

## WHERE FOCUS FORMULAS AND DISCOURSE MARKERS MEET

*Silvie Válková and Jarmila Tárníková*

### Abstract

Our paper, theoretically anchored in functional and systemic grammar, focuses on a relatively marginal type of focus formulas (FFs), referred to by Schmid (2001) as ‘*N-be-that*-constructions’ or constructions with *shell-Nouns* (cf. *The trouble/problem/fact... is that people have short memories.*). When we used corpus data (BNC, COCA) to verify the role of FFs in information packaging in text/discourse, we were faced with their two seemingly contrary manifestations: they occurred either (i) as relatively stable utterance-initial templates or (ii) as looser configurations, co-occurring with various discourse markers (DMs). Our hypothesis is that in the latter case, namely when interlaced into clusters of DMs, the FFs tend to adapt to the communicatively regulative (Leech 1983) roles of surrounding DMs, and extend their role as focalising devices by an additive role, i.e. to participate in overt language manifestations of a number of pragmatically-based communicative strategies associated with facework. Our aim is to verify the validity of our hypothesis by authentic language data.

### Key words

communicatively regulative units, focus formulas, shell-Noun, information packaging, discourse markers, facework, pragmatic enrichment

## 1 Some preliminaries

### 1.1 Intersection of two perspectives

Our topic emerged spontaneously at the intersection of two research perspectives: that of a discourse analyst, interested in a pragmatically-based approach to various communicatively regulative devices in discourse, i.e. *discourse signposts* in general and *discourse markers* in particular (Válková), and that of a syntactician, focusing on the processes of information packaging in sentence complexes, considered from a text/discourse perspective (Tárníková). Both of us, when verifying our working hypotheses with corpus data (BNC, COCA), were faced with the co-occurrence of the same devices, i.e. *focus formulas* and *discourse markers*, but to each of us, the co-occurrence patterns revealed different research values. For the discourse analyst, the focus formulas (e.g. *the thing/point/problem/trouble...is that*) were a by-product on the way to a context-sensitive analysis of left-periphery discourse markers (such as *well*,

*oh, ah, and now*, as in *Well, the thing is...; Now, the problem is...*), their mutual configurations (*Oh well, perhaps the trouble is...*), and their clustering with other discourse signposts (cf. 3.1); for the syntactician, it was the discourse markers that emerged from the data as a by-product on the way to a better understanding of the formal, functional, and distributional properties of the selected ‘*N-be-that-focus formulas*’ in the process of information packaging (e.g. *Well, the thing is er, I, I can...It’s quite nice how...*[BNC, KP5 S\_conv]). Both of us realised that by getting rid of our respective by-products and focusing on the respective targeted devices only, we would cut the devices off from their natural surroundings and underestimate their mutual distribution in discourse.

So the consensus was to consider the selected FFs and DMs together, with greater attention being paid in this phase of our joint research to FFs, using DMs as supportive interpretative cues to some less predictable functions of FFs that were emergent from corpus data.

Below are some preliminaries on which we based our research. These concern: (i) the status of FFs and DMs within the architecture of language (Halliday & Webster 2009: 231), with a distinction made between communicatively constitutive and communicatively regulative units (Leech 1983); (ii) a scalar approach applied in the search for tentative taxonomies, and (iii) the activation of both the paradigmatic axis of alternation and the syntagmatic axis of co-occurrence in context-sensitive data analyses.

## 1.2 Communicatively constitutive and communicatively regulative units

This distinction, based on Leech (1983), enables functional linguists in general and discourse analysts in particular to identify those language units whose primary function is to *constitute* the core of the meaningful potential in utterances (hence the term *constitutive*), and those whose primary role is to *regulate* the flow of communication in various ways. While the constitutive units are *rule-governed*, the regulative ones are *principle-controlled* in nature, with the principles understood as pragmatic principles of cooperation, politeness, and tact (Grice 1975, Leech 1983), subsumed in Halliday (2003) under the *interpersonal metafunction of language* (for details, cf. Tárnayiková 2012).

In our approach, both focus formulas and discourse markers will be looked upon as *communicatively regulative units*, with FFs regulating *information packaging* in discourse, and DMs regulating *the smooth flow of interaction* in various aspects, the most important ones being those associated with facework and the strategies used to avoid face-threatening acts (FTA).

### 1.3 Scalar approach

One of the crucial findings of the functional linguists of the Prague School was the existence within the scope of functions of language units of a *centre* (*core*) and a *periphery*, with a presupposed *gradient* (scale) of transition between the two polarities on an imaginary scale of functional stratification (cf. Daneš 1964).

In the sections below, we will echo the *centre-periphery* scalarity twice: first, in approaching the FF with introductory shell-Nouns (*the thing, problem, fact, issue...is that*) as *marginal*, i.e. peripheral means of information packaging (in comparison with such central (core) constructions for highlighting as *clefts, pseudo-clefts, or extrapositions* (Tárnyiková 2009: 85-112)), and second, in considering various degrees of *templatedness* of the above-mentioned formulas, dynamically strengthened or weakened in the interplay with various contextual factors. Thus, for example, a FF interlaced into a cluster of DMs, as in *Well, as a matter of fact, the thing, I mean, is that...*, reveals a lesser *degree of templatedness* and becomes a less reliable guide to its focalising function than a FF initiating the utterance in its base form, cf.

