

# Topic-Focus Subjectivity within Verbal Complex Substructures on the Romanian Syntax-Prosody Interface

Diana Trandabăț

Mihai Alex Moruz

Iustin Dornescu

dtrandabat@iit.tuiasi.ro mmoruz@iit.tuiasi.ro idornescu@iit.tuiasi.ro

Neculai Curteanu

Cecilia Bolea

curteanu@iit.tuiasi.ro

cbolea@iit.tuiasi.ro

Institute for Computer Science, Romanian Academy, Iași Branch

Corresponding author: Diana Trandabăț dtrandabat@iit.tuiasi.ro

## Abstract

The main contributions enclosed within this paper refer to: (i) A general view on the syntax-prosody interface, the non-isomorphism between the two facets of the language being conjectured as a proper subsumption mapping between syntax and prosody; (ii) Implementation of the Topic-Focus Articulation (TFA) algorithm for the Romanian sentence, with novelties on Intonational TFA and interrogative sentences.; (iii) Devising the verbal group (VG, *viz.* verbal complex) into verbal subgroup syntactic structures; (iv) TFA algorithm refinement on complex Romanian VGs; (v) Subjectivity / Objectivity discovery based on TFA and communicative dynamism disordering, with consequences for text analysis and prosody design in e-learning systems.

## 1 Introduction

This paper has two main roots: results of the Prague School on the Topic-Focus Articulation (TFA), which confronts the *subjectivity* bias of *communicative dynamism* (CD) to the *systemic ordering* (SO) within local text and speech (Hajicova *et al.*, 1995; Hajicova, 2006). The second starting point is the analysis of the *verbal complex* (Monachesi, 2005) (denoted hereafter as *verbal group*, VG) and its interfaces, especially the syntax interface for the Romanian VG and the phonology interface for the Italian restructuring verbs. The paper benefits from results on functional VG FX-bar projection and restructuring as Verbal SubGroups (VSGs), and experience with TFA and *intonational* TFA im-

plementation for Romanian (Curteanu *et al.*, 2007a; 2007b).

The purpose of the present paper is to support the idea that changing the argument SO by a certain CD within a clause, and (de)composing various VSGs (such as auxiliaries, restructuring verbs) within VG may create specific subjectivities and (classes of) emotional states. These can be automatically recognized in (e-learning) text analysis or planned (presupposed) within prosody design for text-to-speech systems.

The main results enclosed within this paper are: (i) A general view on the syntax-prosody interface, the non-isomorphism between the two facets of the language (Monachesi, 2005) being conjectured as a proper subsumption mapping between syntax and prosody; (ii) Implementation of the TFA algorithm for the Romanian sentence, with novelties on Intonational TFA and interrogative sentences.; (iii) (De)composing the VG into VSG (Verbal SubGroup) syntactic substructures; (iv) TFA algorithm refinement on complex Romanian VG; (v) Subjectivity / Objectivity discovery based on TFA and CD disordering, with consequences for text analysis and prosody design in e-learning systems.

## 2 A General Subsumption Mapping on the Syntax–Prosody Interface

(Curteanu *et al.*, 2007b) outlines the image of a general parsing strategy, in parallel for both text and speech on the syntax-prosody interface, consisting of three main components: a *lattice*-type organization of textual *marker* (boundary) *classes*, with the role of syntactic structure delimiting and dependency-establishing (for the IS-driven *syntactic side* of the interface); a similar set of marker classes for the *prosody side*; a *homomorphic* gen-

eral *mapping* between the two main structures with the property of preserving *subsumption* between the involved (sub)structures: the IS-driven *syntax subsumes* the *prosody*. The SCD marker classes and FX-bar projection structures (Curteanu, 2006c), (Curteanu and Trandabăt, 2006b) are concrete steps on this road, taken for both Romanian and English.

We completely agree with the remarks in (Monachesi, 2005) on the “non-isomorphism between prosodic and syntactic constituents”, but we support an operator such as Ghini’s Mapping Algorithm (1993) as a homomorphism of syntactic marker classes on prosodic ones, phonologically-driven, with the main property of subsumption between syntax and phonology-prosody.

