

LEONHARD LIPKA – HANS-JÖRG SCHMID

***To begin with*: degrees of idiomaticity, textual functions and pragmatic exploitations of a fixed expression**

Abstract: In a context like "The story only seems to begin with" the word sequence *to begin with* clearly has a meaning which differs from its typical sentence-initial use as in "To begin with, let me say that..." The first example will probably be understood "literally" while the second will be interpreted as a fixed expression or idiom. It is argued that instead of a dichotomy between the literal and idiomatized reading of *to begin with* there is a gradient of idiomaticity. Various degrees of idiomaticity are illustrated with examples taken from the Lancaster-Oslo-Bergen corpus and described in terms of additional semantic components, textual functions and pragmatic exploitations. It is shown that *to begin with* is an important text-structuring device, and that speakers seem to use *to begin with* as a kind of "linguistic threatening gesture". Finally, the semantic and functional aspects of *to begin with* are compared to the related expressions *to start with* and *for a start*, with reference to an earlier analysis of the semantic difference between *begin* and *start* (Schmid 1993).

1. Introduction

Since the advent of the prototype theory of categorization in the seventies there has been a strong tendency to apply the concepts of prototype categories and family resemblances to linguistic terminology. The result of this widespread enthusiasm for fuzzy category boundaries has been a revision of all kinds of traditionally clear-cut classifications or dichotomies (e. g. phonemes, morphemes, grammatical categories, intonation patterns; cf. e. g. Taylor 1989, Tsohatzidis 1990) in terms of scales and gradients. While this development should certainly be welcomed, because prototype theory does indeed seem to capture the essence of categories more adequately than classical definitions, the new approach also holds the danger of encouraging what Wierzbicka has denounced as "intellectual laziness and sloppiness" (1990: 365). In this paper we will address the traditional distinction between simple lexemes and fixed expressions or idioms. We will use the example of the fixed expression *to begin with* to show that instead of a dichotomy between literal meaning and idiomatic use there are various degrees of idiomaticity. Taking heed of Wierzbicka's criticism of the abuses of prototype theory, we will use an array of semantic, syntactic, textual and pragmatic considerations in order to specify the scale of idiomaticity rather than merely state that it exists.

2. Complex expressions with *start* and *begin*: a scale of idiomaticity

The material for this study is taken from the Lancaster-Oslo-Bergen corpus of English (LOB). In a number of other empirical studies which used the LOB for a database (cf. Leitzke 1989, Grimm 1991) it has been regretted that the collected material is not extensive enough for many purposes. The same quantitative limitation of the LOB manifested itself in the present research: while *to begin with* was at least recorded 11 times in the LOB, *for a start* was found only 3 times and *to start with* did not occur at all in this collection of about 1 million words. Therefore, additional examples were collected from grammar textbooks, dictionaries, and from an introductory textbook on Cognitive Psychology (Sanford 1985). In what follows, we will

focus our attention on the expression *to begin with*, and relate *to start with* and *for a start* to these findings in section 4.

To some readers, the claim that *to begin with* exhibits semantic idiosyncrasies which are caused by the fact that it has undergone a process one might call lexicalization may seem rather premature. And indeed, examples of the word sequence *to begin with* in the LOB can be provided which reveal this claim to be largely unwarranted. Example (1) is a case in point:

- (1) [This] is often ignored so that the story only seems to begin with the otherwise inexplicable lorry-forays against red centres ... (LOB J 56 5)

In this example, any assertion that *begin* or *to begin with* acquires additional semantic elements over and above the prototypical attribute 'inchoative' would run into serious difficulties, because neither syntactic nor semantic clues can be obtained. It seems, then, that in example (1) *begin* is used in a construction which, by mere coincidence, is formally identical with the fixed expression *to begin with*.

Now, by way of comparison, consider example (2).

