THEME MOVEMENT
IN ACADEMIC DISCOURSE[1]

Libuše Dušková

1. This paper examines the tendency of English to retain the same subject in successive clauses with regard to its largely prevalent function of theme. Both these properties of the English subject were first pointed out by Vilém Mathesius[2] and confirmed in further studies[3] especially with regard to the subject's thematic function. The other property of the English subject, the disposition to denote the same entity in successive clauses, subsequently appeared worth examining in connection with the concept of thematic progressions proposed by František Daneš.[4]

The concept of theme is based on Jan Firbas's theory of functional sentence perspective (FSP, i.e., information structure).[5] The FSP theory works with four FSP factors: semantic structure, context dependence/independence (cf. discourse old and discourse

[1] A shortened version of this paper was presented at the Lund conference of the International Association of University Professors of English (IAUPE) in 2007.
[2] Vilém Mathesius, *A Functional Analysis of Present Day English on a General Linguistic Basis* (Prague: Academia, 1975) 100-103; Vilém Mathesius, "O funkci podmětu" (On the Function of the Subject), *Čeština a obecný jazykozpyt* (Czech Language and General Linguistics) (Prague: Melantrich, 1947) 100-103.
[3] Libuše Dušková, *Studies in the English Language*, Part 1, Part 2 (Prague: Karolinum – Charles University Press, 1999) 2:273-80.
[4] František Daneš, "Functional Sentence Perspective and the Organization of the Text," *Papers on Functional Sentence Perspective*, ed. František Daneš, (Prague: Academia, 1974) 106-28.
[5] Jan Firbas, *Functional Sentence Perspective in Written and Spoken Communication* (Cambridge: Cambridge University Press, 1992).

new),[6] linearity and intonation in speech. The interplay of these four factors determines the FSP function (theme, transition, rheme) of the clause elements. In the deep (interpretative) word order the clause elements are arranged according to gradual increase in the information load (i.e., the degree of communicative dynamism). This arrangement constitutes the basic distribution of communicative dynamism, and basically corresponds to what is called the principle of end focus in British grammar.[7] The theme and the rheme are defined, respectively, as the elements carrying the lowest and the highest degree of communicative dynamism irrespective of sentence position. This is the meaning in which the term "theme" is used henceforth. In a more general sense of "the subject matter" the term employed is the "hypertheme."

As regards thematic progressions (TP), the following analysis draws on Daneš, who has proposed the following major types:[8] (1) Simple linear TP (= TP with linear thematization of rhemes): each rheme becomes the theme of the next utterance. (2) TP with a continuous (constant) theme: successive utterances contain the same theme to which different rhemes are added. (3) TP with derived themes: the utterance themes are derived from a "hypertheme" (of a paragraph, or another text section). Of the various combinations into which these types of thematic progressions may enter, Daneš mentions (4) the progression with a split rheme: a rheme composed of more than one (mostly coordinated) component provides successive themes constituted by each component treated separately.

Considering these types of theme development in connection with the tendency of English to preserve the same subject in successive clauses and the subject's largely thematic function, it can be assumed that a prominent role in the text build-up will be played by the second type of thematic progression, viz. the progression with a constant theme.

---

[6] Rodney Huddleston and Geoffrey K. Pullum, *The Cambridge Grammar of the English Language* (Cambridge: Cambridge University Press, 2002) 1370.

[7] Randolph Quirk, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik, *A Comprehensive Grammar of the English Language* (London: Longman, 1985) esp. 1361-62; cf. also the general tendencies regarding information structure in Huddleston and Pullum, *The Cambridge Grammar of the English Language*, 1372.

[8] Daneš, "Functional Sentence Perspective," 118-20.

2. This assumption was tested on a section of text on a natural science topic comprising 100 clauses (both finite and non-finite).[9] The results are given in Table 1 (see Appendix); items with identical superscripts denote recurrent subjects, items without superscripts denote subjects with one occurrence (different subjects) and subjects due to syntactic construction (*it* and *that* as components of a cleft sentence, and *it* in construction with extraposed infinitive).[10] Where discussed, these subjects are referred to by the number of the respective clause.

As a prerequisite of arriving at the data presented in Table 1, it was necessary to determine which subjects were to be regarded as identical (recurrent) and which as different. This appeared to be a matter of degree rather than a clear-cut division. As shown by columns 2 and 3, which list identical subjects, i.e., subjects with more than one occurrence, apart from instances displaying full identity, denoted by numbers alone, these columns also include subjects whose identity is only partial, which are marked by the respective number + a letter. Six items are found here: *label*[(1)], *black-footed squirrel*[(2)], *time*[(3)], general *we*[(4)], *the zoologist*[(5)], and *human species*[(6)].

Table 2: Subjects with More Than One Occurrence

Item	Number of Occurrences
Label[(1)]	2
Black-footed squirrel[(2)]	30
Time[(3)]	2
General subject *we*[(4)]	31
The zoologist[(5)]	3
Human species[(6)]	12
Total	80

The figures in Table 2 show that among the six recurrent items, two display a much higher frequency of occurrence than the rest: item [(2)], *the black-footed squirrel*, the hypertheme of a greater part of the examined text section (30 occurrences); and item [(4)], general *we*,

---

[9] Source text: Desmond Morris, *The Naked Ape* (1967) (London: Triad Grafton Books, 1986) 13-15. See Appendix.

[10] Cf. information packaging constructions in Huddleston and Pullum, *The Cambridge Grammar of the English Language*, 1366 v and viii.

represented by 31 occurrences. Item (6), *the human species*, shows less than a half of this frequency of occurrence (12 instances). However, this is due to the limited length of the text: this item appears in the last part of the examined section, and reflects a shift to a new hypertheme, which is the global theme of the entire text. Items (1) and (3), each with two occurrences, are of interest in that they denote the local and temporal scene on which the first hypertheme is discussed. In this respect they are inherent components of the content build-up, their low frequency of occurrence indicating their background, scene-setting nature.

The four recurrent items with the highest frequencies of occurrence – *black-footed* squirrel(2), general we(4), *the zoologist*(5), *human species*(6) – may be divided into two groups: the two hyperthemes *black-footed squirrel* and *human species* on the one hand, and general *we* and *the zoologist* on the other. Even here, the dividing line is not quite clear-cut: *the zoologist* is associated with both hyperthemes, but like the general *we* and unlike the two hyperthemes, does not represent 'what is spoken about', and in that sense both *we* and *the zoologist* are external to the text and can be classed together.