### 3 TFA / ITFA

TFA theory and algorithm is a procedure intervening exactly on the core of syntax-prosody interface, with semantic, discursive, and pragmatic surgical tools, bringing clearer rules for establishing the accentuated-deaccentuated IS components on the prosodic intonational units of a sentence, for assigning proper intonational stress (ToBI labels). Finding the set(s) of rules by which one could evaluate *the focused* elements of a spoken utterance is *one* of the keys of the *prosody prediction*.

#### 3.1 Ghini’s Algorithm

In the process of assigning intonational units to syntactic structures, Ghini’s (1993) *mapping algorithm* seems to behave better than currently existing algorithms when applied on complex VGs, with VSGs having restructuring verbs as their heads (Monachesi, 2005). This is mainly due to the fact that, in Ghini’s mapping algorithm, the phonologic phrase constituency is driven by *phonological-prosodic principles* rather than syntactic ones.

Ghini’s mapping algorithm consists in adding to the right-hand side boundary of phonological markers the following *prosodic principles*, which are used to determine the correct phonological phrase constituency within the specific domain formation: A. *Uniformity and average weight*; B. *Symmetry*; C. *Increasing units*. Applying these phonological principles, Monachesi (2005) shows that Ghini’s mapping algorithm produces appropriate phonological phrases for complex VGs having

restructuring verbs as their heads and in which the clitic climbing phenomenon occurs.

#### 3.2 Intonational Unit Acquisition from FDG-Trees

Using the functional-dependency (FD) parsing tree of a sentence, we determine the intonational phrase boundaries, by applying the following empirically acquired rules:

- **Clause acquisition:** every subtree that has as its root a verb forms a clause, together with all of its descendents, with the exception of the verbal descendents which are themselves heads of verbal groups, and their subtrees;
- **Main clause acquisition:** all verbs on the highest level in the FD tree of a sentence are heads of main clauses;
- **IP (Intonational Phrase) acquisition:** each main clause, together with all its subordinate clauses, forms an IP;
- **ip (intermediate phrase) acquisition:** each syntactic clause is an *ip*. There are cases when *ip* structures are interrupted by speech break markers (punctuation marks, coordinative conjunctions, etc.)
- **RU (Rhythmical Unit) acquisition:** (i) a parent and its first successor form, together, an RU (regardless whether the successor stands to the left or the right of the parent). The RU also contains the first adjacent descendent of the successor, recursively, to the leaf; (ii) Non-adjacent successors are each in separate RUs, in a similar way as in (i); (iii) If a parent has two successors adjacent with it, the first successor forms an RU with its parent and the other successor forms an RU by itself.

An example of FD-based acquisition of intonational units is:

```
[[[O'Brien apuc•]RU [sticla de
gât]RU]IP [[•i turn• în pahare]RU
[un lichid]RU [de un ro•u
aprints]RU]IP]IP.
```

(En: O'Brien gripped the bottle neck and poured a bright red liquid into glasses.)

### 3.3 Topic-Focus computing for Romanian

We attempt to design an adequate syntax-prosody interface for Romanian by adapting and applying the Prague School’s TFA (Topic-Focus

to IS-based intonational units. Applying transformation rules as those proposed by Steedman’s CCG for the local level or those drafted by (Tseng et al.) for global level, the corresponding ToBI labels could be automatically established with better adequacy.