- (2) I would like to begin with the less developed members of the commonwealth and those territories which are still dependent. (LOB H 21 20)

At first sight, it again seems hardly convincing to call the sequence *to begin with* a lexicalized fixed expression. And yet, one may feel that an additional semantic element is present in (2) which is missing in (1). Hearing a speaker start his speech or lecture with the words *I would like to begin with* we are coaxed into assuming that this is only the beginning of a whole list of issues to be dealt with. From a transphrastic point of view, the sentence functions as an introduction to or an announcement of what is to come later, and *to begin with* contributes a major part to performing this function. More specifically, the 'announcing function' of (2) can be located in the meaning of *begin* (something that begins usually also goes on in some way) and in the fact that some sort of locutive object like *I would like to begin my speech with ...* has been omitted. It thus seems that the sequence *to begin with* in (2) comes a little closer to an idiomatized use than it does in (1), because it conveys one additional semantic element, which can roughly be glossed 'announcing that there is more to come'. Also, the omission of the locutive object is part of the idiomatization process.

Superficially, it is only a small step from example (2) to intuitively more idiomatized uses of *to begin with* such as (3) or (4).

- (3) Intention, I would suggest to begin with, is a term which is applicable when a certain roughly specifiable complex of conditions hold. (LOB G 63 139)
- (4) To begin with, the resurrection is held to be the revelation of the mystery of redemption, the open demonstration of God's saving activity to which all previous sacred history has been leading. (LOB D 08 150)

However, a closer look at the examples reveals a number of interesting observations. As in example (2) above, sentences (3) and (4) are preparing us for a list of several items. It could therefore be conjectured that the announcing function is a constant semantic element of the phrase *to begin with*. Indeed, many dictionaries in some way or other capture this idea of the "first of a number of points". Compare for example (5) and (6):

- (5) *To begin with*: 'the first reason is'
To begin with, it's too cold, and besides, we have no money. (LDCE)
- (6) You use the expression *to begin with* [...] to introduce the first of a number of things that you want to say. (COBUILD)

These paraphrasing dictionary entries are not immune to criticism. In particular, the use of the word "reason" in the entry in the LDCE is disputable. While a causative meaning element might perhaps be postulated for (4) and the example given in the LDCE, for (3) and many other examples it appears to be out of place. In addition, one might be inclined to argue over

the use of the word "first" in these dictionary entries. Although for all practical purposes, the intention of the respective lexicographers seems to be transparent enough, theoretically the relation between a *beginning* and a *first entity* is far from clear. After all, it is not inconceivable that something could be started with, say, the third point in a list.¹ On the other hand, the lexicographers could reply that by virtue of the very fact that a certain point out of many is put at the start of e. g. a lecture, it in fact becomes the first point. This seems to be a rather elegant way of settling the unclear status of the word *first*. All the same, we must not forget that example (2), and for that matter also example (3), have taught us that as often as not the sentence containing the expression *to begin with* does not contain the first point of a list, but introduces or announces the list. We may therefore conclude that in contrast to literal uses as represented by example (1), *to begin with*, as a fixed expression, has acquired one of the two semantic elements 'announcing a list of items' or 'first point of a list of items'. It thus functions as a discourse signal.

A second point which distinguishes the idiomatized uses of *to begin with* from other examples is their formal invariability. Whereas above, it proved perfectly possible to expand example (2) from *I would like to begin with ...* to *I would like to begin my speech with ...* similar insertions into the expression *to begin with* are not feasible in (3) and (4). In fact, the formal invariability of a fixed expression is one criterion for idiomaticity which has been put forward by Russian linguists like I. A. Mel'cuk and N. N. Amosava (cf. Lipka (1974)) and which should always be kept in mind.

Thirdly, it must be noted that in idiomatic uses of *to begin with*, *begin* does not of course function as the main verb of the clause in which it occurs, as is the case in literal uses such as (1) or (2). Instead, the whole fixed expression *to begin with* is usually extracted from the clause pattern and functions as sentence adverbial (cf. (4)) or is put in some kind of commenting parenthesis (cf. (3) and section 3.1).

