Jurate Ruzaite

Academic Precision Reconsidered: A Corpus-based Account

Abstract

Although precision in earlier centuries and decades was often treated as one of the major principles in the world of academia, the present paper argues that sometimes vague language is preferable in academic discourse. To show that vague language can serve positive purposes and that precision is not the ultimate goal in communication, the present investigation studies the use of approximators, which make a number less precise. Approximated numbers are investigated in several types of discourse: leisure, business, political and academic discourse. These four discourse types are analysed in both written and spoken modes. The data for the present investigation have been obtained from the British National Corpus. The results demonstrate that approximators are distributed more or less evenly in business and political discourse, in both written and spoken modes. Approximators are most frequently employed in written academic and spoken leisure discourse, though in spoken academic discourse their frequency is considerably lower. To explain this distribution of approximators, their multiple functions in academic discourse are discussed.

1. Introduction

The present paper briefly reviews the traditional perception of precision as a supposedly necessary requirement for human communication and especially in academic discourse. The principle of precision is questioned and re-evaluated on the basis of corpus data, which illustrate the distribution of imprecise quantities in various types of discourse. The category under investigation is approximators (about, around, approximately, roughly and round), which make a number less specific and, thus, should be avoided in academic discourse if the primary aim of this particular type of discourse is precision. To limit the scope of investigation, the analysis is based only on these five approximators. Since

---

1 I am very grateful to Dr. Milda Danyte for reading this paper and making corrections in it. I would also like to thank the two anonymous referees for very useful criticisms and valuable comments made on a previous draft of this paper.

*SKY Journal of Linguistics 17 (2004), 217–247*
they are a category of high frequency, investigation of a larger number of approximators would hardly be manageable. Besides, these approximators form a homogeneous set on the basis of two other aspects. First, they typically precede a number, whereas some of the other approximating units such as odd or more or less follow a number. Second, when these five approximators precede a number, they indicate that the quantity can be larger or smaller than the given number. Other approximators, meanwhile, indicate a more limited range. For instance, odd, more than and over indicate that the quantity is bigger than the given number; almost and nearly indicate that the quantity is smaller than the approximated number. Finally, the approximators about, around, approximately, roughly and round are typically described in dictionaries as the main approximators and as interchangeable synonyms (see, for instance, Oxford Advanced Learner’s Dictionary of Current English (1987), Longman Dictionary of Contemporary English (1992, 1995)).

The data have been obtained from different subcorpora of the British National Corpus (BNC) to demonstrate the varying frequencies of approximators in different types of discourse, namely, leisure, business, political and academic discourse. Such a comparative approach is expected to reveal some tendencies in the use of approximators in discourse in general. Academic discourse will be contrasted with other discourse types to see how much the frequency of approximators in academic discourse is different from their frequency in other discourse types. Thus leisure, business and political discourse serve as comparative information that helps to highlight the use of approximators in academic discourse.

On the basis of the collected linguistic evidence, it is argued that precision is not the ultimate aim of academic discourse. Approximated numbers are employed in the world of academia with a similar or even higher frequency than in other types of discourse. The high frequency of approximated numbers in academic discourse suggests that imprecision, which can serve a variety of communicative functions, is unavoidable. Indeed, the omission of approximators could even have negative effects on communication. Hence though a certain degree of precision is required in academic discourse, in some contexts imprecise references to numbers are preferred.
2. Vagueness: overview of previous research

Criticism of a lack of precision, or vagueness, goes back to the philosophy of Plato (1914) in *Euthypro* and Aristotle in *Rhetoric* and their attempts to define the ‘essence’ of each word precisely in order to comprehend its ‘true’ meaning. According to Aristotle, when speaking about events we should always aim at specificity and indicate exact dates, numbers, etc. (1946: 1407).

Wittgenstein’s *Philosophical Investigations* (1967) is one of the earliest non-Aristotelian approaches to word meaning. Wittgenstein’s non-rigid approach can be called the basis of the reversed treatment of imprecision, or inexactness, in language. With respect to exactness and inexactness, Wittgenstein criticizes the widely-spread attitude that “‘inexact’ is really a reproach, and ‘exact’ is praise” (1967: 42). In response to this stigmatisation of imprecision, he claims that “we misunderstand the role of the idea in our language” (1967: 45).