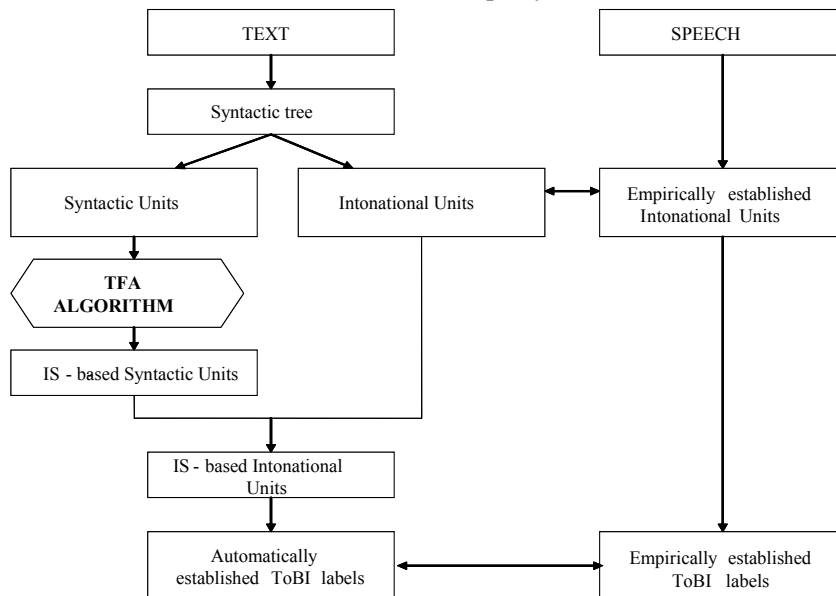


Figure 1. The architecture of the prosody prediction system

Articulation) algorithm (Hajcova et al. 1995). The purpose of this algorithm is to compute the topic-focus (theme-rheme) for a given sentence and obtain, on the basis of sentence IS elements, a better acquisition of the sentence intonational units, and subsequently, an improved naturalness for the assignment of tone and tune phrases to the established intonational units of the sentence.

The starting point of our prosody prediction system (Figure 1) is the morphologically-syntactically tagged text, with SCD markers and FD dependencies. From the FD parsing tree, two directions are derived: **(1)** grouping the FD branches into syntactical constituents, and **(2)** detecting intonational units from the FD tree using the acquisition rules presented above.

The *Topic-Focus Articulation (TFA) Algorithm* receives as input the syntactic units of a sentence, SCD markers (Curteanu 2006c) (such as the definite or indefinite article), and certain semantic features. The outcome consists of IS-based units, which are actually syntactic constituents with topic-focus values. The obtained syntactic IS is used to re-arrange the intonational units according to their relevance for the speech dynamics, leading

The right-hand side of Fig. 1 presents the empirical annotations of the spoken text. The manually determined intonational units and the ToBI labels will be compared to the automatically acquired ones in order to evaluate and adjust the TFA procedure against the (currently empirical) annotated gold corpus.

### 3.4 The TFA Algorithm

The input for the TFA algorithm consists of FD trees for Romanian sentences. Besides the morphological annotation of each word, the *semantic features* of constituents are also classified according to their *specificity degrees*: (i) *general* – low specificity, contextually non-bound; (ii) *specific* – high specificity, contextually non-bound; (iii) *indexical* – mid-specificity, contextually-bound. Examples of specificity degrees for *temporal complements* are: general (“*niodată*”, “*mereu*”, etc), indexical (“*astăzi*”, “*anul acesta*”), specific (“*22 iunie*”, “*într-o frumoasă zi de mai*”). The TFA algorithm utilizes the specificity degrees only for verbs and *temporal / locative* complements. The position of the verbal direct and indirect arguments within the

systemic order (SO) is important when computing the sentence *topic* and *focus*. According to (Hajcova et al. 1995), the SO for the main kind of complementations in English is:

Time - Actor - Addressee - Objective - Origin - Effect - Manner - Directional.*from* - Means - Directional.*to* - Locative

The SO that we shall use for Romanian within the TFA algorithm is similar, being supported by the observation that, in clauses with 3-valenced verbs, the indirect object usually stands *before* the direct object for the purpose of avoiding the possessive ambiguity (I-am dat cartea Ioanei may be interpreted as Ioana being the receiver or the possessor of the book).

Figure 2 presents an example of the algorithm input, where the parsing tree is described in a LISP-style syntax and the output of the procedure consists of the topic-focus assessment for the verb and all its complementations.

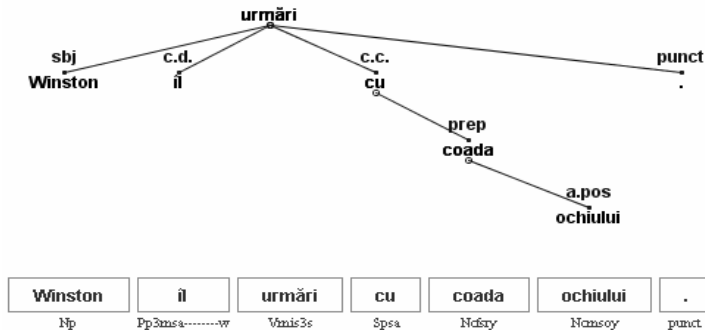


Figure 2a. FD parsing tree for the sentence “Winston îl urmări cu coada ochiului.” (En: Winston followed him from the back of the eye.)

```
verb (topic(f), sem(interm), label(il urmari), ltree(act(topic(f), det(1), so(1), surf(np), label (Winston))), rtree(mod(topic(f), det(1), so(8), surf(np), label (coada), ltree(prepcu), rtree(apos(ochiului))))))
```