The distinction between the two groups moreover appears in their realization forms. As shown in Table 1 (see Appendix), the superscripts 2 and 6, which denote the two hyperthemes, are in a number of instances followed, respectively, by the letters a-f and a-d. These indicate partly identical subjects, and are specific to the hyperthemes; the items forming the other group, the general *we* and *the zoologist*, do not display them. In the case of the two hyperthemes both fully and partly identical subjects are realized by three forms: anaphoric pronouns (personal, relative and demonstrative), zero (ellipsis in finite clauses and structural deletion in nonfinite clauses) and referentially identical noun phrases. General *we* and *the zoologist* have only pronominal and zero form of realization (which is accidental in the case of the latter).

2.1 Examining the two hyperthemes in more detail we arrive at the following findings.

2.1.1 The hypertheme *black-footed squirrel* is realized by pronouns and zero both in the case of fully and partly identical subjects. Fully identical subjects have pronominal realization in 9 instances (in clauses 5, 6, 9, 10, 13, 20, 22, 28, 42); the zero form occurred in one instance (in clause 23); in total 10 instances. See Table 1, items with the superscript [2] in columns 2 and 3. In the case of partly identical subjects anaphoric pronouns are found in three instances (cf. clauses 35 *they* [*mating calls and displays*], 43 *these* [*the markings of its fur – its black feet*], 58 *it* [*every aspect of its behaviour*]); and zero in 5 clauses (15, 30, 32, 34, 59); altogether 8 instances.

The distinction between fully and partly identical subjects results from the differences between the objects being referred to, which are primarily denoted by nouns (antecedents of the pronominal and zero realization forms). Fully identical subjects are found among referentially identical noun phrases reiterating the same head noun (clauses 4, 7, 19, 25; altogether 4 instances). A group of 3 designations by a more general, superordinate term (hypernym) is again a border line case assignable to either fully or partly identical subjects: clause 3 *this animal*, clause 21 *the new form*, clause 36 *a new species*. Superordinate terms have been classed as fully identical. Clause 14 *the ancestors of this animal*, where the head noun denotes a subclass, is connected with this group through its postmodification, and is hence included in this group (accordingly, 4 instances in total).

Partly identical subjects are characterized by a distinctive semantic feature, mostly the relationship of part/whole with respect to the hypertheme, which may also involve an inalienable characteristic (4 instances): *social and sexual behaviour* [of *black-footed squirrels*] in clause 29, *their anatomy* in clause 31, *mating calls and displays* [of *black-footed squirrels*] in clause 33 and *the markings of its feet* in clause 41. See Table 1, items with the superscripts [2, 2a-f] in columns 2 and 3.

2.1.2 The hypertheme *human species* presents a similar picture. Pronominal and NP realization is found in both fully and partly identical subjects: pronouns in clauses 77, 83, 89, 92 and 99 (*that* referring to a partly identical subject *skin*), altogether 5 instances; noun phrases: hypernyms classed as fully identical in clauses 78 (*a strange form of life*) and 96 (*this species*). Partly identical noun phrases

(e.g., *legs, arms*) occurred in clauses 93, 94, 95, 98 and 100. In total the hypertheme *human species* is represented by 6 fully identical subjects and an equal number of partly identical subjects, altogether 12 instances. See Table 1, items with the superscripts [6, 6a-d] in columns 2 and 3.

2.2 Recurrent subjects of the second group, general *we* and *the zoologist*, differ not only from the hyperthemes, but partly also from each other.

2.2.1 As noted in Section 2, the distinctive features of the general *we* consist of the form of realization, semantics and the FSP aspect. Its two realization forms (pronominal – 18 instances, and zero – 13 instances [3 ellipses and 10 deletions]; in total 31 instances) entail the full identity of all occurrences. The semantics of the general *we* – designation of the general human agent, here people interested in zoology, including the text producer and text recipient – is reflected in the different FSP nature of this recurrent subject. Whereas the two hyperthemes in the subject function largely represent the theme both in the sense of the least dynamic element within the respective clause, i.e., the theme as defined in the FSP theory, they also denote what is being spoken about. On the other hand, the general *we* complies with only one of these aspects: it does not constitute what is spoken about, but merely the least dynamic element of the sentence. This distinction is closely connected with the other FSP feature of the general *we*, viz. its non-occurrence as the rheme. On the other hand both hyperthemes, representing what is spoken about, first appear as rhemes, mostly in syntactic functions constituting the verb complementation, but also as the subject. Where they operate as themes in a thematic section composed of several elements, the two aspects of the theme (the theme as the carrier of the lowest degree of communicative dynamism, and the 'aboutness' aspect) are realized by two separate elements, theme proper and diatheme (the most dynamic element within the thematic section[11]). What is spoken about then represents one of the components of the thematic section. The distinction between different thematic elements can be found even in Mathesius' initial general outline of the FSP theory: themes like

[11] Cf. Aleš Svoboda, *Diatheme* (Brno: Univerzita J.E. Purkyně, 1981).

general *we* would be classed as the starting point of an utterance (*východiště výpovědi*).[12]

2.2.2 The recurrent subject *the zoologist* is most similar to the general *we* in the last respect, i.e., in the text under study it does not convey what is spoken about, but represents the starting point for statements about the hyperthemes. However, its more specific meaning does not exclude it from what might be spoken about in another context, which presupposes previous occurrence in the rheme. The more specific semantic structure is moreover reflected in a narrower scope of inclusiveness. While it may (and virtually does) include the 'I' (the producer) of the text (a non-specialist could not write it), it is non-inclusive with respect to the text recipient (the addressee).

The *zoologist* has three occurrences in the text, one pronominal (clause 70) and two zero realizations (clauses 71 and 73).

3. Nonrecurrent (different) subjects.

Table 3

Noun Phrases		Clausal Subjects and Constituents of Syntactic Constructions
8 nothing		11 it (in *it*-cleft)
18 the new trend		12 that (in *it*-cleft)
39 all		16 it (in *it*-cleft)
44 (the rash) that		17 that (in *it*-cleft)
47 (the starting point) telling us		27 it (anticipatory S>S=infinitive)
48 something		51 that (clausal proform)
61 The great advantage		72 it (anticipatory S>S=infinitive)
65 (a fact) which		84 just how odd (S=verbless clause)
66 (humility) that		
67 things		
82 (other species) that		
97 (locomotion) which		
Total	12	8

---

[12] Vilém Mathesius, "O tak zvaném aktuálním členění větném" (On the so-called functional sentence perspective), *Čeština a obecný jazykozpyt*, 234-35.