The belief in the exact meaning of words and a clear delineation between different lexemes is categorically rejected as a misleading attitude to word meaning by other philosophers and linguists as well (e.g. Popper 1945, Janicki 1999). Popper rightly points out that we should aim at clarity and simplicity, although absolute precision is unattainable (Popper 1945). Besides, he observes that the ideal of precise concepts is a myth even in physics (Popper 1957).

More flexible approaches to imprecision can also be observed in the recent trends in logic. Some logicians who deal with the issue of vagueness tend to admit the limitations of classical logic by acknowledging the existence of borderline cases and the lack of sharp boundaries in language (e.g. Keefe 2000). Perhaps the most significant departure from classical logic is the logic of fuzzy sets, which was developed by Zadeh (see Yager et al. 1987). As Zadeh (Yager et al. 1987) observes, most of our reasoning is imprecise and most classes of things in the real world do not have sharp boundaries. Therefore, science has to change attitude to vagueness by rejecting the ‘fetishism of precision’ (Yager et al. 1987).

Though for a long time vagueness was mainly an object of philosophical investigations, it has recently become an important object of analysis in linguistics as well. Uses and misuses of vagueness are investigated following the principle that precision is not the ultimate goal of communication, and that vagueness is an important communicative
strategy of multiple applications. Various categories of vague language have already received some attention from linguists in analyses of varying length (e.g. Crystal and Davy (1979), Dines (1980), Aijmer (1985), Wierzbicka (1986), Channell (1980, 1990, 1994), Dubois, B.-L. (1987), Dubois, S. (1992), Overstreet (1995, 1999), Myers (1996), Overstreet and Yule (1997, 2002), Overstreet and Yule (2002), Drave (2002), Cheng and Warren (2001, 2003), Jucker et al. (2003), etc.). These investigations demonstrate that vagueness is a natural language feature which performs multiple functions in communication. Moreover, the results of previous investigations demonstrate that vagueness cannot and should not be avoided, since overprecision can lead to communicative breakdowns, as will be argued later in the present paper.

Vague language and approximators in particular are frequently treated as one of numerous hedging strategies. Hedging helps the speaker avoid categorical and straightforward assertions by allowing the speaker to distance him/herself from a claim and so reduce his/her commitment to the claim. As Myers (1989) points out, hedges are employed when authors of scientific texts cautiously present new research results to other researchers. In such a situation, caution is necessary, since the new data or results might involve lack of certainty. Thus hedging helps the author follow the principle of honesty when presenting uncertain claims to other scientists (Myers 1989). The broad category of hedges as a means of expressing doubt and uncertainty in scientific research articles has been investigated by Hyland (1998, 2000). Approximators as hedges are touched upon by Stenström (1994), who equates the use of sort of and approximately in spoken interaction (1994: 129). Prince et al. (1982) investigate approximators as a category of hedging in physician-physician discourse and arrive at the conclusion that uncertainty is the main underlying cause of hedging as well as approximating. Similarly, Clemen (1997) observes that vague statements function as a hedging strategy when exact data or precise information are not available, and thus the speaker is uncertain about a particular claim. Approximators as a category of hedges are discussed in relation to various politeness strategies by Brown and Levinson (1994). Itani (1996) analyses approximators as a type of hedges in English and Japanese. Importantly, Itani (1996) argues that approximators do not always function as hedges.

It has to be admitted that approximators do have a mitigating effect and are used in cases of uncertainty just as many other types of hedges.
However, as will also be argued later in this paper, hedging and expressing uncertainty is just one of the multiple functions of approximators (cf. Channell (1990, 1994)). Besides, hedging, being a very broad issue, is beyond the scope of the present paper and thus will not be discussed here in greater detail.

Investigations of imprecision in academic discourse include analyses carried out by such linguists as Dubois (1987), Myers (1996) and Channell (1990). Myers (1996) studies the strategic uses of vagueness in various types of academic writing such as articles, textbooks, prospectuses, grant proposals, reviews and popularisations. The categories under Myers’ (1996) investigation include polysemic terms, general terms, non-numerical references, elliptical expressions and others, but approximated numbers are not touched upon. The investigations by Dubois (1987) and Channell (1990) account for the usage of imprecise quantity references, including approximated numbers. Dubois (1987) focuses on imprecise numerical expressions in biomedical slide talks. Her investigation is especially important for the present analysis, since it also focuses on spoken academic discourse. Dubois (1987) attributes approximators to the broad category of hedges, which express the speaker’s stance. Channell (1990) investigates precise and vague quantities in written texts on economics. On the basis of her results, she points out the main functions of vague quantities performed in academic discourse: (a) giving the right amount of information, (b) deliberately withholding information, (c) persuading the reader, (d) lacking specific information, and (e) downgrading and highlighting (Channell 1990). Channell (1990) relates these functions of vague quantities to the Gricean maxims of quality, quantity and manner, which are to be observed in order to achieve successful communication. Thus Channell’s investigation clearly demonstrates that vagueness can serve important communicative purposes, and the belief in academic precision and accuracy is ‘a mere prejudice’ (Channell 1990: 95).

