON FAMILIARITY:**

For any sign:

- (i.) The value of the SYNSEM|LOCAL|CONTEXT|INDICES| NEW-INDICES path \cap SYNSEM|LOCAL|CONTEXT|INDICES| PREVIOUS-INDICES path = \emptyset ;
- (ii.) The value of the SYNSEM|LOCAL|CONTEXT|INDICES| OLD-INDICES path must be a subset of the value of the SYNSEM|LOCAL|CONTEXT|INDICES| PREVIOUS-INDICES path.

The former principle will handle the management of discourse files while the latter will serve to formalize the old information/new information distinction.

The Discourse Principle places constraints on a discourse, which must consist of a set of sentences. Since these sentences must necessarily be ordered, the DP handles information exchange in terms of these ordered pairs (sentence S and sentence $S+1$). Part (a) of the DP states that the first sentence (S) shares its list of updated indices (the ‘file’) with the second sentence ($S+1$); these indices are incorporated as the list of previous indices for sentence $S+1$. Part (b) of the DP states that its updated indices list consists of the set of previously-introduced indices unified with the set of indices introduced by the sentence itself. This principle accounts for the ‘passing’ of information from sentence to sentence: each sentence builds on a file encoded as PREVIOUS-INDICES, which were taken from the previous sentence’s UPDATED-INDICES store. This sentence in turn consults its NEW-INDICES store and adds any contents to the file, storing the result as its UPDATED-INDICES

store. This serves as the basis for the next sentence’s PREVIOUS-INDICES, and so forth. Assuming an empty discourse file, partial AVMs of the first two sentences of a discourse might look like those in (16) and (17):

(16) *Initial sentence representation:*

$$\left[\begin{array}{l} \text{SYNSEM | LOCAL | CONTEXT | INDICES} \\ \text{OLD-INDICES } \{ \} \\ \text{NEW-INDICES } \boxed{1} \\ \text{PREVIOUS-INDICES } \boxed{2} \\ \text{UPDATED-INDICES } \boxed{3} = \boxed{1} \cup \boxed{2} \end{array} \right]$$

(17) *Subsequent sentence representation:*

$$\left[\begin{array}{l} \text{SYNSEM | LOCAL | CONTEXT | INDICES} \\ \text{OLD-INDICES } \boxed{4} \\ \text{NEW-INDICES } \boxed{5} \\ \text{PREVIOUS-INDICES } \boxed{3} \\ \text{UPDATED-INDICES } \boxed{6} = \boxed{5} \cup \boxed{3} \end{array} \right]$$

The first sentence has the empty set as its value for the SYNSEM|LOCAL|CONTEXT|INDICES| OLD-INDICES path. The Conditions on Familiarity require only that the OLD-INDICES be a subset of the PREVIOUS-INDICES; as the empty set is a subset of every set, this is acceptable. The first sentence combines its PREVIOUS-INDICES set ($\boxed{1}$) and the set of indices it itself introduces ($\boxed{2}$); these sets are combined as the UPDATED-INDICES ($\boxed{3}$). These updated indices are then structure-shared with the PREVIOUS-INDICES ($\boxed{3}$) of the next utterance, as shown in (17).

The Conditions on Familiarity (15) are very similar to Heim’s Extended Novelty-Familiarity Condition. The first condition requires that the new indices introduced by some signs not be members of the discourse file. The second condition states that old indices, if any, must be members of the discourse file. Unlike Heim’s NFC, however, no reference to indefinite or definite NPs has been made. We will return to this notion later.

4 The Theme Construction

Having formalized old and new information in terms of file change, we still need to insure that the theme will precede the rheme in our Russian examples. It is clear that the linear realization of words is somewhat free in Russian, and thus the phonology value of the phrasal sign may undergo the function ‘interleave-constituents’ in the sense of Pollard and Sag (1987) which “allows the

pieces of the phonology strings of daughter signs to be interleaved with those of other daughters in the construction of larger signs” (Pollard and Sag 1987:189). This function is proposed to be subject to language-specific Linear Precedence constraints, so we need to propose one to give us free word order but will maintain that the old information must precede the new information.