Finally, as a fixed expression *to begin with* exhibits phonological properties which do not occur in non-idiomatized and half-idiomatized uses such as (1) and (2) respectively. In (1) and (2), the preposition *with* is undoubtedly of very low intonatory prominence; in addition, its final fricative will, in assimilation with the ensuing homorganic /ð/, obtain a voiced quality. In contrast, the comma in example (4) indicates a pause between the two sounds, separating the expression *to begin with* from the rest of the sentence. Listening to native people reading out sentences containing *to begin with*, one also realizes that they tend to emphasize the whole phrase by ending it with a slight rise in pitch on the word *with*. Concomitantly, the second syllable of the word *begin* acquires a much higher degree of tonic prominence than in (1) and (2), because the phrase *to begin with* functions as a tone unit in its own right. Similar phonological characteristics apply for the word sequence inserted between commas (*I would suggest to begin with*) in example (3).

We have so far singled out some semantic, formal, syntactic and phonological properties of the sequence *to begin with*, which justify our claim that it is an idiomatized or fixed expression. In addition, the fact that there are uses of *to begin with* which have only acquired one additional semantic element (cf. (2)) has shown that instead of a two-fold distinction between "literal" and "idiomatic", there is a gradient of idiomaticity running from clearly literal to more or less idiomatic uses.²

The examples we have provided so far give evidence that in idiomatic uses, *to begin with* acquires additional functions which necessitate a step beyond simple semantic and syntactic considerations. In particular, we have seen that *to begin with* has an important textual function and is used as a discourse signal. In the following section we will discuss this textual function and try to show how it is exploited by speakers for pragmatic purposes (in a literal and a linguistic sense).

3. *To begin with* as a discourse signal

3.1. Textual functions

It has already been mentioned that *to begin with* is frequently used in an announcing function. What the discourse signal *to begin with* eventually announces or introduces, and how it achieves this function, varies to a considerable extent. In example (3) the parenthesis containing *to begin with* serves as a kind of heading under which a number of arguments (which we have not provided) are ordered. As we have said, in these cases *to begin with* does not constitute the first element of a list, but introduces the list which follows. This has not been recognized by lexicographers.

In other examples in our corpus, *to begin with* does accompany the first point in a list of items, e. g. in (7):

(7) The new book [...] is a case in point. The book is, to begin with, curiously non-critical.

[...] There is, further an irreducible sentimentalism about the book. (LOB G 36 144)

Here, the first sentence of the paragraph provides the heading while the sentence containing *to begin with* already mentions the first argument.

As far as the content of the "list" is concerned, the variation is still greater. While it is true that sometimes a series of reasons is given (cf. (5)), this is not necessarily the case. In fact, *to begin with* can also introduce lists of other items such as points to be discussed or arguments put forward in favour of an assertion. This can be illustrated by providing more context for example (4) quoted above:

(8) These stages of development reflect the Church's changing outlook. To begin with, the resurrection is held to be the revelation of the mystery of redemption, the open demonstration of God's saving activity to which all previous sacred history has been leading. Secondly, the teaching of Jesus is held to be an essential part of the revelation, though its true significance was only known to the "elect". [...]. The third and final stage is the claim that [...]. (LOB D 08 150)

Example (8) gives a good idea of how speakers, or more often writers, employ *to begin with* in order to structure their textual output. After the topic has been introduced in the first sentence of the paragraph, *to begin with* serves as a marker for the first of the stages mentioned in the introduction, which is later taken up by the words *secondly* and *the third and final stage*. In cases like this, Keller (1981: 97f) speaks of the "semantic framing" of the "topic" by means of verbal signals which he calls "gambits". However, it is not really the topic that is structured by these signals, but the message itself, and in our opinion, the term "semantic framing" seems a little pompous for the simple establishment of a list of items. We will therefore steer clear of Keller's terminology altogether and use the more general term "discourse marker" introduced by Schiffrin (1987).