- (1) *The point is that none of those was accompanied by an earthquake.*  
[BNC CKC W\_fict\_prose]

If, in addition to this, the content of the THAT-clause is rather vague (cf. the underlined part in Example 2 below), with hesitation markers and false starts, and there is in fact nothing to highlight, the shift of FFs from their focalising function to *pragmatic* marking is even more obvious. Example 3 below illustrates how the templatedness can be diminished by the extension within the FF of some of its components, so that the basic skeleton of the FF is, as it were, scattered throughout the textual chunk; cf.

- (2) *Well I mean you see the problem is that how the rich erm rich peasants...*  
[BNC KM6(1050)]
- (3) *Right, erm, ah, I think the best thing we can do is look at the rate book...*  
[BNC JK S\_unclassified]

#### 1.4 Context-sensitive approach

What is emergent from corpus data (the BNC and COCA, representing two varieties of English, British and American), is the need for a *dynamic, context-sensitive approach* to both focus formulas and discourse markers, based on an ongoing process of the *negotiation* of meaning (in the Hallidayan (1975) sense of ‘learning how to mean’) and responsive to the interplay of communicatively constitutive and communicatively regulative units in discourse (cf. Section 4).

#### 1.5 Activation of two axes: vertical and horizontal

When processing the data by which we verify our working hypothesis about the expanding functions of the targeted FFs, we will activate both the *vertical*, paradigmatic axis of *alternation* (e.g. the choice within the spectrum of shell-Nouns used to initiate the focus formulas, such as *the thing/problem/fact/trouble/point...is that*) and the horizontal, syntagmatic axis of co-occurrence of the FFs with DMs, as in Example 4, where the FF is sandwiched, as it were, by the DMs *well* and *I mean*.

(4) *Well the thing is I mean he plays for England* [BNC KD6 S\_conv]

The mirror procedure will be applied to utterance-initial (left-periphery) DMs, which are mostly prefixed to the focus formulas and will be looked upon as alternating possibilities (*Oh/Ah/Well/Now*), mutually co-occurring with each other (*Oh, well; Well now...*) and in configurations with other discourse signposts (cf. 3.2).

#### 1.6 Terminological notes

So far, the research on FFs conducted within the theoretical framework of a functional systemic grammar has prevalently focused on the role of FFs in the distribution of communicative dynamism in utterances or utterance complexes – either within the theoretical model of the Prague School tradition, known as the *Functional sentence perspective* (FSP), cf. Mathesius, Firbas, Dušková, Svoboda, Chamonikolasová, Adam, and others, or as *Topic-Focus Articulation* (TFA) in the model developed by Sgall, Hajičová, and their collaborators and followers.

In order to find a referential term tailored to our partial goal, i.e. to avoid framework-specific connotations, we opted for the term *information packaging*, leaving aside such possibilities as *information staging*, *grounding of the text*, or *thematic variations of clauses*.

The general term *focus formulas*, borrowed from Tuggy (1996: 724-726), will be modified in the subsequent sections into the more specific term *focus formulas with shell-nouns (FFSNs)* to reflect the narrowing of our scope to one specific type of focus formulas (or *constructions* in Dik 1980: 215-229).

The term *shell-Nouns*, understood as a convenient shorthand for ‘use-as-shell nouns’ (Schmid 2000: 4), will be preferred here to other existing labels, such as *abstract Nouns* or *catch-all Nouns*.

The terms *left-periphery* and *right-periphery discourse markers* will be used with reference to DMs used in the utterance-initial position (cf. *Well, yes, I mean, you know* [+ proposition]), or appended to the end of the utterance (cf. *It was a welcome party, or whatnot/or what you may call it/or something like that/and stuff like that/and all that/and suchlike...*).

### 1.7 Aim and research questions

Our aim is twofold: to contribute to a more ‘delicate’ (in the Hallidayan sense of the word) delimitation of the status, form, and function(s) of the selected FFSNs, and – by tracing their interplay with DMs – to contribute to new insights into the interface between discourse and grammar (Aijmer 2007) in this particular domain.

The following research questions reflect the above-mentioned intersection of the two perspectives:

- a. why the base form of the FFSNs (cf. 2.3 below), expected to be a good guide to information packaging, i.e. a kind of a fixed focalising template, occurs in the data with so many structural and lexico-semantic variables;
- b. how the frequent co-occurrence of FFSNs with discourse markers, which either precede the formula or are interlaced into it, contributes to a pragmatic enrichment of the role of FFSNs in discourse.

Guided by our research tasks, we will first focus on the FFSN and discuss its formal, functional, and distributional properties (Section 2), then briefly survey the functional properties of utterance-initial DMs and their co-occurrence in Section 3, and after that consider the impact of the mutual co-existence of DMs with FFSN on the pragmatic enrichment of the use of FFSNs, supported by the interplay of various contextual factors (cf. Section 4).

### 1.8 Why bother about peripheral phenomena?

The peripheral status of FFSNs in information packaging leads many linguists either to neglect them or refer to them as *other* constructions for rhematising

(focalising/highlighting) parts of the information structure of a given utterance or a textual chunk. But, as we hope to show, even a seemingly dispensable linguistic unit or structure can be rich in offering a variety of parameters to trace and functions to disclose. Moreover, as is evident from the dynamism of language development (Vachek 1976), the processes of coming into existence, dying out, revitalisation, and/or re-evaluation are inseparable from the continuity of language use.