The eight items in the right-hand column include four clauses constituting *it*-clefts and two instances of *it* anticipating extraposed infinitives (their subjects are listed in the third column of Table 1). Both constructions are derived from underlying simpler structures, from which they differ in their FSP. The two *it*-clefts represent the more frequent type (called stressed-focus *it*-clefts by Prince;[13] discourse-old presupposition *it*-clefts in Huddleston and Pullum[14]), in which the *that*-clause conveys context-dependent, given information.[15] The focused element is in both cases the interrogative pronoun *what*, which in the underlying non-cleft clause – in the absence of the prosodic factor, operating in speech – implements a thematic subject, cf.

Clauses 11, 12: (7 No black-footed squirrel has ever been found in that continent before. 8 Nothing is known about it. 9 It has no name. 10 For the zoologist it presents an immediate challenge.) 11 What is it about its way of life 12 that has made it unique? < What has made it unique?

Clauses 16, 17: (14 the ancestors of this animal must have split off from the rest and 15 established themselves as an independent breeding population.) 16 What was it in the environment 17 that made possible their isolation as a new form of life? < What made possible their isolation as a new form of life?

Whereas in the non-cleft forms the subject interrogative pronoun is thematic, in the clefts it becomes the rheme owing to the splitting of the propositional content into two clauses.

In the case of subject infinitives the extraposed construction appears to be the basic structure, since it is the extraposed component that usually contains new information. The principle of end focus here as a rule operates jointly with the principle of end weight. On the other hand, initial placement of a subject infinitive, regardless of the degree of the subject's weightiness, assigns it thematic function. Hence the FSP structures of the extraposed and non-extraposed construction usually differ in that both constructions have the rhematic component at the end, cf. *to become*

---

[13] Ellen F. Prince, "A comparison of *it*-clefts and *wh*-clefts in discourse," *Language*, 54 (1978): 883-906.

[14] Huddleston and Pullum, *The Cambridge Grammar of the English Language*, 1424.

[15] In the less frequent type the *that*-clause contains new, context-independent information. For more detailed treatment, see my *Studies in the English Language*, 318-31.

isolated from possible contamination by their neighbours in clause 27, but *advantageous* in the non-extraposed form, added for comparison; and similarly in clauses 72-73.

Clauses 27, 28: 27 it would be advantageous for them 28 to become isolated from possible contamination by their neighbours. x To become isolated from possible contamination by their neighbours would be advantageous for them.

Clauses 72, 73: 72 Even for the zoologist [...] it is difficult 73 to avoid the arrogance of subjective involvement. x Even for the zoologist [...] to avoid the arrogance of subjective involvement is difficult.

Of more immediate relevance to the present discussion, however, appear to be the subjects realized by noun phrases. Four of these are general categorial expressions (indefinite pronouns 8, 39, 48, and 67 *things*), which are not specific to any one type of text or any subject matter.

Six subjects of this group are anaphoric pronouns or zero (44, 47, 65, 66, 82, 97) constituting the second component of a simple linear progression, whose rhematic component occurs in the preceding clause as a context-independent complement of the verb (object, subject complement, adverbial), cf. 4.2.

4. Before proceeding to discuss the types of thematic progressions that occur in the text, all subjects need to be considered with respect to their FSP function. While Types 2 and 3 (progressions with a constant and derived theme) involve thematic subjects, Type 1 (simple linear progression) may have a subject with the FSP function of rheme. This is due to the configuration of the FSP functions required by this type: the thematic subject is preceded by a rhematic antecedent, which may appear in any syntactic function. Such configurations may also occur in Type 4 (progression with split rheme).

Both the largely thematic character of the English subject and its tendency not to vary in successive clauses suggest that thematic subjects are primarily to be sought among recurrent subjects.

4.1 Considered from the viewpoint of their realization form, all recurrent (identical) subjects implemented by anaphoric pronouns

are, unless given prominence prosodically, by a focalizer or a syntactic construction (e.g., a cleft sentence) inherently thematic; in the case of zero subjects the thematic function is the only possible one. Accordingly, when occurring in succession, these subjects induce Type 2 of thematic progression, progression with a constant theme. Fully identical subjects realized by noun phrases, whether by identical nouns or a more general term, may have either thematic or rhematic function. The former again give rise to progressions with a constant theme, but where the subject has rhematic function and is followed by a thematic one, the resulting configuration constitutes Type 1, simple linear progression.

Partly identical subjects present a similar picture: anaphoric pronouns (except for the instances specified above) and zero form basically induce Type 3 of TP, progression with derived themes. This type of progression appears to serve as a major device of thematic ramification. Partly identical subjects expressed by nouns, like fully identical substantival subjects, induce two types of thematic progressions. Where they have thematic function, they form Type 3 of TP, progression with derived theme, while subjects with rhematic function again give rise to Type 1 of TP, simple linear progression (see 4.3).

Apart from the FSP function of the subject, a relevant factor of thematic progression is the distance between clauses, i.e., the section of the text across which the notion denoted by the subject remains activated. In Daneš's framework of thematic progressions this question is not explicitly discussed. However, from his formulation

> TPs are often complicated by various insertions (supplements, explanatory notes) or asides. They may also occur in an incomplete or somewhat modified form. [...] Our types of TPs are to be considered as abstract principles, models, or constructs,[16]

immediate succession does not appear to be a necessary condition. Moreover, unspecified distances between clauses are indicated in some of his examples by dots [...].[17]

Nevertheless, the distance between clauses is relevant from the

---

[16] Daneš, "Functional Sentence Perspective," 121.
[17] "Functional Sentence Perspective," 119, 121.

viewpoint of the retrievability span,[18] i.e., the length of text within which individual components of the text build-up remain derivable from the preceding context. The retrievability span of thematic elements has been shown to extend, in general, to seven clauses.[19] Obviously, the retrievability of an item derives from all its occurrences in any syntactic function, not just from its being construed as the subject. In the case of *the black-footed squirrel* the retrievability span of seven clauses applies even when restricted to the subject: here all distances except one fall well within this span, the majority of clauses with identical subjects occurring in direct contact. In terms of the intervening clauses the distance was 1 to 4; the sole exception to the seven-clause retrievability span was reiteration of this identical subject after a gap of 15 clauses (43, 58). However, this has no effect on the retrievability span since this hypertheme remains activated by occurring in other syntactic functions.