To sum up the discussion so far, we have seen that from a textual point of view *to begin with* can mark either the announcement or the actual beginning of a list of various kinds of items. In cooperation with other textual signals it therefore enhances the internal structure of texts and creates text cohesion. In the by now classic account of cohesion presented by Halliday & Hasan (1976), *to begin with* and the like are treated as surface signals for an underlying semantic relation labelled "conjunction". According to Halliday & Hasan, conjunction appears in an "internal" and an "external" variety. In essence, the distinction captures the recognition that sometimes conjunctive items only concern the way speakers present their discourse (cf. (9)),³ and sometimes they relate real events to each other (cf. (10)).

(9) Next, he was incapable of inserting the key into the lock.

(10) Next, he inserted the key into the lock. (Halliday & Hasan 1976: 239)

In all the examples we have quoted so far, *to begin with* belongs to the internal type of conjunc-

tion and indeed in the lists provided by Halliday and Hasan (1976: 267) it occurs under "temporal relations (internal)" only. Yet the data in our corpus clearly show that *to begin with* is used in external function as well, as in (11):

(11) By degrees. Centigrade v. Fahrenheit.

The fight is on. *To begin with*, both temperatures will be put in the ring together.

The press and the broadcasting authorities are asked to help. (LOB B 03 160)

In this example, *to begin with* is not just an element of the "speaker's organization of his discourse" (Halliday & Hasan 1976: 239), but it also describes the order of the events as they occur in reality. Somewhat more confusing is example (12), which epitomizes the use of *to begin with* favoured by Sanford in his introductory textbook.⁴

(12) *To begin with*, we shall consider some important data on learning which can be interpreted in terms of the working memory hypothesis. Then we will move on to more specific ideas which have been associated with the working memory concept.

Comparing this example to (8), we realize that here *to begin with* is not used to put a structural grid on the ongoing part of the text, but rather on the passages that are to follow. This is reflected in the use of the main verb of the clause (*we shall consider ...*). So, although *to begin with* in (12) also refers to the organization of language, it differs from (8) in that a future linguistic act is predicted. In (8) it is part of the structuring of the immediate text passage itself, while in (12) *to begin with* refers to the structuring of the later text.

The difference between the internal and the external use of *to begin with* as a conjunctive signal is paralleled both in the dominant language function and in the type of speech act that is performed. The internal use exemplified by (8) is closely related to the interpersonal function of language, since it reflects "the speaker's own 'stamp' on the situation" (Halliday & Hasan 1976: 240). In effect, *to begin with* here serves as a metalinguistic structural device. In contrast, the *to begin with* in (12) is a linguistic element that reflects the experiential function of language. Likewise we could argue that, applying Searle's (1976) taxonomy of speech acts, the utterance in (8) must be assigned to the group of representatives, because it is a statement. The first utterance in (12), on the other hand, makes predictions on a future act of the speaker and must therefore be classified as a commissive. So on a deeper level, the two seemingly so similar occurrences of *to begin with* in (8) and (12) reveal a number of interesting differences.

How does the position of *to begin with* in the sentence in which it occurs tie in with its textual function? Generally speaking, *to begin with* can be put into an initial, a medial or a final position. The data in our corpus indicate that while no correlations between position and the type of textual function can be established, there is a clear preference to use *to begin with* sentence-initially or, in parentheses, right after the subject as in (3) above. And this is of course hardly surprising, for a number of reasons. Since the most frequent syntactic function fulfilled by *to begin with* is that of a sentence conjunct (in the sense of Quirk et al. 1985), it tends to be found in the canonical conjunct position, i. e. at the beginning of the sentence. From a textual point of view it is again only sensible to put structural signals sentence-initially, because this is obviously a perceptually salient position. If a speaker envisages a strong structural grid of the type *To begin with ... Secondly ... Thirdly ...* etc., it is recommendable to employ these discourse markers at the beginning of their respective sentences.

3.2 Pragmatic exploitations

In approximately 50 percent of the examples of *to begin with* that we have collected, the announcing function misfires, or in other words, *to begin with*, is not followed by a list. This raises the question whether in these cases speakers pursue other aims by using *to begin with* and if so, what they are. We will put forward two possible ways in which speakers take advantage of the textual function of *to begin with* for other purposes.