Figure 2b. TFA algorithm input for the sentence “Winston îl urmări cu coada ochiului.”

```
([[[Winston(t), îl urmări(t/f)]RU, [cu coada ochiului (t(f))]RU]IP]IP)
```

Figure 2c. TFA algorithm output for the sentence “Winston îl urmări cu coada ochiului.”

In order to better determine Topic-Focus for Romanian interrogative questions, the original

TFA algorithm for affirmative sentences was modified, and is presented below. Let **S** be the set containing the Verbal Group (VG) (Monachesi 2005), (Curteanu et al. 2006a) and all its complements from a sentence **s**. Let **f** denote the *focus*, **t** - the *topic*, and **t/f** denote ambiguous topic-focus elements; **t(f)** is a special kind of ambiguity (this element is **f** only when there is no other **f** in the reading of the sentence). **S(i)** denotes the *i*-th complementation in the surface order, **verb(S)** is the VG, and **last(S)** is the last element in the surface word order. The **Interrogative TFA Algorithm** is exposed below:

```
1. f = S;
   /*all elements are presumed to be focus*/
2. if (first(S)=verb(S))
   2.1. if (subject(S) is definite)
       f=f-{subject(S)};
       t=t∪{subject(S)};
       f=f∪{verb(S)};
       goto 3.2;
   2.2. else if (subject(S) is indefinite)
       f=f-{verb(S)}; t=t∪{verb(S)};
       f=f∪{subject(S)};
       goto 3.2;
   2.3. else
       f=f∪first(S);
       f=f-{verb(S)}; t=t∪verb(S);
       goto 3.2.
3. else
   3.1. if (verb(S) is general)
       f=f-{verb(S)};
       t=t∪{verb(S)};
       else if (verb(S) is specific)
       f=f∪{verb(S)};
       else
       f=f-{verb(S)};
       t/f=t/f∪{verb(S)};
   3.2. for (i=pos (verb(S)) to
         pos (last(S)))
       /*for all complements behind the verb*/
       if (S(i) is touched by 2)
           continue;
       if ((S(i) is indefinite subject)
           or (S(i) is specific local complement))
       /*specific means it is neither general, nor indexical*/
           f=f-{S(i)}; t/f=t/f∪{S(i)};
       else
           f=f-{S(i)}; t=t∪{S(i)};
   3.3. /*to the left of the verb*/
   i.if (pos (first(S))=pos (verb(S))-1 and
        (first(S) is definite noun or personal
         pronoun))
       /*the verb has only one complement to the left, and that complement is an definite noun or a personal pronoun*/
       f=f-{last(S)};
       t/f=t/f∪{last(S)};
   ii.if ((first(S) is temp or loc) and
        first(S) is specific)
       f=f∪{last(S)};
       else if ((first(S) is temp or loc) and
        first(S) is not specific)
```

```

f=f-{last(S)};
t=t∪{last(S)};
last(S) is moved to the right of the
verb in the CD
iii.let (S(i),S(j)) be the leftmost
pair that fails to follow SO
for(a=i to pos(verb(S))-1)
f=f-{S(i)}; t=t∪{S(i)};
for(a=1 i-1)
f=f-{S(i)};
t/f=t/f∪{S(i)};
f=f∪{first(S)};
iv.if(ii not met and iii not met and
first(S) is indefinite)
f=f∪{last(S)};
v.for all S(i) where i>pos(verb(S)) and
S(i) not touched by i, ii, iii, iv
f=f-{S(i)}; t/f=t/f∪{S(i)};
4. if all x in S-{verb(S)} are either t or
t/f
4.1. if((verb(S) is t/f) and (first(S)
is definite or indexical or personal pro-
noun))
f=f-{first(S)}; t(f)=first(S);
4.2. else
f=f-{verb(S), first(S)};
t/f=t/f∪{verb(S), first(S)};
5. if at this step, the sentence still con-
tains no focus, then it does not bring any new
information (it is only a comment on
old information).