4.1.1 As stated above, the hypertheme *black-footed squirrel* has the thematic function in all instances of pronominal and zero realization, e.g.:

Clauses 5, 6: 5 It [Th] has black feet and 6 it [Th] comes from Africa.

Clauses 22, 23: 22 it [Th] would be no more than a race of the basic species and 23 [zero/Th] could be swamped out, reabsorbed into the mainstream at any point.

Among the 12 instances realized by noun phrases seven are thematic (in clauses 3, 14, 19, 21, 25, 29, 41) and five rhematic (in clauses 4, 7, 31, 33, 36), e.g.:

Clause 3: This animal [Th] is new to science.

Clause 4: Inside the cage there sits a small squirrel [Rh].

---

[18] Firbas, *Functional Sentence Perspective in Written and Spoken Communication*, 23-24.

[19] The retrievability span as determined so far is subject to special contextual configurations which may extend it to much longer distances. Cf. Jan Firbas, "Substantiating Daneš's View of Givenness as a Graded Phenomenon," *The Syntax of Sentence and Text*, ed. Světla Čmejrková and František Štícha (Amsterdam and Philadelphia: John Benjamins, 1994) 119-29; Jan Firbas, "Retrievability Span in Functional Sentence Perspective," *Brno Studies in English* 21 (Brno: Masaryk University, 1995) 17-45; Libuše Dušková, "Some Thoughts on Potentiality in Syntactic and FSP Structure," *Proceedings of the Conference on the Dynamics of the Language System* (České Budějovice: Jihočeská univerzita) (forthcoming).

On the whole among the 30 instances of this hypertheme there are 25 thematic and 5 rhematic subjects.

4.1.2 A similar picture is presented by the hypertheme the *human species*: all pronominal and zero realizations serve as themes, cf.

Clauses 77, 78: [76 by approaching the human being] 77 as if he [Th] were another species, a strange form of life on the dissecting table, 78 [zero/Th] awaiting analysis.

Thematic subjects also prevail, even more prominently, among substantival subjects (in clauses 93, 94, 95, 96, 100); the only rhematic subject occurs in clause 98 (five and one instances, respectively). For example:

Clauses 93-96: 93 The legs [Th] are too long, 94 the arms [Th] are too short and 95 the feet [Th] are rather strange. 96 Clearly this species of primate [Th] has developed a special kind of locomotion …

Clause 98: There is yet another characteristic [Rh] (99 that [Th] cries out …)

Altogether, among the 12 instances of this hypertheme in the subject function 11 subjects are thematic and one subject is rhematic.

4.1.3 Both recurrent subjects conveying the local and temporal scene, *label* and *time,* have two occurrences each, one of which is thematic and the other rhematic:

Clauses 1, 2: 1 There is a label [Rh] on a cage at a certain zoo 2 that [Th] states simply

Clauses 24, 26: If, 24 as time [Th] passed, (25 the squirrels became more and more perfectly tuned-in to their particular environment,) 26 the moment [Rh] would eventually arrive [27 when it would be advantageous for them 28 to become …].

4.1.4 The recurrent subjects, the general *we* and *the zoologist* are all thematic as a result of their pronominal or zero form and initial position.

Clauses 37-38: 37 When we [Th] look at our unidentified squirrel in its zoo cage, 38 we [Th] can only guess about these things.

Clauses 70, 71, 73: Even for the zoologist 70 who [Th] is used 71 [zero/Th] to calling an animal an animal (72 it is difficult) 73 [zero/Th] to avoid the arrogance of subjective involvement.

4.1.5 As regards non-recurrent (different) subjects, although the feature of continuous recurrence is lacking, even these subjects have been found to be largely thematic.[20] As follows from Table 3 and the attached comment, the subjects relevant with respect to thematic progressions are the twelve subjects realized by noun phrases. Of these, ten are thematic and two rhematic. Apart from anaphoric pronouns (in clauses 44, 65, 66, 82, 97; five instances) and one zero realization (clause 47), thematic subjects are also realized by nouns and an indefinite pronoun (clauses 18 *the new trend*, 39 *all*, 61 *the great advantage*, 67 *things*). Indefinite pronouns are also the realization forms of the two non-recurrent subjects with the rhematic function (*nothing* in clause 8 and *something* in clause 48). Compare:

Clause 44: ... the rash that [Th] gives a doctor a clue

Clause 47: ... starting point [zero/Th] telling us

Clause 18: The new trend [Th] must have started out in a small way

Clause 8: Nothing [Rh] is known about it.

Clause 48: there is something [Rh] worth pursuing.

4.1.6 As noted in Section 1, the assignment of the FSP functions is based on the three factors operating in written language: linearity (word order), context in/dependence and semantics. In the case of thematic subjects the FSP function is largely indicated by the initial position and context dependence signalled by anaphoric devices (pronouns and zero).

In the case of rhematic subjects, whether recurrent or nonrecurrent (ten in total), final or postverbal position is found only where they appear after existential *there* – a special device for moving a rhematic subject to the postverbal part of the sentence. This was found in four instances:

Clause 1: There is a label [Rh] on cage at a certain zoo

Clause 4: Inside the cage there sits a small squirrel [Rh].

Clause 48: there is something [Rh] worth pursuing

Clause 98: there is another characteristic [Rh]

The six other rhematic subjects occur in the initial or a preverbal position. Here linearity is counteracted by semantic structure and context in/dependence. Let us first consider the fairly

---

[20] Cf. Dušková, *Studies in the English Language*, 2: 273-80.

straightforward instances in clauses 26, 36 and 31, 33.

Clause 26: the moment [Rh] would eventually arrive (27 when it would be ...)

Clause 36: At last, a new species [Rh] would have evolved, separate and discrete, ...

Here an important role is played by the semantics of the verb which expresses appearance on the scene. The subject in clause 26 (*the moment*) is partly context-independent in that it adds a specifying feature to the preceding concept of *time,* which assigns it a higher degree of communicative dynamism than is conveyed by a verb of this semantic class. In this configuration the subject acquires the function of rheme, but preserving its regular preverbal position, it displays a discontinuous structure as regards its postmodification, the *when*-clause, which constitutes the most dynamic part of the rheme, the rheme proper.[21] Clause 36 also contains a verb of appearance on the scene, and the novel aspect of the recurrent item *species* is indicated by the nongeneric indefinite article.