As HPSG makes use of constraints on the phonologies of constituents to predict word order, the word order of stylistically neutral Russian sentences can be captured with the following constraint:

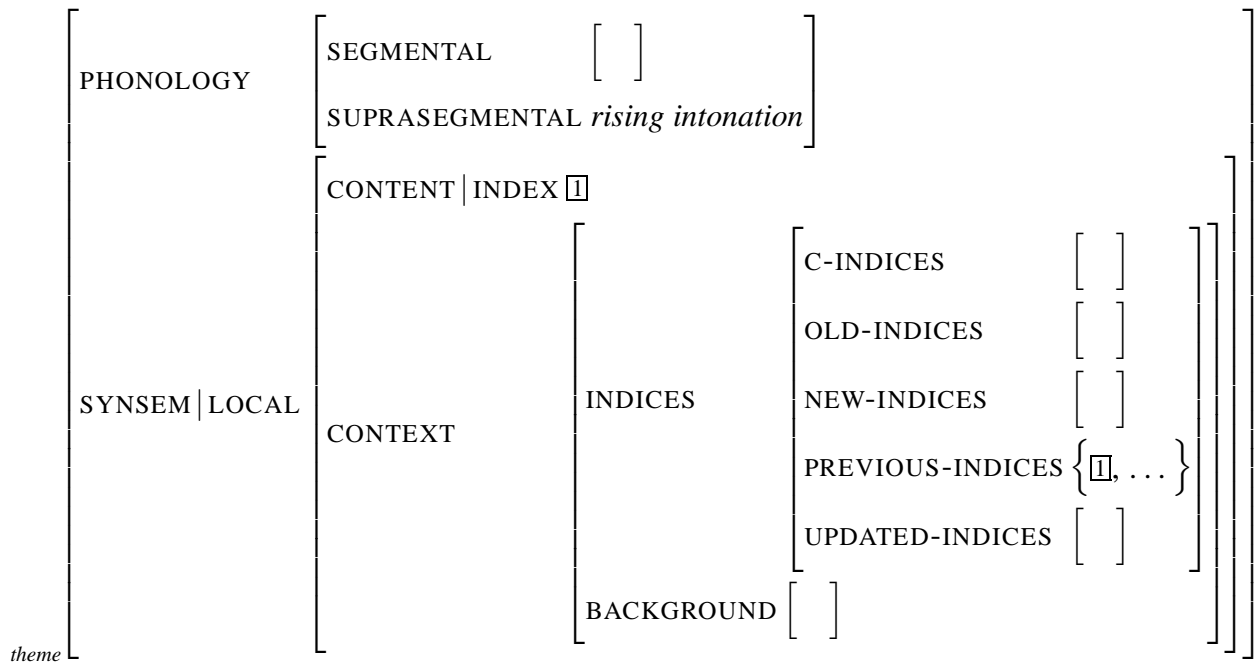
(18) **THEME/RHEME LINEAR PRECEDENCE CONSTRAINT:**

$$\text{sign}_{\text{theme}} \ll \text{sign}_{\text{rheme}}$$

This constraint requires signs of sort *theme* precede signs of sort *rheme*. By the Theme Principle, a sign of sort *theme* must have as a referent some member of the discourse file (PREVIOUS-INDICES). This constraint dictates that a phrasal sign which contains the old information, if present, must precede phrasal signs which contain the new information.

When the theme is uttered, it is uttered with a particular intonational pattern (this again helps to distinguish theme and rheme in ‘emotive’ speech, when our linear precedence constraint does not always apply). To arrive at the correct phonological representation, we need to somehow link the theme with the correct intonational pattern. We might do this by means of a construction like those of Construction Grammar, where a ‘theme construction’ unifies a sign of sort *theme* with the intonation usually associated with the theme. A hybrid HPSG-CG construction might look like this:

(19) **THE THEME CONSTRUCTION:**



Here the index of the theme is token-identical to some element previously introduced into the discourse file, and is marked for the proper intonational pattern. Other parts of the AVM would remain unspecified to allow unification with any other sign that met these criteria. Left for further specification and investigation is precisely how to formalize the phonological path of the AVM: just how is ‘rising intonation’ mapped onto the sentence? While certainly still sketchy, this construction begins to suggest how we capture the facts above. The rheme would not undergo such a construction and would not therefore be marked with the proper intonational pattern or point to some previously established element in the domain of discourse. By our linear precedence constraint, it would also then have to follow the theme in stylistically neutral examples.

5 Definiteness, Old Information, and Themes

Heim’s analysis of definite and indefinite NPs in terms of file change semantics reduced the distinction between the two types of NPs to old versus new information. Indefinites introduced new ‘file cards’ into the discourse, whereas definites needed to refer to old ‘file cards’. As this is precisely the sort of thing this paper looks to address in discussing themes and rhemes, it would be interesting to look at the distinction between them in terms of definiteness and indefiniteness. While Russian provides little insight here, lacking determiners like a or the, other languages with such determiners also have theme/rheme distinctions. For this reason, I will briefly examine themes in Arabic.

Arabic also has pragmatically determined themes which also typically occur sentence-initially in a language that is most often VSO. In the following Arabic examples, the theme is in bold type:

- (20) **Zaydun** ?abuu-hu mariidun
Zayd-NOM father-GEN-of-him ill-NOM
 ‘Zayd’s father is ill.’

- (21) **Zaydun** al-lisaanu ?atwalu min al-yadi.
Zayd-NOM the-tongue-NOM longer from the-hand-NOM
 ‘Zayd is a man of words rather than deeds.’