```

The workings of the TFA algorithm (both the affirmative and interrogative ones) are described in the examples below, which analyze the sentence “Ion ar fi vrut să fi fost dus la muzeu.” (En: John would have wanted to be taken to the museum.)

Ion(t), ar fi vrut să fi fost  
dus(t/f), la muzeu(f)

Fig. 3a Affirmative TFA algorithm output

Ion(t/f), ar fi vrut să fi  
fost dus(t/f), la muzeu(f)

Fig. 3b Interrogative TFA algorithm output

Ar fi vrut(t/f), Ion(t/f) să  
fi fost dus(t/f), la muzeu(f)

Fig. 3c Affirmative TFA algorithm output

As one can see, there are differences in the computed topic-focus values for the sentences analyzed in Fig 3a and 3b, and this is due to the differences between the TFA algorithms for affirmative and interrogative questions. The sentence in Fig 3c is a different reading for the same sentence, but, in this case, the verbal group has been split in two by the subject. For the purposes of the algorithm, the part of the verbal group which does not contain the predicational verb is considered a special type of complement, which can never break SO.

After computing topic/focus for the syntactic constituents, the natural step to follow is to assign topic/focus values to the intonational units computed from the FDG tree. We have called this process the Intonational TFA Algorithm (ITFA), since we conceived it as an extension of the Prague School’s TFA. The output of the TFA algorithm is postprocessed in order to define the  $t(f)$  value: if the clause contains an IS unit  $a$  labeled as  $t(f)$ , and no other IS unit in that clause is labeled as bearing a focus  $f$ ; in all other instances, the IS unit labeled as  $t(f)$  will be labeled as  $t/f$ . This processing help defining the ITFA algorithm for merging intonational units and IS based syntactic units.

```

if (IU unit == IS unit)
#the IU and the IS have exactly the same elements
topic/focus(IU unit) := topic/focus(IS
unit)
else if (IU unit contains more than one IS
unit)
if (all are t/f)
topic/focus(IU unit) := t/f
else
count occurrences of topic and focus re-
spectively
#an IS having t/f yields a topic and a focus values
if (dominant topic/focus can be deter-
mined)
topic/focus(IU unit) := dominant
topic/focus
else
#dominant topic/focus cannot be determined, for instance if we
have one focus and one topic
topic/focus(IU unit) := t(f)
else if (IS unit contains more than one IU
unit)
headIU := the IU containing the head of
the IS unit
topic/focus(headIU) := topic/focus(IS
unit)
topic/focus(rest of IUs) := t

```

The intonational units extracted from the sentence “Winston îl urmări cu coada ochiului.” is:

[[[Winston îl urm•ri]<sub>RU</sub> [cu coada ochiu-  
lui]<sub>RU</sub>]<sub>IP</sub>]<sub>IP</sub>.

When applying the ITFA Algorithm, we obtain the following distribution:

[Winston îl urm•ri]<sub>T</sub> [cu coada ochiu-  
lui]<sub>T/f</sub>.

One can see that the final rhythmic unit has changed its value from  $t(f)$  to  $t/f$  due to the post-processing of the TFA Algorithm. Also, the RU “Winston îl urmări” has become topic since, when computing the number of topics and foci from the IS units, we had 2 topics and one focus.

#### 4 Substructures of the Romanian Verbal Predicate

We investigated the syntactic / semantic substructures of the Romanian *verbal group* (VG), or *verbal complex* (Monachesi, 2005), starting from the instruments and arguments in the literature, and melted into the device of (direct and inverse) FX-bar projections and theory (Curteanu et al., 2006a), (Curteanu and Trandabat, 2006b), (Curteanu, 2006c).

The *verbal group* (VG) can be further divided into several stable syntactic substructures (subgroups): the Tense Auxiliary SubGroup (TASG), the Passive Tense Auxiliary SubGroup (PTASG), the Modal Verb Subgroup (ModVSG).

The most frequent verbal subgroup (VSG) is the **Tense Auxiliary SubGroup (TASG)**. This subgroup accepts, beside the tense auxiliary, the VG negation and special adverbs (*mai, cam, prea, și, tot*), for example:

Ion [[va mai]<sub>TASG</sub> pleca]<sub>VG</sub> ?