## 2 Focus formulas with shell-Nouns (FFSNs)

### 2.1 Characterising FFSNs

As in many places in this paper, we will benefit from the studies of Schmid (2000, 2001) and Stvan (2007, 2014), selectively taking into view the approaches of Tuggy (1996), Kelzer (2013), and Delahunty (2011).

Schmid (2001: 1529) describes FFSNs as ‘*N-be-that-constructions*’, or, alternatively, as constructions with *shell Nouns*. Delahunty (2011) speaks about ‘*Thing Sentence Matrixes*’, while Stvan (2014) focuses on the first component of the formulas, i.e. the shell-Noun, making a distinction between *full shell Nouns* and *bare shell Nouns* on the basis of the presence or absence of the Determiner preceding the shell-Noun (cf. *The thing is...* vs. *Thing is...*). Kelzer (2013) prefers to describe the focus formula as a sequence ‘*The X is (is)*’ construction, to explicitly indicate the spectrum of possibilities in the lexico-semantic representation of shell-Nouns.

The shell metaphor explicates the role of the FFs in the act of interaction, i.e. to function as a *shell* into which the propositional content, i.e. a chunk of information, is encapsulated, as in *The thing is [that the girl was alone with her father in the house.]* [COCA, 2009 FIC].

While to Schmid (2001: 1538), shell-Nouns are above all “cataphoric signposts pointing to the complementing *THAT-clause*”, our arguments go a step further by stating that in the linear arrangement of the whole utterance, the shell-N shifts the shelled content of the complementing *THAT-clause* towards the end of the utterance, where, as a result of the *principle of end-focus*, it can be focalised as a “played-out” piece of information.

### 2.2 Definition of shell-Nouns and their taxonomy

Inspired by Schmid (2000), we will apply the following characteristics of shell-Nouns in our data processing:

“Shell nouns make up an open-ended functionally-defined class of abstract nouns that have, to varying degrees, the potential for being used as conceptual shells

for complex, proposition-like pieces of information.” ...“The property of shell-nounhood is a functional property” (2000: 4).

Schmid (2001: 1549) offers a long list of nouns that can function as conceptual shells (ranging in his frequency arrangement from *problem, thing, truth, fact, trouble, point, result, view ...* to *suggestion, consensus, and opinion*).

In looking for classifying criteria by which to reflect the scalarity of semantic emptiness/saturation of the shell-Nouns, we found Schmid’s (2001: 1532) proposal of four classes (and subclasses), based on the distinctions of *factual, linguistic, mental, and modal* shell-Nouns, too specific for our purposes, since the scalarity of semantic emptiness/saturation, which is relevant for our hypothesis about the primary (focal) and secondary (discourse marking) roles of FFSNs, goes across the proposed classes.

Consequently, in the present state of our research, such distinctions as inherently positive/negative connotations of shell-Nouns (*hope vs. problem*) and their semantic vagueness/saturation (*thing vs. point, reason...*) were found more relevant.

### 2.3 Structural properties of FFSNs

The base structure of the FFSN can be described as  
 [(Det)+N<sub>abstract, vague, sg.</sub> +V<sub>linking</sub> + (THAT)-clause<sub>content</sub> ]

**Det** stands for the Determiner, i.e. *the, my*, or a functional zero (e.g. *the point is, my point is, Δ point is*). The presence or absence of the Det reflects a creeping option, restricted to a limited set of shell-Nouns (cf. Stvan 2014: 600-601 for her COCA-based list of ‘bare’ shells said to visualise the grammaticalising process in the inner structure of FFSNs, cf. *Truth is, Trouble is, Thing is*).

**N** stands for the shell-Noun, which is abstract, semantically rather vague (unspecific), and used in sg.; it can be extended by pre-modifiers and/or post-modifiers (*the important point, the point I would like to make*).

**V** is the linking verb BE, mostly used in the simple present, less frequently in the simple past (cf. e.g. the BNC results for the query “point is”, with 6,429 occurrences of BE<sub>pres</sub> and 2,552 occurrences of BE<sub>past</sub>, with the ratio being 2.5:1 in favour of the simple present tense forms).

**THAT** stands for the connective signalling the subordinate status of the following content clause. If THAT is deleted, the content clause is merely juxtaposed to the FFSN, as in *But the thing is she must have seen him*.

**Clause types:** besides the prototypical THAT-clause, INF(itive)-clauses, HOW-clauses, IF-clause, or WHERE-clauses also emerge from the data, as in

*The main thing is to get in early.* [BNC CEF]. *The point is how to keep the religious fanatics from executing Gella?* [BNC, 1991 FIC BkSF:StarfireDown]. Our focus will be on prototypical THAT-content clauses and their juxtaposed variants.

Unlike Schmid (2001), who uses the ‘*N-be-that*-constructions’ as a base, we have also inserted Det into the base, arguing that the construction with Det (i.e. a *full shell*) is significantly more frequent in the corpus data, and hence perceived as unmarked, compared to *bare shells*, whose occurrence in the corpora is rather limited but – as Stvan (2014) pointed out – is increasing significantly in number, at least in American English.