In contrast to clauses 26 and 36, the verbs in clauses 31 and 33 do not express existence or appearance on the scene; hence determination of the FSP functions of the clause elements rests on their degree of context in/dependence. The verbs in both clauses, *change* and *differ,* are derivable from *undergo special modification*s in clause 29. Clause 33 moreover contains a context-dependence indicator *also.* Although the subjects are partly context-dependent in being derived from the hypertheme *black-footed squirrel,* they contain novel features defining a particular aspect, which contrast with the preceding derived theme *their social and sexual behaviour* in clause 29, and are consequently more dynamic than the fully context-dependent verb.

Clause 31: At first their anatomy [Rh] may have changed

Clause 33: later their mating calls and displays [Rh] would also differ.

Clauses 7 and 8 have been classed as having rhematic subjects owing to the universal negator which constitutes, or is a component of, the subject. Negators as such largely act as rhematizers.

---

[21] Discontinous postmodification may serve as a device for reconciling the conflict between the grammatical principle and the principle of end focus (cf. Dušková, *Studies in the English Language,* 2:105-119).

Clause 7: <u>No black-footed squirrel</u> [Rh] has ever been found in that continent before.

Clause 8: <u>Nothing</u> [Rh] is known about it.

The subject in clause 7 constitutes the rheme only potentially. The negator implements the determiner of a context-dependent (recurrent) head noun, i.e., the subject is a composite element heterogeneous with respect to context-in/dependence. Having clausal scope, the negative determiner negates verbal polarity, which is in fact the most important novel feature, the rheme proper, of this clause. If we resort to prosody for indication of the carrier of the main intonation centre (nuclear tone), which as a rule falls on the rheme, we find three potential candidates: the subject, the lexical component of the verb *found* and the final adverbial *before*). As in other instances of potentiality,[22] in writing the FSP structure of the clause remains ambiguous.

Clause 8 is similar in that *Nothing*, having clausal scope, again negates the verb, the rheme proper being negative polarity. However, this subject is entirely context-independent and will in speech carry a stronger stress than the verb.

As regards the percentage of thematic and rhematic subjects in the group of recurrent subjects on the one hand, and in the group of different subjects on the other, it appears that recurrent thematic subjects are more frequent, viz. out of 80 recurrent subjects, 8 (10%) are rhematic and 72 (90%) thematic. (If the subject in clause 7 is assigned thematic function, the percentage of thematic subjects rises to 91.2.) In the case of different subjects, out of the 12 relevant for the present discussion (see Table 3 and Section 4.1.5), 2 (16.7%) are rhematic and 10 (83.3%) thematic. This result appears to confirm one of the starting points of this paper, the assumption of a close relationship between recurrence and thematicity.

4.2 As a final point the foregoing findings are considered with respect to the assumed connection between the tendency of the English subject not to vary and Type 2 of thematic progression.

4.2.1 In the examined text section clauses with identical subjects were found to form three types of thematic progression: progressions with a constant theme, progressions with a derived theme and simple linear progressions.

---

[22] Cf. Dušková, "Some Thoughts on Potentiality in Syntactic and FSP Structure."

4.2.1.1 The most frequent type is the progression with a constant theme. In the case of the general *we* and *the zoologist* it is the only type of progression registered. This is not accidental, since the absence of progressions with a derived theme follows from the full identity of these subjects in all their occurrences. The non-occurrence of the simple linear progression, on the other hand, is consequent only in the case of the general *we*, whose rhematic function, required as a constituent of this progression, is virtually ruled out by its semantics. This does not apply to *the zoologist*, whose more specific meaning allows it to appear in contrastive contexts, conducive to rhematic function.[23] Thematic progressions with a constant theme containing the general *we* and *the zoologist* account, respectively, for 31 and 3 clauses (*we* in clauses 37, 38, 40, 45, 46, 49, 50, 52, 53, 54, 55, 56, 57, 60, 62, 63, 64, 68, 69, 74, 75, 76, 79, 80, 81, 85, 86, 87, 88, 90, 91; *the zoologist* in clauses 70, 71, 73).

The two hyperthemes form, apart from progressions with a constant theme, also progressions with a derived theme and simple linear progressions. In the case of *black-footed squirrel*, progressions with a constant theme are again the most frequent; they are found in 13 clauses (clauses 3, 5, 6, 9, 10, 13, 19, 20, 21, 22, 23, 25, 28). As regards *the human species*, progressions with a constant theme are as frequent as progressions with a derived theme, each accounting for 4 clauses, the former in clauses 77, 78, 83, 96.

Accordingly, the total number of clauses forming thematic progressions with a constant theme is 51. For example:

Clauses 45, 46, 49, 50, 52, 53, 54: 45 To really [zero/Th] understand this new species, 46 we [Th] must use these clues only as a starting point [47 telling us 48 there is something worth pursuing]. 49 We [Th] might try 50 [zero/Th] to guess at the animal's history [51 but that would be presumptuous and dangerous]. 52 Instead we [Th] will start humbly 53 by [zero/Th] giving it a simple and obvious label: we [Th] will call it the African black-footed squirrel.

Clauses 70, 71, 73: 70 (for the zoologist) who [Th] is used 71 to [zero/Th] calling an animal an animal, [72 it is difficult] 73 [zero/Th] to avoid the arrogance of subjective involvement.

---

[23] The difference in this respect between the two notions is shown by the in/applicability of a focalizer, cf. *even a zoologist*, but **even we* (applicable only to 1st person plural, in contrast to other persons, cf. *even I*; but not to the general sense 'one').

Clauses 9, 10: 9 <u>It</u> (= *the black-footed squirrel*) [Th] has no name. 10 For the zoologist <u>it</u> [Th] presents an immediate challenge.

Clauses 77, 78, 83, 96: 77 as if <u>he</u> (= *the human species*) [Th] were another species, a strange form of life on the dissecting table, 78 [<u>zero</u> Th] awaiting analysis. ... 83 <u>he</u> [Th] is obviously a primate of some sort. ... Clearly <u>this species of primate</u> [Th] has developed a special kind of locomotion.