Imprecision in academic discourse is closely related to such notions as modesty, honesty and caution. In relation to medical English written discourse, Salager-Meyer rightly points out that “argumental arrogance and exuberance are not well regarded by the scientific community; whereas in contrast, humility, coyness, and cautiousness are” (1994: 150). Swales (1990) notices that assertiveness and precision are avoided in academic discourse when the principles of honesty and modesty are followed. According to Swales, hedging serves then as a useful means of “projecting
honesty, modesty and proper caution in self-reports and for diplomatically creating research space in areas heavily populated by other researchers” (Swales 1990: 175).

Previous research has shown that vagueness can be treated as a natural feature of language that is sometimes crucially important for successful communication. Vagueness as a natural feature is observed by Channell (1994) and Wardhaugh (1985, 1993). People usually comprehend vague language effortlessly and do not even notice it in everyday interactions. Wardhaugh, for instance, claims that “we seek only occasional clarification of remarks made to us” (Wardhaugh 1985: 8). Moreover, “one of the basic principles of communication is that you cannot, and should not, always ‘tell it like it is’” (Wardhaugh 1985: 36). Generally, vagueness has to be tolerated in a conversation, since “vagueness rather than precision will prevail” (Wardhaugh 1993: 181).

According to some scholars, the increase of precision in communication might have negative effects (Tannen 1989, Popper 1992). Inappropriate use of details, according to Tannen (1989), can be completely ineffective, as they can be boring (mainly observable in interactions between the old and the young), insulting (especially when expressing criticism in a highly specific way) or a basis for humour (Tannen 1989). Popper also notices that precision, paradoxically enough, “usually leads to loss of clarity, and to waste of time and effort on preliminaries which often turn out to be useless, because they are bypassed by the real advance of the subject: one should never try to be more precise than the problem situation demands” (1992: 24, emphasis added).

Vagueness, although it usually causes no hindrances to communication, requires mutual tolerance from interlocutors. As Wardhaugh (1985) observes, speakers often do not present very detailed information, and instead use the anaphoric this and that, or such vague expressions as something like, sort of, kind of, etc. Therefore, listeners inevitably become active participants in constructing the meaning of the speaker’s message. Such cooperation in meaning construction is treated as one of the advantageous aspects of communication by Tannen (1989). One of the involvement strategies in communication that Tannen points out is indirectness. When interlocutors work harder to ‘decode’ the meaning, they comprehend it better and feel the communication process and the result more rewarding. In this way hearers or readers become active participants and are, as she says, ‘meaningfully, mythically involved’ (Tannen 1989:
17). Vagueness as a trigger of active cooperation between the participants of communication is also observed by Danell (1978), who calls vague and ambiguous utterances “cues proposed to the listener.” By this label he suggests that when listeners hear a non-specific utterance, their conceptual store is activated. Listeners’ brains instantaneously rely upon their contextual and background knowledge and thus listeners become actively involved in the interpretation of the given cue.

3. Data and method

The present investigation is a corpus-based analysis of linguistic variation across different types of discourse. The data have been obtained from several subcorpora that constitute the BNC, each illustrating a different type of discourse. As can be seen in Table 1, four major categories of texts within the subcorpora of spoken and written texts have been selected for the present analysis.