Firstly, there can be no doubt that English speakers are well aware of the fact that not every *to begin with* is followed by more points. On the other hand, they nevertheless seem to assume that somehow or other *to begin with* is often or at least should in fact be followed by one or more further items or arguments. So, even though as speakers they tend to neglect the list themselves, as hearers or readers they expect some sort of continuation all the same. Thus, one might say that *to begin with* evokes an expectant state of mind in the hearer, who believes that there is more to come. It can therefore be exploited as a so-called "incompletion marker" (cf. Coulthard 1985: 64), functioning analogously to more elaborate signals like *Yd like to make two points* or simpler markers such as *if, since or firstly*. When a speaker has uttered a discourse marker that, at least in principle, implies a further continuation and elaboration on the speaker's part, he or she manages to reserve the right to speak for an extended amount of time. Since after an announcement like *to begin with* it would be rather impolite to interrupt the current speaker right at the end of his or her first sentence, it functions in effect as a kind of floor-securing device.

Secondly, according to the COBUILD dictionary, *to begin with* is used

especially when you want to correct something that someone else has just said EG *To begin with, the invitation for eight really means eight-thirty to nine.*

This correcting function of *to begin with* is probably restricted to spontaneous spoken discourse and is therefore not recorded in our corpus, which is derived from written texts. Ultimately, it is a matter of conjecture whether the correcting function, which is undoubtedly sometimes intended by speakers, is derived from the meaning of the idiom or from its textual function, perhaps the function as an incompletion marker. Be that as it may, it seems to be the case that the correction function and the use as incompletion marker taken together, have led to uses of *to begin with* which we would like to interpret as "linguistic threatening gestures". By this term we refer to situations — probably familiar to the reader — where we are desperately seeking arguments for refuting someone else's statement. To gain time and to intimidate the other person we begin our utterance with *to begin with*, thus creating the impression that we have a whole list of arguments in store which we will advance to support our point of view. Since, however, we are in fact particularly short of arguments, what we do is in effect impose on the other person. Under such situational circumstances, *to begin with* is used to feign arguing power; it can be said to function as a linguistic threatening gesture.

4. *To begin with, to start with, for a start*

It is rather obvious that *to begin with* resembles two other formally complex discourse markers, viz. *to start with* and *for a start*. How are the three related to each other? First of all, it must be mentioned that due to the shortage of examples in the LOB and Sanford (1985), our observations concerning this question have to rely on the examples provided by dictionaries and on the judgment of native speakers. In the LOB, *to start with* does not occur at all and *for a start* was recorded only three times. In Sanford (1985), there are two examples of *to start with* and none of *for a start*. How can this significant difference in the frequency of occurrence of *to begin with* on the one hand and *to start with* and *for a start* on the other be explained? In this section, we will consider a) semantic, b) stylistic and c) functional aspects which seem to be responsible for the different frequency of *to begin with, to start with* and *for a start*.

a) What, in semantic terms, is the difference between the verbs *begin* and *start*? In an earlier study (Schmid 1993), the internal category structure of the two verbs was investigated. The data for this research consisted of 318 examples of *start* and 472 examples of *begin* taken from the LOB. All 790 occurrences of the two verbs were subjected to a computer-aided contextual anal-

ysis which took into account various semantic and grammatical features deducible from their linguistic environment. The results of this study are summarized in figure (1), where tentative labels for the subcategories of the verbs are rendered in quotation marks, while their meaning is schematically specified by means of semantic features or attributes.

Figure 1 shows that both *start* on the left hand side and *begin* on the right hand side display networks of subcategories which overlap in a family resemblance fashion (cf. Rosch & Mervis 1975). In terms of frequency of occurrence, both verbs merge in a common prototypical core represented in the diagram by the bold box. The meaning of this shared prototypical subcategory can be characterised by the attributes 'inchoative' and 'dynamic', with *start* tending towards 'agentive' contexts and *begin* towards 'non-agentive' contexts. In other words, both *start* and *begin* most frequently denote beginnings of actions (*start*: 67.6%, *begin*: 93.4% of the respective total of occurrences); however, even within this shared meaning (whose frequency of occurrence leads to the "naive" view that *start* and *begin* are synonyms), *begin* differs from *start* in that *begin* often takes the agent out of focus. For an illustration compare examples (13) and (14):

(13) The race began.