(Moutaouakil 1989:101)

As in the Russian examples, the Arabic theme “specifies the universe of discourse with respect to the subsequent predication is presented as relevant” (Dik 1978:19). Moutaouakil (1989) discusses at length the properties of the Arabic theme, comparing it with other constructions like those of

topic and focus. Those properties of the theme are much as those shown for Russian; he concludes as I have that the theme is a pragmatic function.

Interestingly enough, Arabic themes must be definite, as demonstrated by the following examples:

- (22) a. * *rajalun qara?tu kitaaba-hu.*
a-man-NOM read-1S book-ACC-of-him
 ‘A man - I read his book.’
- b. *ar-rajalu qara?tu kitaaba-hu.*
the-man-NOM read-1S book-ACC-of-him
 ‘The man - I read his book.’
- (23) a. * *fataatun kallamtu-haa.*
a-young-girl-NOM spoke-1S-to-her
 ‘A young girl - I spoke to her.’
- b. *al-fataatu kallamtu-haa.*
the-young-girl-NOM spoke-1S-to-her
 ‘The young girl - I spoke to her.’

(Moutaouakil 1989:103)

The contrast between the (a) and (b) examples would seem to suggest that definiteness is important because it establishes a link to the universe of discourse. In the terms outlined above, the definite theme refers to some entity (or index) already introduced. This is precisely Heim’s Novelty-Familiarity Condition:

(24) **THE EXTENDED NOVELTY-FAMILIARITY CONDITION:**

For [a logical form] to be felicitous w.r.t. [a file] F it is required for every NP_i in that:

- (i) if NP_i is [- definite], then $i \notin \text{Dom}(F)$;
- (ii) if NP_i is [+ definite], then
 - (a) $i \in \text{Dom}(F)$, and
 - (b) if NP_i is a formula, F entails NP_i .

(Heim 1982:369-70)

Whether it is the definiteness that determines novelty or vice-versa I will leave for further research. What is important is that familiarity, taken to be a necessary condition for being a theme, is encoded by definiteness in Arabic. This would seem to support both Heim’s theory and the framework outlined above.

6 Verbal Themes and Further Issues

As a final consideration in our analysis of theme constructions, I will turn to verbal themes. These are allowed by our sort hierarchy; we have made *theme* and *rheme* separate subsorts of type *sign*, so that we could have both phrasal and lexical themes and rhemes. However, the framework described thus far has talked in terms of indices, presumably introduced by NPs into the discourse. This leaves open the question of how to treat verbal themes like that in (26):

- (25) Lev Tolstoj – avtor romana “Vojna i mir.”
Lev Tolstoj author novel-GEN “War and Peace.”
‘Lev Tolstoj was (is) the author of the novel *War and Peace*.’

- (26) On napisal etu knjigu v 1869 godu.
he wrote that book-ACC in 1869 year
‘He wrote that book in 1869.’

Given solely sentence (25) as the previous discourse, the theme of (26) would be *on napisal etu knjigu* ‘he wrote that book’. Our model thus far can account for the establishment of the NPs *he* and *that book* as picking out old indices and thus being signs of sort *theme*. It is the verb *napisal* ‘wrote’ that is problematic.

Intuitively, it seems clear is that *napisal* is old news because we know Lev Tolstoj was the author of *War and Peace*. Being the author of a book means that one wrote it; the fact that we can **infer** from the statement ‘Lev Tolstoj was the author of *War and Peace*’ that ‘Lev Tolstoj **wrote** *War and Peace*’ seems to cause *napisal* in (26) to be encoded as old information.

To this end, we can modify our theory slightly to incorporate such facts. First of all, we probably want to ensure that the BACKGROUND psos are also stored in the discourse file. These would be roughly equivalent to the information stored on Heim’s file cards. We will accomplish in a manner similar to our INDICES-STORE. BACKGROUND psos will be collected and stored much as were indices. First, we will modify our feature declaration for sort *context*:

$$(27) \quad \left[\begin{array}{l} \text{INDICES} \\ \text{BACKGROUND} \end{array} \left[\begin{array}{l} \text{CONTEXTUAL-INDICES } c\text{-indices} \\ \text{OLD-INDICES } \quad \textit{set (indices)} \\ \text{NEW-INDICES } \quad \textit{set (indices)} \\ \text{PREVIOUS-INDICES } \quad \textit{set (indices)} \\ \text{UPDATED-INDICES } \quad \textit{set (indices)} \\ \text{PSOAS } \quad \textit{set(psoas)} \\ \text{PREVIOUS-PSOAS } \textit{set(psoas)} \\ \text{UPDATED-PSOAS } \textit{set(psoas)} \end{array} \right] \right]$$

context

It is quite possible that we would want to add additional paths after BACKGROUND to more closely mirror the attributes for INDICES. However, this should suffice for the present. We can now revise our Discourse Principle to account for parameterized-states-of-affairs:

(28) **THE DISCOURSE PRINCIPLE (REVISED):**

Let [sentences $\boxed{1}$] be of type *discourse-sign*. Then:

- (i.) For any S_i, S_{i+1} in $\boxed{1}$: the value of the SYNSEM|LOCAL|CONTEXT|INDICES| UPDATED-INDICES path for S_i is structure-shared with the value of the SYNSEM|LOCAL|CONTEXT|INDICES|PREVIOUS-INDICES path for S_{i+1} ;
- (ii.) For any S_i in $\boxed{1}$: the value of the SYNSEM|LOCAL|CONTEXT|INDICES| UPDATED-INDICES path for S_i is the union of the values of the SYNSEM|LOCAL|CONTEXT|INDICES|PREVIOUS-INDICES and SYNSEM|LOCAL|CONTEXT|INDICES| NEW-INDICES paths for S_i ;
- (iii.) For any S_i in $\boxed{1}$: the value of the SYNSEM|LOCAL|CONTEXT|BACKGROUND| UPDATED-PSOAS path for S_i is the union of the values of the SYNSEM|LOCAL|CONTEXT|BACKGROUND|PREVIOUS-PSOAS and SYNSEM|LOCAL|CONTEXT|BACKGROUND| PSOAS paths for S_i .

The revised DP now allows us to pass both psogas and indices from sentence to sentence, maintaining our ‘file’. We may require themes to either have an index already present in the file, or restrict themes to describing a subset of psogas in the file.

(29) **THE THEME PRINCIPLE (REVISED):**

A sign of sort *theme* must:

- (i.) Have a SYNSEM|LOCAL|CONTENT| INDEX value token-identical to a member of the SYNSEM|LOCAL|CONTEXT|BACKGROUND| PREVIOUS-INDICES value; or
- (ii.) Have a SYNSEM|LOCAL|CONTEXT|BACKGROUND-PSOAS value token-identical to a member of the SYNSEM|LOCAL|CONTEXT|BACKGROUND| PREVIOUS-PSOAS value.

As of yet we have not accounted for the phenomena in examples (25) and (26). We need simply make a statement much like that of (b) of the second part of Heim’s Novelty-Familiarity Condition. Let us introduce the following additional principle:

(30) THE FAMILIARITY INFERENCE PRINCIPLE:

A sign whose *context* values are such that the value of the SYNSEM|LOCAL|CONTEXT|BACKGROUND|PREVIOUS-PSOAS path entails the value of the SYNSEM|LOCAL|CONTEXT|BACKGROUND-PSOAS path is of sort *theme*.

This principle will allow us to account for (25) and (26): as the psoas in the discourse file after (25) was uttered will include ‘Lev Tolstoj is the author of *War and Peace*’, and as *author(x,y)* entails *wrote(x,y)*, the verb *napisal* ‘wrote’ is considered old information and must therefore be of sort *theme* by principle (30).

7 Final Remarks

This paper has only begun to address the questions that a satisfactory analysis of theme and rheme in languages like Russian and Arabic would have to answer. First, a theory like HPSG would need a much better theory of discourse; the groundwork for such a theory has been presented here. Without such a theory, there can be no clear basis for unification of the theme and an already-established element of the discourse. This theme then needs to be assigned the proper intonational pattern; this might be achieved via some sort of construction like that outlined above. Finally, without a more detailed explanation of how the ‘interleave-constituents’ function might work, it is unclear what mechanics are involved that keep the theme and rheme in their proper order but still manage to keep other elements in their ‘usual’ order, other than to suggest, as I have above, that the Theme-Rheme Linear Precedence Constraint serves to further constrain signs based on the pragmatic notion of old versus new information. Evidence from Arabic has been shown to support the notion that themes must be definite, old information, and further suggestions show verbal themes are readily accounted for after slightly modifying the framework presented, and by using entailment relations between parametrized states of affairs.

References

- Dik, S.C. (1978) *Functional Grammar*. Amsterdam: North-Holland. [Third printing (1981), Dordrecht: Foris.]
- Heim, Irene. (1982) *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts at Amherst.
- Krylova, Olga, and Serafima Khavronina. (1988) *Word Order in Russian*. Moscow: Russky Yazyk.
- Moutaouakil, Ahmed. (1989) *Pragmatic Functions in a Functional Grammar of Arabic*. Dordrecht: Foris.
- Pollard, Carl, and Ivan A. Sag. (1987) *Information-based Syntax and Semantics, Vol. 1*. Stanford University: CSLI Publications. [Distributed by University of Chicago Press.]
- Pollard, Carl, and Ivan A. Sag. (1994) *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.