## 2.4 Narrowing the scope

The following shell-Nouns were selected from the BNC corpus on a quantitative basis: *thing* (6,667 results), *problem* (6,447), *point* (6,425), and *fact* (5,841). This frequency-based list of preferences also prompts the existence of a potential spectrum of semantic saturation of the shell-Nouns (cf. the semantically vague *thing* vs. inherently negative *problem*, inherently focal *point*, or objectivising *fact*).<sup>1</sup>

## 2.5 Templatedness of FFSNs (a scalar approach)

Our first step is to consider the *inner configurations* of the elements within the template of the FFSN and pay attention to corpus-emergent *constants* and *variables*. The constants correspond to the above-mentioned base structure [Det+N+BE+THAT-clause], as in

(5) *The problem is that the law is artificial.* [BNC A2P (692)].

The base structure is looked upon as a sequence with the highest degree of structural templatedness but with a whole spectrum of lexico-semantic alternations within the shell-Nouns.

The variables are potential alternations and consequent modifications within each of the base components (cf. Table 1 below), changing the degree of templatedness in various ways.

Delahunty (2011: 109) speaks about “contextually determined fixity and flexibility in *Thing sentence matrixes*”, as prompted by the very title of his study.

To illustrate the spreading of variables within the template, let us compare the following BNC and COCA data, with the template in bold. The query was “*Thing is*” (with *thing* intentionally chosen as the semantically vaguest shell, used by many authors as a prototype of shell-Nouns (cf. Delahunty 2011)).

	Data source	DMs	Det	Premod.	shell	Postmod.	link	connective	Shelled content
1	COCA_2009 FIC		<b>The</b>		<b>thing</b>		<b>is</b>		the girl was alone with her father in the house
2	BNC_CJR W_pop		<b>The</b>		<b>thing</b>		<b>is</b>	<b>that</b>	killing got to be a habit with Mario, adds his friend
3	BNCA7C W_bibliography		<b>The</b>	curious	<b>thing</b>		<b>is</b>	<b>that</b>	Lewis, though so very widely read in other areas, had read...
4	BNCADR W_bibliography		<b>The</b>	important	<b>thing</b>	with this tour	<b>is</b>	<b>that</b>	it <u>is</u> personal
5	BNC_KD5 S_conv				<b>Thing</b>		<b>is</b>		it's all right to tell people
6	BNC_KCV S_conv	Right, right (unclear)			<b>thing</b>		<b>is</b>		she's got her new pencil case.
7	BNC JK7 S_unclassified	Right, erm, ah, I think	<b>the</b>	best	<b>thing</b>	we can do	<b>is</b>		to look at the rate book for this question
8	BNC HUN (724)	Erm, well	<b>the</b>		<b>thing</b>		<b>is</b>	<b>that</b>	I still erm...
9	BNC K7 1(336)	Well the I mean, the	<b>the</b>		<b>thing</b>		<b>is</b>	<b>that</b>	erm over the year...
10	BNCJA9 S_meeting				<b>Thing</b>		<b>was</b>		we had lots of complaints last July
11	COCA, 1994 SPOK	Yeah, I think	<b>the</b>	important	<b>thing</b>	to remember	<b>is</b>	<b>that</b>	Buthelezi is the only one

Table 1: Structural variety in the FFSN template for “thing”

### *Constraints on modification*

Our survey above is far from complete and concerns only one out of many shell-Nouns of the relatively open-ended class. Although analogous surveys could map the situation with other shell-Nouns, with some of them there are constraints on the choice of modifying variables.

To exemplify the case, the shell-Noun *problem* can alternate with the more personal *my problem* (cf. ...*my problem with her book is precisely that...* [BNC EF0 W\_religion]), while there was no occurrence in our data of *\*my thing* to alternate with *the thing*, or *\*my fact* alternating with *the fact*.

While the shell-N *problem* frequently occurs with the partitive quantifier *part of*, as in *Part of the problem is...* [BNC CM8 (716)], there was no occurrence of this partitive premodification with e.g. *point*, since it would be strange to expect the *point* to be cognitively perceived as something that is segmentable, i.e. as *\*part of the point*.

The inherently negative shell-Nouns (*problem*, *trouble*) will resist the positively evaluative attributes (*\*the best problem*).

Here, however, we are in the domain of collocability, which is beyond the scope of the present study.

### *Tentative conclusions*

FFSNs occur either as utterance-initial base forms (cf. Example 2 in Table 1 above) or in a variety of structural configurations, which partly manifest the *extension* of the base form and partly its *reduction*. The extension concerns the possible pre- and post-modifications of the shell-Nouns, restricted by the lexico-semantic nature of the respective shell-Nouns. The reduction concerns the Det deletion and the THAT-connective deletion and its result is a focus formula with a bare shell and a juxtaposed content clause, cf. *Thing is she is not home*.

The high frequency of modified FFSNs with some of the shell-Ns (*thing*, *problem*), and their interlacing with DMs, leads us to agree with Schmid (2001: 1538) that the FFSN can “simply tell readers/hearers to watch out for what is going to follow”. Put differently, the more extended the FFSN is, the less transparent the templatedness that emerges, so that the number of variables and the length of the modifying parts in FFSN contribute to the suppression of its primary focalising function, opening up space to a context-sensitive pragmatic enrichment, specifically in the milieu of the co-occurring discourse markers.