(14) The sprinters started to run.

Clearly, the two verbs are interchangeable in the two sentences, but the statistical evidence shows that English speakers prefer *begin* in non-agentive and ergative contexts like (13) and *start* in the canonical agentive construction like (14).

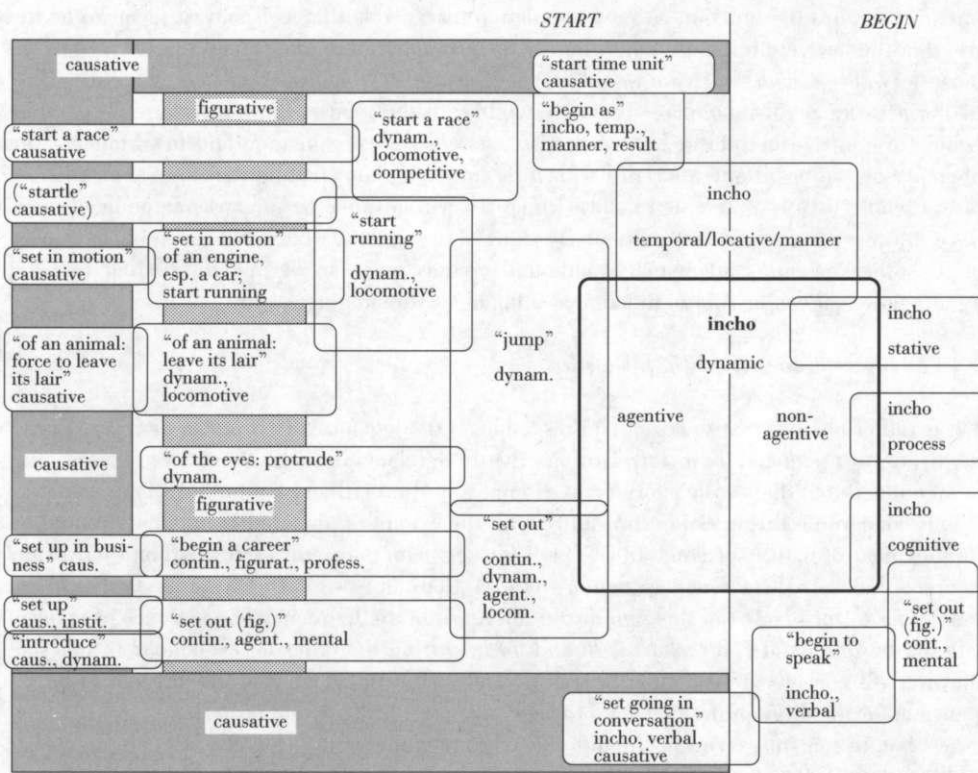


Figure 1: Schematic representation of the internal category structure of *start* and *begin* (Schmid 1993: 268)

Yet this difference in meaning is minute compared to the semantic contrast recognizable in the rest of the category network. Quantitatively, even a cursory glance at the diagram will show that whereas the category structure of *start* is highly complex, branching out in various polysemous chains, the use of *begin* is confined to a rather limited area. Semantically speaking, *begin* is used more often than *start* for gradual beginnings, mostly in the cognitive or emotive domains, and for contexts involving speaking and talking. Thus, the LOB data indicate that statistically people prefer (15), (16), (17) and (18) to (15'), (16'), (17') and (18').

(15) The pattern began to emerge ... (LOB G 49 105)

(16) Then I began to think. (LOB G 09 161)

(17) He began to feel limp ... (LOB K 27 95)

(18) 'Look here', he began,... (LOB L 11 124)

(15') The pattern started to emerge ...

(16') Then I started to think.

(17') He started to feel limp ...

(18') *'Look here', he started, ...