(En: Will John leave?)

[[Nu prea aș mai fi]<sub>TASG</sub> cântat]<sub>VG</sub>.

(En: I would not have sung.)

[[Nu ar mai fi]<sub>TASG1</sub> trebuie [[să îl]<sub>TASG2</sub> plătești]<sub>VG1</sub>]<sub>VG2</sub>.

(En: You would not have had to have paid it.)

The analysis of the linguistic data could suggest two forms of the TASG: a *basic* form and an *applied* form. The basic form of the TASG is represented by all possible tense auxiliaries, in all inflected and personal forms, with or without negation and special adverbs, such as *sunt, nu mai sunt, voi fi, prea vei fi, aș fi, nu ar mai fi, am fi*, etc. The VG applied (or *instanced*) form represents the mapping of the *basic* TASG to the VG *semantic head* (usually, the predicational verb). This operation entails the applied TASG to be enriched with *object (pronominal) clitics*. TASG is a non-saturated VSG that needs as semantic head a noun (N), adjective (Adj), or non-finite verb (V) form. For V finite forms, TASG is considered to be enclosed, by default, within the synthetic inflection of the verb.

The *second* VSG is the **Passive Tense Auxiliary SG (PTASG)**. PTASGs are non-saturated VSGs that need a non-finite verb as semantic head to complete its (copular) meaning.

[[[Am]<sub>TASG</sub> fost]<sub>PTASG</sub> gonit]<sub>VG</sub>.

(En: I was banished.)

Premiul [[[îi va]<sub>TASG</sub> fi]<sub>PTASG</sub> dat]<sub>VG</sub> lui Ion.

(The prize will be given to John.)

Special attention must be granted to the PTASG when the passive semantic diathesis has the appearance of a reflexive syntactic diathesis, as in:

Ușa [[s-a]<sub>TASG</sub> deschis]<sub>VG</sub>.

(En: The door opened.)

The correct marking of the VSGs in this example would be:

Ușa [[[s-a]<sub>TASG</sub>]<sub>PTASG</sub> deschis]<sub>VG</sub>.

Another VSG is the modal one. The **modal verbal subgroup (ModVSG)** is a TASG derivation, having as semantic head a modal verb (*a trebui* – en: must, *a putea* – en: can). ModVSG accepts special adverbs and negations, and may contain, in its applied form, object clitics, if its argument is not a clause.

Nu prea aș fi putut să i-o mai fi împrumutat

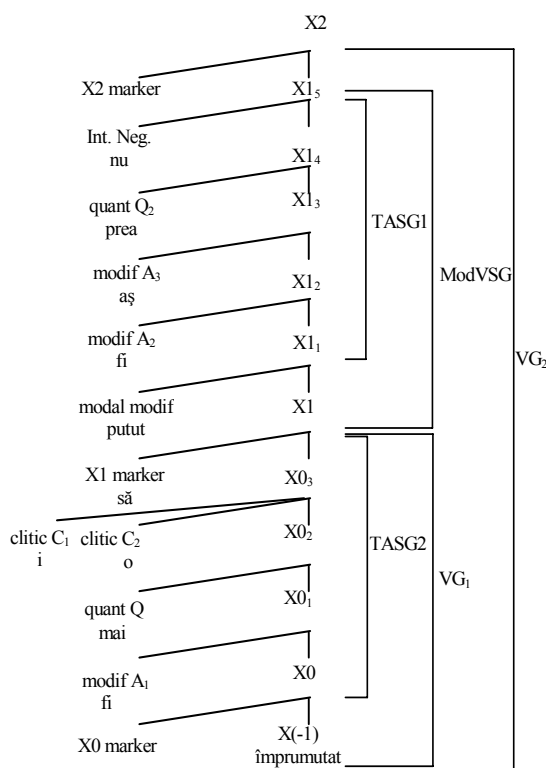


Figure 4. Example of VSGs

[[Nu trebuie]<sub>ModVSG</sub> [să i-o mai dai]<sub>VG1</sub>]<sub>VG2</sub>.

(En: You must not give it to her.)

[[Nu prea aș fi putut]<sub>ModVSG</sub> [să i-o mai fi împrumutat]<sub>VG1</sub>]<sub>VG2</sub>.