### 3 Characterising DMs

Discourse markers, as understood here, are part of a larger group of discourse signposts belonging to overt language manifestations of Halliday's interpersonal metafunction of language, i.e. what Crystal (1994: 257) characterises as "Language used for...maintaining social contact rather than for exchanging information or ideas". As mentioned before, their participation in interaction is, above all, regulative and their share in this regulation is associated with the smooth flow of communication and avoidance of face-threatening acts. Their high frequency of occurrence is prototypical of the spoken mode in highly colloquial everyday encounters. Being meta-discourse signposts, they do not belong to any particular part of the propositional content of the utterance, which makes their distribution relatively flexible. Those operating as interpretative cues to various communicative intentions (e.g. hesitation, tentativeness, camouflaging a knowledge deficit, signals of turn-taking, turn-passing, topic-shifting) – cf. Schiffrin (1988) and Aimer (2002, 2007, 2009) – tend to occur in the left-periphery of the proposition. Those used as mitigating afterthoughts tend to occur in the right-periphery (cf. 1.6 above).

The semantic load of DMs is not inherently rich but rather situation-retrievable. In other words, the same DM can be used to signal various facets of communicative verbal monitoring and verbal manoeuvring. This flexibility opens up space to both advantages and disadvantages. Such statements, however, are too general to apply – and in a way irrelevant to our partial goals, i.e. to identify which DMs prototypically occur in the utterance-initial position, how symbiotic their co-occurrence with FFSNs is, and what the impact of this co-occurrence is on the strengthening or weakening of the focalising function of the FFSNs.

#### 3.1 Left-periphery discourse markers and their clustering

The left-periphery distributional requirement is met by four subtypes of DMs: *discourse connectives* (*and, but, or...*), *contact words*, referred to here as *gambits* (*oh, ah, well, and now*), *markers of cognition* (*I think, I mean*), *markers of personal involvement* (*you know, you see*) – and their mutual configurations in clusters either preceding the FFSN (cf. *oh, well, you know, I mean, the problem is that...*) or interwoven in it, as in

- (6) *Well, well, see how it goes, but erm I don't, the thing is you see they can teach you so much about it.* [BNC KCY(112)]

The four prototypical utterance-initial DMs mentioned above tend to co-occur in chains, initiated by a sentence connective (if any), and followed by the gambits, markers of cognition, and markers of personal involvement, as in

- (7) *But I mean you see the problem is that how the rich erm rich peasants...*  
[BNC KM6(1050)]

Expressions of agreement and disagreement and interjections can also be integrated into the left-periphery discourse signposts, as in

- (8) *Oh yes, oh gosh, yes, well the problem is...* [BNC KDM 8320]  
(9) *Well no they're the thing is I mean you're talking of five or six hundred.*  
[BNC KCG(1892)]

#### *In-group configurations of gambits*

Since the four alternating gambits, i.e. *oh*, *ah*, *well*, and *now*, can mutually co-occur and consequently lengthen the chain of utterance-initial DMs, we traced the frequency-based preferences emergent from the BNC data (inclusive of the repetition of the same DM, such as *oh oh*, *well well...* but also *oh*, *well*, *well ...*), as in

- (10) *Well, well see how it goes, but erm I don't, the thing is you see...*  
[BNC KCY (112)]

The in-group configurations, however, reflect only a language potential, which need not necessarily currently be met with, as the following survey shows:

<i>Oh well</i> (1533)	<i>Now now</i> (16)	
<i>Oh oh</i> (369)	<i>Ah now</i> (15)	
<i>Ah well</i> (236)	<i>Well well well</i> (15)	
<i>Ah ah</i> (222)	<i>Well ah</i> (8)	
<i>Well now</i> (184)	<i>Now oh</i> (6)	
<i>Well well</i> (141)	<i>Now well</i> (6)	
<i>Oh now</i> (26)		(taken from Válková 2012: 225)

#### **4 Co-occurrence of DMs with FFSNs in the left-periphery position**

The co-occurrence of discourse markers and focus formulas comprises a little-discussed but communicatively relevant area of our everyday interaction based on spontaneous and highly colloquial encounters. In Table 1 above, we identified

rich variation in the base formula. Such an extended template was considered a step to a less transparent focalising function and a possible prerequisite for pragmatic enrichment. The next source of support for the pragmatic enrichment within the functions of the FFNS is their co-existence with the above-mentioned discourse markers.