Start, on the other hand, is used to denote dynamic and often sudden beginnings of actions. In addition, and this is in fact related to its etymological origin OE *styrtan*, 'jump' etc., *start* leaves the inchoative area and covers a whole range of purely dynamic, figurative and causative meanings.

This major semantic difference between *begin* and *start*, which manifests itself in a higher frequency of occurrence for *begin* in mental context and when the beginning of speaking is denoted, fits in nicely with our findings concerning *to begin with* and *to start with*. We have seen that either of the fixed expressions can be used to refer to the structuring of discourse or for other linguistic and metalinguistic activities. It can therefore hardly come as a surprise that the fixed expression derived from the verb which specializes in this type of context is more frequent. In fact, the parallel in the contrast between *begin* and *start* on the one hand, and *to begin with* and *to start with* on the other, can be regarded as further support for our claim that the ties between the original and the lexicalized meaning of a fixed expression are not completely severed.

b) In addition to these semantic considerations, the style of the texts we have used could be a determining factor for the difference in frequency between the three fixed expressions. According to the judgment of native speakers, *begin* is considered to be slightly more formal than *start*, and is therefore even today preferred in written language, when the two are semantically interchangeable. The same probably applies to the fixed expressions derived from the two verbs,⁵ and so the shortage of *to start with* and *for a start* in our written text corpora may not be a coincidence.

c) Finally, it seems reasonable to differentiate between the various functions that all three discourse markers can have. Take the message-structuring function. Undoubtedly, the mental chunking and structuring of an utterance into various points that will be made requires at least a minimal amount of preplanning activity. We would therefore expect that the announcing function of *to begin with*, *to start with* and *for a start* is particularly dominant in highly-edited written texts and maybe in the speech of experienced speakers such as politicians and lecturers. From the point of view of text genres, however, this kind of textual output again belongs to rather formal text types where *begin* and consequently *to begin with* will tend to occur more often than *to start with* and *for a start*. Conversely, the correcting function and the imposing threatening gesture will more often be exploited in spontaneous spoken speech, where *start* is more frequently chosen. So everyday unplanned discourse is the area where we would also expect *for a start* and *to start with* to crop up most often.

Altogether, we have provided a number of explanations for the higher frequency of occurrence of *to begin with* over *to start with* and *for a start*, which, however, can hardly be proved by reference to our working corpus.

5. Summary and conclusion

What have we been able to find out about the fixed expression *to begin with*? We have tried to show that it would be misleading to think of *to begin with* as an expression which has one specific idiomatic meaning. Instead we would like to propose a cline of idiomaticity, ranging from literal uses of the sequence *to begin with* over uses which only introduce or announce a list, to fully idiomatized uses, which do not only exhibit specific semantic, formal, syntactic and phonological properties, but also perform one or more textual and pragmatic functions. Even among the clearly idiomatic uses, it seems possible to distinguish between more or less idiomatized examples, e. g. between those involving the correcting function or what we have called the linguistic threatening gesture on the one hand, and those which simply introduce an enumerating grid, structuring the text, on the other.

The idea of a gradient of idiomaticity of course raises the question whether the diachronic lexicalization process, which has led to the idiomatic use of *to begin with*, might or even must have gone through the same stages. The earliest record of *to begin with* in the OED dates from the year 1531:

And, to begin withal, they said *Confiteor*.

As indicated by the quotation and also pointed out in the OED, the earlier version of *to begin with* was *to begin withal*. This can be seen as a clue that historically, *to begin with* is not derived from „normal“ uses of *to begin with* but has probably functioned as an adverbial phrase right from the start. As far as the various semantic, grammatical, textual and pragmatic properties of the modern *to begin with* are concerned, the OED provides no further evidence of their evolution. Whether the progression towards semantic and pragmatic complexity that we have outlined has a diachronic parallel must therefore remain a matter of speculation.