(En: I could not have really lent it to her.)

When the argument of ModVSG is a clause, the pronominal clitics are naturally embodied into the finite VG argument of the ModVSG, and ModVSG may receive at most the reflexive (or impersonal) clitics:

[[Nu se poate]<sub>ModVSG</sub> [să i-o dau]<sub>VG1</sub>]<sub>VG2</sub>.

(En: I can't give it to her.)

Ordering those subgroups, we obtain the following VSGs, for a sentence like:

[[[Nu prea aş fi]<sub>TASG1</sub> putut]<sub>ModVSG</sub> [[să i-o mai fi]<sub>TASG2</sub> împrumutat]<sub>VG1</sub>]<sub>VG2</sub>.

The linear construction for the sentence above is presented in Figure 4.

## 5 IS-Driven VG Mapping on the Syntax-Prosody Interface

The TFA algorithm considers the VG as a verbal complex, not as a sum of VSGs. Therefore, the whole VG receives a single *topic*, *focus* or *ambiguous* value.

Frăția(f) [i se spune](t).

(En: It is called The Brotherhood.)

Winston(t) [îl urmări](t/f) cu coada ochiului t(f).

However, in the case of complex VGs, a breath break can interrupt the verbal structure. We noticed that this pause never breaks a VSG; instead, it joins one or more VSGs into a rhythmic unit (though the ITFA algorithm presented above). Several examples are:

[Nu prea aş fi putut]<sub>RU1</sub> [să i-o mai fi împrumutat]<sub>RU1</sub>.

Ion [trebuie]<sub>RU1</sub> [să înceapă]<sub>RU2</sub> [să scrie]<sub>RU3</sub> deja scrisoarea.

(En: John must already start to write the letter.)

In those cases, the topic-focus values of the VG are transferred to all its compounding RUs.

### 5.1 Subjectivity in VSGs

The standard pronunciation assigns to each RU from the VG the topic/focus value obtained through the TFA algorithm for the VG. The topic/focus value for each RU in the VG can be modified if applying scenario indications.

When trying to focus a VG-compounding RU, the following elements can be stressed:

- the semantic head (it usually carries accent in normal pronunciation).
- the modal/passive verbs

- the special adverbs
- the negation.

Thus, the RUs of the VG from the sentence:

Nu ar fi trebuit să-i fi fost dată cartea.

(En: The book should not have been given to him.) can be focused in the following manner:

[Nu ar fi trebuit] [să-i fi fost dată] [cartea].

[Nu ar fi trebuit] [să-i fi fost dată] [cartea].

[Nu ar fi trebuit] [să-i fi fost dată] [cartea].

[Nu ar fi trebuit] [să-i fi fost dată] [cartea].

The first two cases, which stress the first RU, indicate a personal implication of the speaker towards the modality expressed (thus a subjectivity mark). The last two examples place the focus on the informational content rather than on the subjective one. A similar situation can be created for the other modal verb “a putea”:

[Puteai] [să i-o dai] (En: You could have given it to him.) # but you did not want it.

[Puteai] [să i-o dai] # but there was no restriction in giving Y to Z

The accentuation of the first RU indicates the modality, the subjective opinion of the speaker, while the accentuation of the second RU indicates the objective implication.

### 5.2 Interrupted VGs

The VG, contiguous in affirmative pronunciation, can be split by different arguments/adjuncts if in an interrogation. For instance, the sentence:

[Ion]<sub>t</sub> [va putea]<sub>t/f</sub> [să scrie]<sub>t/f</sub> [din nou]<sub>t</sub>.

(En: John will be able to write again.)

can be split by the subject or the temporal complement as follows:

[Va putea]<sub>t/f</sub> [Ion]<sub>t/f</sub> [să scrie]<sub>t/f</sub> [din nou]<sub>t</sub>?

[Va putea]<sub>f</sub> [să scrie]<sub>f</sub> [Ion]<sub>t</sub> [din nou]<sub>t</sub>?

[Va putea]<sub>f</sub> [să scrie]<sub>f</sub> [din nou]<sub>t</sub> [Ion]<sub>t</sub>?

[Va putea]<sub>f</sub> [Ion]<sub>t</sub> [din nou]<sub>t</sub> [să scrie]<sub>t/f</sub>?