4.2.1.2 Progressions with a derived theme rank second in the frequency of occurrence among the clauses with recurrent subjects. Being based on partly identical subjects, they are found only with the two hyperthemes, *the black-footed squirrel* in 7 instances (in clauses 14, 15, 29, 30, 43, 58, 59; within these clauses the derived themes form progressions with constant theme); and *the human species* in 4 clauses (93, 94, 95, 100). Altogether, progressions with a derived theme account for 11 clauses. For example:

Clauses 29, 30: 29 At this stage <u>their</u> (= *the black-footed squirrels'*) <u>social and sexual behaviour</u> [derived Th] 30 would undergo special modifications, [<u>zero</u>/derived Th] making interbreeding with other kinds of squirrels unlikely and eventually impossible.

Clauses 93, 94, 95, 100: 93 <u>The legs</u> (= of *the human species*) [derived Th] are too long, 94 <u>the arms</u> [derived Th] are too short 95 and <u>the feet</u> [derived Th] are rather strange. ... 100 <u>the skin</u> [derived Th] is virtually naked.

4.2.1.3 As regards simple linear progression, apart from the two hyperthemes (*the black-footed squirrel* in clauses 4, 5 (6); 31, 32; 33, 34 (35); *the human species* in clauses 98,99) it is also found with the scene-setting subject *label* (in clauses 1, 2); altogether in 10 clauses. In the group of identical subjects it appears to have a comparable frequency of occurrence with the sequence with a derived theme (11 clauses, cf. 4.2.1.2). For example:

Clauses 4, 5 (6): 4 Inside the case there sits <u>*a small squirrel*</u> [Rh]. 5 <u>It</u> [Th] has black feet (and 6 <u>it</u> [Th] comes from Africa).

Clauses 33, 34 (35): 33 <u>their mating calls and displays</u> [Rh] would also differ, 34 [<u>zero</u>/ Th] ensuring (that 35 <u>they</u> [Th] attract only mates of the other type).

Clauses 98, 99: 98 But there is <u>another characteristic</u> [Rh], 99 <u>that</u> [Th] cries out for attention:

Clauses 1, 2: 1 There is <u>a label</u> [Rh] on a cage at a certain zoo 2 <u>that</u> [Th] states simply

As shown through examples with a clause number in brackets, the simple linear progressions often open progressions with a constant theme.

4.2.2 Clauses with different subjects (cf. Table 3) provide favourable conditions for thematic progressions only where the subjects have non-subject identical antecedents (the other potential configuration with non-subject identical postcedents does not occur in the text section under study). Sequences of different subjects with non-subject antecedents are found in clauses 43, 44; 46, 47; 64, 65, 66; 81, 82; 96, 97. In all these clauses the subjects are anaphoric devices (pronouns or zero) with the thematic function whereas their non-subject antecedents occur in the rheme. Accordingly, the resulting sequences constitute simple linear progressions, cf.

Clauses 43, 44: 43 But these are only the symptoms, <u>the rash</u> [Cs=Rh] 44 <u>that</u> [S=Th] gives a doctor a clue about his patient's disease.

Clauses 46, 47: 46 we must use these clues only <u>as a starting point</u> [Co=Rh] 47 [S/<u>zero</u>=Th] telling us ...

Clauses 64, 65, 66: 64 we ourselves are not black-footed squirrels – <u>a fact</u> [appositive of Cs=Rh] 65 <u>which</u> [S=Th] forces us <u>into an attitude of humility</u> [O$_{prep}$= Rh] 66 <u>that</u> [S=Th] is becoming to proper scientific investigation.

Clauses 81, 82: 81 by comparing him <u>with other species</u> [O$_{prep}$=Rh] 82 <u>that</u> [S=Th] appear to be most closely related.

Clauses 96-97: 96 this species of primate has developed <u>a special kind of locomotion</u> [O=Rh] 97 <u>which</u> [S=Th] has modified its basic form.

It should be noted that while simple linear progressions mostly terminate with the second component or open progressions with a constant theme, clause 65, whose subject forms the second (thematic) component of the simple linear progressions 64, 65, contains a rhematic prepositional object with a thematic subject postcedent in clause 66, so that the three clauses give rise to a sequence of two simple linear progressions.

These instances (11 clauses in total) do not in principle differ from simple linear progressions found among recurrent subjects in

that both types comply with the required sequence of FSP functions: a rhematic element is followed by a co-referential thematic one. As a result, the total number of clauses forming simple linear progressions amounts to 21.

Considering the frequency of occurrence of thematic progressions on this basis and taking clauses with both identical and different subjects together, the first place on the frequency scale belongs to progressions with a constant theme (owing to the prevalence of identical subjects; 51 clauses), with simple linear progressions ranking second (21 clauses), and progressions with a derived theme in the last position (11 clauses). This leaves 17 clauses to be accounted for.

4.2.3 This group contains clauses with both identical and different subjects.

4.2.3.1 The group of identical subjects displays three clauses that appear to stand outside thematic progressions: clauses 36 and 24, 26.

Clause 36: At last, a new species would have evolved, separate and discrete, a unique form of life, a three hundred and sixty-seventh kind of squirrel

Clauses 24, 26: If, as time [Th] passed, (25 the new squirrels became more and more perfectly tuned-in to their particular environment,) the moment [Rh] would eventually arrive when ...

Clause 36 is a clear instance of a presentation scale, with a verb of appearance on the scene and a rhematic subject, within a progression of derived themes. Consequently, it stands outside the progression.

Clauses 24, 26 form a sequence of two partly identical subjects, the first having the function of theme, while the second constitutes the rheme, again in a clear case of a presentation scale. A configuration of this kind fails to align itself with a thematic progression.

4.2.3.2 As in the foregoing section, out of all the different and excluded subjects, listed in Table 3, only those in the first column are considered here. The excluded clauses in the second column comprise, besides those constituting parts of syntactic constructions

(discussed in Section 3) two clauses (51 and 84) whose subjects are clause equivalents. The role of all these devices in the text build-up and theme development remains to be studied.

The clauses containing anaphoric subjects with non-subject rhematic antecedents have been added to simple linear progressions (cf. 4.2.2). The remaining nonrecurrent subjects, those realized by nouns and indefinite pronouns (in clauses 8, 18, 39, 48, 61, 67) appear to stand outside thematic progressions. It should be noted that except clause 67 the clauses with different thematic subjects, which considerably prevail over the different rhematic ones, are the most informative because they contain a novel item even in the subject.