Finally, abstracting from the examples that we have dealt with, we hope to have shown that an isolated semantic and/or syntactic analysis of fixed expressions clearly is not adequate any more in the nineties. As Fillmore et al. (1988) have demonstrated for the case of *let alone*, it is not only useful and instructive, but even necessary, to take textual and pragmatic considerations into account, when such apparently simple idioms are investigated.

- 1 It is not uncommon, for example, in a discussion following a lecture, that the speaker could answer a series of questions by saying: to take up your last point first.
- 2 For various linguists who have recognized that there are degrees of idiomaticity, as for example in Fraser's "frozenness hierarchy" postulated in 1970, cf. Lipka (1974).
- 3 Note that the comma after *next* indicates a pause, as in (4), and as opposed to (10).
- 4 It is worth mentioning that in fact seven of the ten examples of *to begin with* in Sanford (1985) follow the same pattern. This shows the danger of relying too much on a small number of authors, because obviously writers have predilections for certain wordings they habitually use. On the other hand, this is a strong argument for using text corpora which consist of language samples of different origins as e. g. the LOB corpus.
- 5 Allen (1990) and Quirk et al. (1985) mark *for a start* as *colloq.* and *informal* respectively.

References

- Allen, Richard E. (1990), *The Concise Oxford Dictionary of Current English*, 8th. ed., Oxford.
- Coulthard, Malcolm (1985; 1977), *An Introduction to Discourse Analysis*, 2nd. ed., London–New York.
- Fillmore, Charles J., P. Kay, & M.C. O'Connor (1988), "Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone", *Language* 64, 501–538.

- Grimm, Ursula (1991), *Lexikalisierung im heutigen Englisch am Beispiel der ER-Ableitungen*, Tübingen.
- Halliday, Michael A. K. & Ruqaiya Hasan (1976), *Cohesion in English*, London–New York.
- Keller, Eric (1981), “Gambits: Conversational Strategy Signals”, in F. Coulmas, ed., *Conversational Routine*, The Hague etc., 93–113.
- Leitzke, Eva (1989), *(De)nominale Adjektive im heutigen Englisch. Untersuchungen zur Morphologie, Syntax, Semantik und Pragmatik von Adjektiv-Nomen Kombinationen der Typen atomic energy und criminal lawyer*, Tübingen.
- Lipka, Leonhard (1974), „Probleme der Analyse englischer Idioms aus struktureller und generativer Sicht“, *Linguistik und Didaktik* 20, 274–285.
- Lipka, Leonhard (1991), Review of Schiffrin (1987), *Anglia* 109, 148–151.
- Quirk, Randolph et al. (1985), *A Comprehensive Grammar of the English Language*, London–New York.
- Rosch, Eleanor & Carolin B. Mervis (1975), “Family Resemblances: Studies in the Internal Structure of Categories”. *Cognitive Psychology* 8, 573–605.
- Sanford, Anthony J. (1985), *Cognition and Cognitive Psychology*, New York.
- Schiffrin, Deborah (1987), *Discourse Markers*, Cambridge.
- Schmid, Hans-Jörg (1993), *Cottage und Co., Idea, Start vs. begin. Die Kategorisierung als Grundprinzip einer differenzierten Bedeutungsbeschreibung*, Tübingen.
- Searle, John R. (1976), “A Classification of Illocutionary Acts”, *Language in Society* 5, 1–24.
- Sinclair, John et al. (1987), *Collins Cobuild English Language Dictionary*, London. (Cobuild)
- Summers, Della et al. (1978; 1987), *Longman Dictionary of Contemporary English*, 2nd ed., Harlow etc. (LDCE)
- Taylor, John R. (1989), *Linguistic Categorization. Prototypes in Linguistic Theory*, Oxford.
- Tsohatzidis, Savas L., ed. (1990), *Meanings and Prototypes. Studies on linguistic categorization*, Oxford.
- Wierzbicka, Anna (1990) “Prototypes Save’: on the Uses and Abuses of the Notion of ‘Prototype’ in Linguistics and Related Fields”. In: S.L. Tsohatzidis, ed., *Meanings and Prototypes. Studies on Linguistic Categorization*, Oxford, 347–367.