[Ion]<sub>t</sub> [va putea]<sub>t</sub> [din nou]<sub>t</sub> [să scrie]<sub>t/f</sub>?

By breaking the VG, the resulting VSGs receive different topic/focus values, used to accentuate

differently the pronunciations, thus to indicate the subjectivity that can be extracted from the text.

## 6 Discussions and Conclusions

After briefly describing the subsumption mapping between syntax and prosody, this paper has presented the implementation of the Topic-Focus Articulation algorithm for Romanian sentence, with novelties on Intonational TFA and interrogative sentences. The Romanian verbal group division into the tense, modal and passive subgroups was also addressed, since we intended to apply the TFA algorithm for complex verbal groups (viz. composed of multiple subgroups). The major discovery of our paper is that the verbal subgroups are differently articulated in order to account for the subjectivity / objectivity and communicative dynamism disordering, with consequences for text analysis and prosody design in spoken e-learning systems.

A possible application of determining the subjectivity/ objectivity of Romanian VSGs in eLearning is the acquiring and improvement of the Romanian language by non-native speakers. An application for learning the manner in which subjectivity/ objectivity can be expressed in Romanian VGs can be designed as follows: for a VG “standard” shape, *syntactic variations* may be applied, with their modified meanings attached. For instance, we can describe the differences between: “am mers” (en: “I went”), “am mai mers” (EN: I already went) and “am tot mers” (EN: I kept on going), which are all subjective variations on the VG “I went”, for a better understanding of the Romanian language. Another example is represented by the syntactic (classical syntax and topic-focus syntax) and the corresponding semantic variants and their sibling meanings in §5.2.

Another possible application of the objectivity/subjectivity detection using the TFA algorithm in eLearning considers the ability of a text to be considered a *learning material*. We intend to test the hypothesis that a subjective text (i.e. communicative dynamics disordering) has more attraction for the learners than an objective one. The text subjectivity/objectivity is computed by taking into account the differences between the systemic order and the communicative dynamics for each sentence of the text.

## References

- Eva Hajičová, H. Skoumalova, and P. Sgall. 1995. *An Automatic Procedure for Topic-Focus Identification*, Computational Linguistics, 21 (1): 81-94.
- Eva Hajičová. 2006. *Old Linguists Never Die, They Only Get Obligarily Delete*. Computational Linguistics, 43 (4):457-469.
- Mark Steedman. 2000. *Information structure and the Syntax-Phonology Interface*. Linguistic Inquiry, MIT Press, Cambridge, (34): 649-689.
- N. Curteanu, D. Trandabăț, and M. Moruz. 2006a. Substructures of the (Romanian) Predicate and Predication using FX-bar Projection Functions on the Syntactic Interface, ECIT 06, Iasi, Romania
- Neculai Curteanu and Diana Trandabăț. 2006b. *Functional (F)X-bar Projections for Local and Global Text Structures. The Anatomy of Predication*. Revue Roumaine de Linguistique, Romanian Academy Editorial House, Bucharest, No. 4 (to be published).
- N. Curteanu, D. Trandabăț, and M.A. Moruz. 2007a. *Syntax-Prosody Interface for Romanian within Information Structure Theories*. (C. Burileanu and H.N. Teodorescu Eds.) Proceedings of SpeD-2007 Conference (Speech Technology and Human - Computer Dialogue), Romanian Academy Editorial House, Bucharest, pp. 217-228.
- N. Curteanu, D. Trandabăț, and M.A. Moruz. 2007b. *Syntax-Prosody Interface for Romanian within Information Structure Theories*. Proceedings of TSD-2007 (Text, Speech and Dialogue) Conference, Plzen, Czech Republic, Springer Verlag, Berlin, LNCS / LNAI Series (to be published).
- Neculai Curteanu. 2006. *Local and Global Parsing with Functional (F)X-bar Theory and SCD Linguistic Strategy*. (I.+II.), Computer Science Journal of Moldova, Academy of Science of Moldova, Vol. 14 no. 1 (40):74-102, Vol. 14, no. 2 (41):155-182.
- Paola Monachesi. 2005. *The Verbal Complex in Romance. A Case Study in Grammatical Interfaces*. Oxford University Press, Oxford Studies in Theoretical Linguistics.
- Robert D. Van Valin, Jr. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.