Clause 8: Nothing [Rh] is known about it.

Clause 18: The new trend [Th] must have started out in a small way, …

Clause 39: All [Th] (we can be sure about) is that …

Clause 48: … there is something [Rh] worth pursuing.

Clause 61: The great advantage [Th] … is that …

Clause 67: How different things [Th] are, how depressingly different (68 when we attempt to study the human animal.)

The textual role of these instances also calls for further study.

4.2.4 The relative frequency of occurrence of the thematic progressions found in the text under study has been given in terms of clauses for two reasons. First, as appears from some of the examples, progressions with derived themes and simple linear progressions often combine with progressions with constant themes. Secondly, in strings of progressions with a constant theme it may be difficult to determine the terminal component where the clauses containing identical subjects cease to stand in immediate succession.

Another point to be noted in respect of the quantitative data is that since thematic progressions also occur among clauses with different subjects, there is no correlation between the number of clauses with identical and different subjects on the one hand, and the number of clauses forming thematic progressions on the other.

5. In conclusion, the foregoing discussion may be summed up as follows. The ascertained quantitative data appear to confirm the assumption that the tendency of English to preserve the same subject in successive clauses, together with its largely thematic character, provides favourable conditions for thematic progressions with a constant theme. However, this is largely due to the recurrence of the general *we*, which has been characterized as external to the subject matter dealt with in the text. Moreover, it is at variance with a prominent feature of formal academic writing – frequent use of the passive voice. As shown in previous studies (of the passive voice and *man-Sätze*),[24] the active with the general *we* as the subject and the passive are often interchangeable. In the text under study, the general *we* is used as a starting point partly owing to its inclusive meaning, but more importantly because it allows context-independent, rhematic elements, construed in the active as a verbal complement, to stand in the regular postverbal position, and thus to comply with the principle of end focus. It is in this clausal position that the subject matter – in our text the two hyperthemes *the black-footed squirrel* and *the human species* – is developed. In this respect, the use of the general *we* appears to be an important device in the text build-up.

However, the results obtained from one text of limited length cannot be generalized. A cursory glance at another text that could not be included for considerations of space has displayed frequent occurrence of Type 4 of thematic progression, progression with split rheme. This suggests that the strategy and devices employed in the text build-up depend both on the subject matter and the author's approach to it.

Nevertheless, some general points about thematic progressions have emerged even from this limited probe: simple linear progressions appear to suit the opening part of a text or smaller textual units, which introduces the subject matter to be dealt with. For this reason, these progressions mostly consist of only two members, the first rhematic and the second thematic. Elaboration of the subject matter largely depends on its nature, which may favour any one of the other thematic progressions: with a constant or derived theme, with split rheme, or their combinations. Of these, progressions allowing the largest number of members are to be

[24] Cf. Dušková, *Studies in the English Language*, 1: 113-48; 2: 41-86.

sought among progressions with a constant theme. Overall, a complete picture of the role of thematic progressions in the text build-up calls for an analogous examination of the concluding part of text, which was outside the scope of the present discussion.

APPENDIX

Table 1

	Subjects with More Than One Occurrence		Subjects with One Occurrence	
Clause	Finite	Nonfinite	Finite	Nonfinite
1 there is	label[1]			
2 states	that[1]			
3 is	this animal[2]			
4 there sits	A small squirrel[2]			
5 has	it[2]			
6 comes	it[2]			
7 has been found	no blackfooted squirrel[2]			
8 is known			Nothing	
9 has	it[2]			
10 presents	it[2]			
11 is	it (it-cleft)			
12 has made	that (it-cleft)			
13 does it differ	it[2]			
14 must have split	the ancestors of this animal[2a]			
15 and established	zero (ellipsis)[2a]			
16 was	it (it-cleft)			
17 made	that (it-cleft)			
18 must have started			the new trend	
19 becoming		with a group of squirrels[2]		
20 would be	they[2]			
21 would be	the new form[2]			
22 would be	it[2]			
23 could be	zero (ellipsis)[2]			
24 passed	Time[3]			

25 became	the new squirrels[2]			
26 would arrive	the moment[3]			
27 would be	It (anticipatory)			
28 to become		for them[2]		
29 would undergo	Their social and sexual behaviour[2b]			
30 making		zero[2b]		
31 may have changed	Their anatomy[2c]			
32 at coping		zero[2c]		
33 would differ	Their mating calls and displays[2d]			
34 ensuring		zero[2d]		
35 attract	they[2d]			
36 would have evolved	A new species[2]			
37 look	we[4]			
38 can guess	we[4]			
39 is			All	
40 can be	we[4]			
41 indicate	the markings of its fur – its black feet[2e]			
42 is	it[2]			
43 are	These[2e]			
44 gives			(the rash) that	
45 to understand		zero[4]		
46 must use	we[4]			
47 telling				(starting point) zero
48 there is			something	
49 might try	we[4]			
50 to guess		zero[4]		
51 would be			that (clausal proform)	
52 will start	we[4]			
53 by giving		zero[4]		
54 will call	we[4]			
55 must observe	we[4]			
56 (must) record		zero[4] (ellipsis)		
57 (must) see		zero[4] (ellipsis)		

58 differs	it[2f] (every aspect of its behaviour)		
59 is		zero[2f] (ellipsis)	
60 can piece	we[4]		
61 is			the great advantage
62 have	we[4]		
63 studying		zero[4]	
64 are not	we[4]		
65 forces			(a fact) which
66 is becoming			(humility) that
67 are			things
68 attempt	we[4]		
69 to study		zero[4]	
70 is used	(the zoologist) who[5]		
71 to calling		zero[5]	
72 is	It (anticipatory)		
73 to avoid		zero[5]	
74 can try	we[4]		
75 to overcome		zero[4]	
76 by approaching		zero[4]	
77 as if were	he[6] (the human being)		
78 awaiting	zero[6] (a strange form of life)		
79 can begin	we[4]		
80 can start	we[4]		
81 by comparing		zero[4]	
82 appear to be			(other species) that
83 is	he[6] (the human being)		
84 becomes			just how odd (subject clause)
85 lay out	we[4]		
86 try	zero (ellipsis)[4]		
87 to insert		zero[4]	