#### 4.1 DMs as left-periphery ‘prefixes’ to FFSNs

There are two basic manifestations of the co-occurrence of DMs with FFSNs. First, there is a transparent borderline between the DMs and the FFSN, visually supported by the compact, unmodified base form of the FFSN, as in

- (11) *Oh right! So, I mean (pause) the thing is, they are (pause) erm (pause) they are minority group.* [BNC KDW S\_conv]

or a slightly modified FFSN template, as in

- (12) *But I think the important thing is that we never forget, as Americans, what happened on September 11...* [COCA 2004 SPOK MSNBC\_Matt]

In such cases the prefixed DMs do not have any significant weakening impact on the focalising function of the FFSNs. If, on the other hand, the propositional content of the THAT-clause is vague, and there is in fact nothing to encapsulate into the shell-N, and, consequently, nothing to draw attention to, the FFSN loses its focalising potential, as in Examples 13 and 14:

- (13) *Yes. You see well the point was when you pick’em up erm we had a sm— we had a ...*

[COCA H5G S\_interview\_oral\_history]

- (14) *Mm. well I mean there’s, the, the other thing is the differential around the, the county, cos there was a time when when* [BNC KLX S\_meeting]

Second, the DMs not only precede the FFSN but extend their scope and penetrate into the FFSN and even further on, into the propositional content of the THAT-clause, cf.

- (15) *You see the thing is you see I mean wh wh when this law can...*  
[BNC KB7(3341)]

- (16) *Well the thing is you see I think what from what I’ve gathered that...*  
[BNC J8D(1996)]

- (17) *Well, well see how it goes, but erm I don't, the thing is you see they can teach you so much about it...* [BNC KCY(112)]
- (18) *Well, the thing is er; I, I can...It's quite nice how...* [BNC, KP5 S\_conv]
- (19) *I mean the thing is the (pause) they're not (pause) you know, some of them are not...* [BNC KE6 s\_conv]

In such situations, the focalising function of the FFSN is also weakened and the FFSN adapts to the discourse-signposting role of the surrounding DMs to participate in creating an atmosphere of familiarity and social nearness, in softening the austerity of the communication, in postponing a bad piece of news, camouflaging a lack of knowledge, or simply in keeping the channel of communication open while conveying almost nothing.

*What favours the focalising function of FFSNs?*

To answer this question, let us observe the behaviour of the shell-Noun *point*, which, as emergent from the BNC and COCA data, represents the highest degree of templatedness of all the five selected shell-Nouns (followed by *fact*).

In the total of 200 random samples obtained for the query '*The point is*' in the BNC, the modification of the shell-Nouns was sporadic, and, moreover, restricted to pre-modifiers with a semantic load of importance (*crucial, essential, first, (very) important*), which is support for the focalising function of a focus formula; only four times was the FFSN preceded by DMs: *now* (2) and *you see* (2), and only in 21 samples out of the total was the shell-Noun *point* preceded by discourse connectives: *but* (13), *and* (3), *anyway* (1), *however* (3), *either way* (1), cf.

- (20) *The point is that they are all vital young men* [BNC A06(262)]
- (21) *But the point is that it was permanently visible.* [BNC A9K(32)]
- (22) *The crucial point is that* [deleted: proposition] *is negative.*  
[BNC J12 W ac nat science]

In the COCA data, the fixity of the template was even more visible, since the choice of modifiers was restricted in the random 200 samples to *main* and *important*.

Conclusion: the focus formula with the shell-Noun *point* is mostly used in its base form as a transparent signal of focalisation of the encapsulated content clause. The factors contributing to the focalising function of the FFSN '*The point is*' can be summed up as follows:

- high degree of templatedness (FFSN is mostly used in its base form, as a compact, uninterrupted whole);
- rare co-occurrence of the FFSN with initial DMs and their clusters;
- absence of DMs interlaced into the FFSN or the following THAT-clause;
- semantically rich content of the encapsulated proposition (cf. the content clause);
- degree of semantic load of the shell-Noun (*point, fact* vs. *thing*).

This matrix is also applicable to the shell-Noun *fact* in our data.

*What favours the discourse-signposting function?*

On the basis of Table 1 and the query *The thing is*, the following priorities seem to hold:

- loose templatedness of the FFSN, with pre- and post-modifications of the shell-N;
- semantically vague shell-N (*thing*);
- vague content of the encapsulated proposition (where there is in fact nothing to put a focus on and the presupposition about the focalising effect of the FFSN triggers wrong expectations);
- interlacing of FFSNs with introductory gambits and other DMs with left-periphery, i.e. utterance-initial, distribution;
- penetration of DMs into the propositional content of the encapsulated clause.

Considering the relevance of the content conveyed by the complementing (THAT) clause, we can echo the well-known saying that *there are two sides to everything and not all the conceptual shells promise to encapsulate 'silver pearls'*, as in

(23) *The point is (unclear) I mean the thing is were (unclear) some good conversation.* [BNC KDA S\_conv]

The first visual impression of the utterance-initial marking might be a presuppositional 'bluff', to use Schmid's (2001) appropriate description of the situation illustrated by Example 23.

In Example 23 there are even two focus formulas (cf. *the point is* and *the thing is*) used to shell the proposition, from which we might deduce a strong emphasis on the propositional content. There are, however, devices in the utterance going against the grain of such expectations (cf. the substitution of the more specific shell-N *point* by the semantically vague *thing*, the mitigating DM *I mean* inserted in between, and the 'supposed to be' content clause), of which only some lexical hints can be retrieved. Such a configuration of structural, lexico-semantic, and

textual properties undermines the focalising function of the FFSN in a significant way, giving us at the same time an argument in favour of a dynamic, context-sensitive approach to this whimsical domain of human interaction.