88 put	we[4]			
89 looks	(the human pelt) it[6a]			
90 are driven	we[4]			
91 to position		zero[4]		
92 is	it[6]			
93 are	the legs[6b]			
94 are	the arms[6c]			
95 are	the feet[6d]			
96 has developed	this species[6]			
97 has modified			(locomotion) which	
98 there is	Another characteristic[6a]			
99 cries out	that[6a]			
100 is	the skin[6a] (=another characteristic)			

Desmond Morris, *The Naked Ape*, pp. 13-15

There (**1**) <u>is</u> a label on a cage at a certain zoo that (**2**) <u>states</u> simply, 'This animal (**3**) <u>is</u> new to science.' Inside the cage there (**4**) <u>sits</u> a small squirrel. It (**5**) <u>has</u> black feet and it (**6**) <u>comes</u> from Africa. No black-footed squirrel (**7**) <u>has</u> ever <u>been found</u> in that continent before. Nothing (**8**) <u>is known</u> about it. It (**9**) <u>has</u> no name.

For the zoologist it (**10**) <u>presents</u> an immediate challenge. What (**11**) <u>is</u> it about its way of life that (**12**) <u>has made</u> it unique? How (**13**) <u>does</u> it <u>differ</u> from the three hundred and sixty-six other living species of squirrels already known and described? Somehow, at some point in the evolution of the squirrel family, the ancestors of this animal (**14**) <u>must have split off</u> from the rest and (**15**) <u>established</u> themselves as an independent breeding population. What (**16**) <u>was</u> it in the environment that (**17**) <u>made</u> possible their isolation as a new form of life? The new trend (**18**) <u>must have started out</u> in a small way, with a group of squirrels in one area (**19**) <u>becoming</u> slightly changed and better adapted to the particular conditions there. But at this stage they (**20**) <u>would</u> still <u>be</u> able to inter-breed with their relatives nearby. The new form (**21**) <u>would be</u> at a slight advantage in its special region, but it (**22**) <u>would be</u> no more than a race of the basic species and (**23**) <u>could be swamped out</u>, reabsorbed into the mainstream at any point. If, as time (**24**) <u>passed</u>, the new squirrels (**25**) <u>became</u> more and more perfectly tuned-in to their particular environment, the moment (**26**) <u>would</u> eventually <u>arrive</u> when it (**27**) <u>would be</u> advantageous for them (**28**) <u>to become</u> isolated from

possible contamination by their neighbours. At this stage their social and sexual behaviour (29) <u>would undergo</u> special modifications, (30) <u>making</u> inter-breeding with other kinds of squirrels unlikely and eventually impossible. At first, their anatomy (31) <u>may have changed and become</u> better at (32) <u>coping</u> with the special food of the district, but later their mating calls and displays (33) <u>would</u> also <u>differ,</u> (34) <u>ensuring</u> that they (35) <u>attract</u> only mates of the new type. At last, a new species (36) <u>would have evolved,</u> separate and discrete, a unique form of life, a three hundred and sixty-seventh kind of squirrel.

When we (37) <u>look</u> at our unidentified squirrel in its zoo cage, we (38) <u>can</u> only <u>guess</u> about these things. All we (40) <u>can be</u> certain about (39) <u>is</u> that the markings of its fur – its black feet – (41) <u>indicate</u> that it (42) <u>is</u> a new form. But these (43) <u>are</u> only the symptoms, the rash that (44) <u>gives</u> a doctor a clue about his patient's disease. (45) <u>To</u> really <u>understand</u> this new species, we (46) <u>must use</u> these clues only as a starting point, (47) <u>telling.us</u> there (48) <u>is</u> something worth pursuing. We (49) <u>might try</u> (50) <u>to guess</u> at the animal's history, but that (51) <u>would be</u> presumptuous and dangerous. Instead we (52) <u>will start</u> humbly by (53) <u>giving</u> it a simple and obvious label: we (54) <u>will call</u> it the African black-footed squirrel. Now we (55) <u>must observe</u> and (56) <u>record</u> every aspect of its behaviour and structure and (57) <u>see</u> how it (58) <u>differs</u> from, or (59) <u>is</u> similar to, other squirrels. Then, little by little, we (60) <u>can piece</u> together its story.

The great advantage we (62) <u>have</u> when (63) <u>studying</u> such animals (61) <u>is</u> that we ourselves (64) <u>are not</u> black-footed squirrels – a fact which (65) <u>forces</u> us into an attitude of humility that (66) <u>is</u> becoming to proper scientific investigation. How different things (67) <u>are,</u> how depressingly different, when we (68) <u>attempt</u> (69) <u>to study</u> the human animal. Even for the zoologist, who (70) <u>is used</u> (71) <u>to calling</u> an animal an animal, it (72) <u>is</u> difficult (73) <u>to avoid</u> the arrogance of subjective involvement. We (74) <u>can try</u> (75) <u>to overcome</u> this to some extent by deliberately and rather coyly (76) <u>approaching</u> the human being as if he (77) <u>were</u> another species, a strange form of life on the dissecting table, (78) <u>awaiting</u> analysis. How (79) <u>can</u> we <u>begin?</u>

As with the new squirrel, we (80) <u>can start</u> by (81) <u>comparing</u> him with other species that (82) <u>appear to be</u> most closely related. From his teeth, his hands, his eyes and various other anatomical features, he (83) <u>is</u> obviously a primate of some sort, but of a very odd kind. Just how odd (84) <u>becomes</u> clear when we (85) <u>lay out</u> in a long row the skins of the one hundred and ninety-two living species of monkeys and apes, and then (86) <u>try</u> (89) <u>to insert</u> a human pelt at a suitable point somewhere in this long series. Wherever we (88) <u>put</u> it, it (89) <u>looks</u> out of place. Eventually we (90) <u>are driven</u> (91) <u>to position</u> it right at one end of the row of skins, next to the hides of the tailless great apes such as the chimpanzee and the gorilla. Even here it (92) <u>is</u> obtrusively different. The legs (93) <u>are</u> too long,

the arms (**94**) <u>are</u> too short and the feet (**95**) <u>are</u> rather strange. Clearly this species of primate (**96**) <u>has developed</u> a special kind of locomotion which (**97**) <u>has modified</u> its basic form. But there (**98**) <u>is</u> another characteristic that (**99**) <u>cries out</u> for attention: the skin (**100**) <u>is</u> virtually naked.