### *Prospects for the future*

The following tasks lie ahead of us: more evidence is needed to address genre-specific and mode-specific (spoken vs. written; monologic vs. interactive discourse) occurrences of FFSNs in authentic language data; a deeper insight is needed into the British and American varieties of English, since, as reported by Stvan (2014), the processes of grammaticalisation and pragmatic enrichment of FFSNs emerge from the COCA data in a more intensive way, resulting not only in an increased number of bare shells (*Problem is, Thing is*), but also in innovative patterns of their use and ongoing changes in their status (cf. the spread of condensed forms, such as *Problem being*, reported by Stvan (2014).

## **5 Concluding remarks**

Both focus formulas with shell-Nouns and discourse markers belong among the overt language manifestations of discourse management (meta-discourse markers). Both participate in the overt language manifestation of interpersonal metafunction, so that their co-occurrence in discourse is predictable.

As communicatively regulative units of language, they are principle-controlled, so that the number of their occurrences in discourse is not so relevant as the strategy of saying the right thing at the right time in order to guarantee the smooth flow of communication, the avoidance of face-threatening acts, and the creation of a 'happy situation' for our everyday encounters.

Though FFSNs primarily participate in regulating information packaging, while DMs regulate the smooth flow of communication, their mutual co-occurrence in the utterance-initial (left-periphery) position creates a pre-condition for their stronger symbiosis, in which FFSNs, when interwoven into the clusters of initial discourse markers, lose the visual support of a fixed template and tend to assimilate to the strategic manoeuvring typical of utterance-initial discourse markers.

The response to the title of this paper is relatively simple: focus formulas with shell-Nouns can meet discourse markers in the left-periphery, utterance-initial position, in which the FFSNs are either linearly preceded by DMs or interwoven into their clusters.

Our theorising, inspired by the studies on the topic that already exist (cf. References) aimed at verifying the validity of the hypothesis, i.e. that even marginal language devices, such as focus formulas, can extend their functional

load over time, and undergo inner hierarchisation into a core function and a spectrum of add-ons.

The application of the context-sensitive and scalar approach to authentic corpus data enabled us to understand the reasons for the dual nature of FFSNs, oscillating between focalisers and discourse signposts, and to reveal the factors favouring one or the other of the regulative functions.

Although the conceptual categories are drawn from earlier research, our approach is innovative in taking into view not only the paradigmatic axis of alternation within the FF variables but also the syntagmatic axis of co-occurrence of the FFs with other left-periphery communicatively regulative language devices (cf. 3.1). This mutual data-based co-occurrence enabled us to reveal the correlation between the focalising effect of the FFs and the degree of their structural fixity in the chain of left-periphery discourse signposts, and the correlation between the focalising effect of FFs and the semantic load of the shelled proposition.

We hope that our findings, together with our tentative conclusions, will provoke further debate in the community of scholars prioritising a multifaceted approach to language data, sensitive to the interface between grammar and discourse.

## References

- Aijmer, K. (2002) *English Discourse Particles*. Amsterdam and Philadelphia: John Benjamins.
- Aijmer, K. (2007) 'The interface between discourse and grammar: *The fact is that*.' In: Celle, A. and Huart, R. A. (eds) *Connectives and Discourse Landmarks*. Amsterdam and Philadelphia: John Benjamins. 31-46.
- Aijmer, K. (2009) 'The pragmatic marker *well*: A text study.' In: Dontcheva-Navratilova, O. and Povolná, R. (eds) *Coherence and Cohesion in Spoken and Written Discourse*. Cambridge: Cambridge Scholars Publishing. 4-20.
- Chamonikolasová, J. (2010) 'Discourse markers in Czech and English conversation.' In: *Languages in the Integrating World*. LINCOM Studies in Communication München: Lincom Europa. 207-218.
- Crystal, D. (1994) *A Dictionary of Linguistics and Phonetics*. Oxford: Blackwell Publishers.
- Daneš, F. (1964) 'A three-level approach to syntax.' *Travaux linguistiques de Prague* 1, 225-240.
- Delahunty, G. (2011) 'Contextually determined fixity and flexibility in thing sentence matrixes.' In: Kuiper, K. (ed.) *Yearbook of Phraseology* 2. Berlin: Mouton de Gruyter. 109-136.
- Delahunty, G. (2012) 'An analysis of The thing is S sentences.' *Pragmatics* 22/1, 41-78.
- Dik, S. (1980) *Studies in Functional Grammar*. London: Academic Press.
- Dušková, L. (2010) 'Syntactic construction, information structure and textual role: An interface view of the cleft sentence.' *Brno Studies in English* 36/1, 29-45.

- Grice, H. P. (1975) 'Logic in conversation.' In: Cole, P. and Morgan, J. (eds) *Syntax and Semantics 3, Speech Acts*. New York: Academic Press. 41-58.
- Halliday, M. A. K. (1975) *Learning How to Mean*. London: Edward Arnold.
- Halliday, M. A. K. (2003) 'On the 'architecture' of human language.' In: Webster, J. (ed.) *On Language and Linguistics. Collected Works of M. A. K. Halliday*, Vol. 3. London and New York: Continuum.
- Halliday, M. A. K. and Webster, J. J. (eds) (2009) *Continuum Companion to Systemic Functional Linguistics*. London and New York: Continuum.
- Kelzer, E. (2013) 'The X is (is) construction: An FDG account'. In: Mackenzie, J. L. and Olbertz, H. (eds) *Coursebook in Functional Discourse Grammar*. Amsterdam and Philadelphia: John Benjamins.
- Leech, G. N. (1983) *Principles of Pragmatics*. London: Longman.
- Povolná, R. (2010) *Interactive Discourse Markers in Spoken English*. Brno: Masaryk University.
- Povolná, R. (2012) 'Causal and contrastive discourse markers in novice academic writing.' *Brno Studies in English* 38/2, 131-148.
- Schiffrin, D. (1988) *Discourse Markers. Studies in Interactional Sociolinguistics 5*. Cambridge: Cambridge University Press.
- Schmid, H. J. (2000) *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition. Topics in English Linguistics*. Berlin and New York: Mouton de Gruyter.
- Schmid, H. J. (2001) 'Presupposition can be a bluff: How abstract nouns can be used as presupposition triggers.' *Journal of Pragmatics* 33, 1529-1552.
- Stvan, L. S. (2007) 'The functional range of bare singular count nouns in English.' In: Stark, E., Leill, E., and Werner, A. (eds) *Nominal Determination: Typology, Context, Constraints, and Historical Emergence*. Amsterdam and Philadelphia: John Benjamins. 171-187.
- Stvan, L. S. (2014) 'Truth is, sentence-initial bare shell nouns are showing up bare.' In: Veselovská, L. and Janebová, M. (eds) *Complex Visible Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*. Olomouc Modern Language Series, Vol. 4. Olomouc: Palacký University. 591-606.
- Válková, S. (2012) *Regulating Discourse: Compliments and Discourse Signposts (English-Czech Interface)*. Saarbrücken: LAP Lambert Academic Publishing.
- Tárnyiková, J. (2009) *From Text to Texture. An Introduction to Processing Strategies*. Olomouc: Univerzita Palackého v Olomouci.
- Tárnyiková, J. (2012) 'Halliday's interpersonal component reconsidered.' In: Hopkinson, C., Tomášková, R. and Zapletalová, G. (eds) *The Interpersonal Language Function Across Genres and Discourse Domains*. Ostrava: Universitas Ostraviensis. 26-36.
- Tuggy, D. (1996) 'The thing is is that people talk that way. The question is why?' In: Casad, E. H. (ed.) *Cognitive Linguistics in the Redwoods. The Expansion of a New Paradigm in Linguistics*. Berlin: Mouton de Gruyter. 713-752.
- Vachek, J. (1976) *Selected Writings in English and General Linguistics*. Praha: Academia.

### Corpora

- Davies, M. (2008) *The Corpus of Contemporary American English (COCA): 400+ Million Words, 1990-present*. URL: <http://corpus.byu.edu/coca/>.
- The British National Corpus*, version 3 (BNC XML Edition). (2007) Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk>.

(Endnotes)

<sup>1</sup> Quantity is perceived in the present state of our analysis as a supportive means for tracing scalarity and its opposite poles. This makes us operate with rather vague relational quantities, such as more/less or prevailing/marginal, since we are interested in emergent tendencies of principle-controlled features that have no prescribed rules for their occurrence. It is more important for us in this phase to identify structural and semantic fixity and flexibility, and to consider statistical matrixes and graphs a necessary prospect for the future. We hope that having used '200' random samples from the BNC and COCA respectively makes us qualified to notice some creeping tendencies in use and identify language means that – emergent from the data – support our initial working hypothesis about the partial pragmatization of FFSNs in a well-defined contextual milieu.

**Silvie Válková** is a lecturer at Palacký University, Olomouc, Czech Republic. Her fields of interest are interlanguage pragmatics and cross-cultural communication, focusing mainly on politeness strategies and their language manifestations. She is the author of two monographs: *Politeness as a Communicative Strategy and Language Manifestation* (2004) and *Regulating Discourse: Compliments and Discourse Signposts (English-Czech Interface)* (2012).

**Address:** Mgr. Silvie Válková, Ph.D., Centre of Language Education, Faculty of Arts, Palacký University, Vodární 6, 771 80 Olomouc, Czech Republic. [e-mail: silvie.valkova@upol.cz]

**Jarmila Tárnyiková** is Professor Emerita of English linguistics in the Department of English and American Studies, Faculty of Arts, Palacký University, Olomouc. Her main research areas embrace Present-day English syntax (*Sentence Complexes in Text. Processing Strategies in English and in Czech*, 2007), Text/Discourse analysis (*From Text to Texture. An Introduction to Processing Strategies*, [2002], 2010), Pragmatics (Chapter 6 in *Rudiments of English Linguistics*, 2000) and the Language of fiction. Her present research focuses on corpus-based analysis of vague language and a cross-language analysis of loanwords. She has been nominated member of the Prague Linguistic Circle. Her monograph on *Sentence Complexes in Text* received Rector's Award for the best monographs of 2007. She is Associate editor of *Linguistica Pragensia*, co-editor of *Anglica III Linguistica* (2009).

**Address:** Prof. PhDr. Jarmila Tárnyiková, CSc., Department of English and American Studies, Faculty of Arts, Palacký University, Olomouc, Křížkovského 10, 771 80 Olomouc, Czech Republic. [e-mail: jarmila.tarnyikova@upol.cz]