<table>
<thead>
<tr>
<th>Text type</th>
<th>Size in word units</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spoken</strong></td>
<td></td>
</tr>
<tr>
<td>Educational</td>
<td>1 265 318</td>
</tr>
<tr>
<td>Business</td>
<td>1 321 844</td>
</tr>
<tr>
<td>Public/Institutional</td>
<td>1 345 694</td>
</tr>
<tr>
<td>Leisure</td>
<td>1 459 419</td>
</tr>
<tr>
<td><strong>Written</strong></td>
<td></td>
</tr>
<tr>
<td>Natural and pure sciences</td>
<td>3 752 659</td>
</tr>
<tr>
<td>Commerce and finance</td>
<td>7 118 321</td>
</tr>
<tr>
<td>World affairs</td>
<td>16 507 399</td>
</tr>
<tr>
<td>Leisure</td>
<td>9 990 080</td>
</tr>
</tbody>
</table>

Table 1. Text types of the BNC subcorpora selected for the investigation

All the subcorpora consist of more than 1 million words. Thus, quantitatively, the corpus provides a sufficient amount of authentic and representative material to collect reliable evidence to study the usage of approximators in different types of discourse. However, as can be seen in Table 1, written and spoken types of texts are not represented evenly. Although written texts evidently outweigh the spoken ones, the subcorpora are comparable, since in the present investigation the frequency of approximators is expressed per 1 million words.
The structure of each category of spoken texts encompasses specific types of events. Business events include sales demonstrations, trades union meetings, consultations and interviews. Institutional and public events consist of sermons, political speeches, council meetings and parliamentary proceedings. Leisure events cover sports commentaries, after-dinner speeches, club meetings and radio phone-ins. Finally, educational events cover such in-class events as lectures and lessons, tutorials, seminars, discussions, demonstration lessons, writing groups, workshops, training sessions, and classroom interactions. The topics of the lectures and lessons include a variety of subjects such as computers, mathematics, oceanography, microbiology, chemistry, social studies, politics, economics, geography, bioenergetics, psychology, English, communication skills, music, drama, history, religion, art, food and etching. Thus the topics are highly varied and include both the humanities and natural sciences.

Written texts consist mainly of books, which constitute 60% of the written subcorpus. 25% of the written subcorpus is periodicals (e.g. newspapers, journals). The rest of the texts are miscellaneous published material (e.g. brochures, advertising leaflets), which make up 5–10%; unpublished material (e.g. personal letters, diaries, essays, memoranda), which also make up 5–10%; and a small amount (less than 5%) of material written to be spoken (e.g. political speeches, play texts, broadcast scripts). The last category of written texts to be spoken is the most marginal one. Since it is likely to contain numerous spoken language features alongside written language features, it is not really representative of written texts. However, the amount used of these texts is very small and thus cannot influence the results to a great extent. Besides, these texts belong to different genres and form an even smaller part of the different subcorpora, for example, political speeches belong to political discourse, while play texts and broadcast scripts belong to leisure discourse. When distributed across various discourse types, the occurrence of these texts becomes too small to influence the general results.

Written academic texts include books and periodicals. The subjects that are discussed in written academic texts cover various topics from the field of natural and pure sciences, e.g. microbiology, chemistry, physics, medicine, geology, sedimentology, mathematics, engineering, astronomy, oceanography, ecology and geography. Thus the scope of written academic texts is narrower than the scope of spoken academic discourse. However, some of the subjects occur in both types of academic discourse, including
mathematics, oceanography, microbiology, chemistry, economics and geography.

The labels for the text types within the written and spoken subcorpora in Table 1 might suggest that the text types are different in the two subcorpora. However, though the texts have different labels in spoken and written subcorpora, they can be attributed to the same general categories of discourse types. First, the spoken category of educational texts and the written text category of natural and pure sciences will be referred further to as academic discourse. Another text type under investigation is business discourse, which is called commerce and finance texts in the written subcorpus. Third, the spoken public and institutional texts and written texts of world affairs will be referred to as political discourse. The final category, whose label is the same in both written and spoken subcorpora, is leisure discourse. Hence the text types under investigation can be attributed roughly to broader categories of discourse types so that they appear to be comparable.

4. Research questions and hypotheses

The present analysis of approximators in different types of discourse attempts to answer the following questions:

1. If the frequency of approximators is discourse-specific, are approximators less frequent in academic discourse than in other types of discourse?
2. Do the speakers of English demonstrate any preferences for certain approximators in different discourse types?
3. What purposes do approximators serve in academic discourse?

In relation to these research questions, the present investigation offers the following hypotheses:

Hypothesis 1. according to the existing belief in academic precision, approximators should be less frequently employed in academic discourse than in other types of discourse.

Hypothesis 2. Speakers of English prefer different approximators in different discourse types, and these differences are most evident when spoken and written academic discourse is contrasted.
Hypothesis 3. In academic discourse approximators perform specific functions, which differ in spoken and written discourse.

The present investigation will test these hypotheses on the basis of corpus data.

5. Results: approximators in various discourse types

The distribution of approximators in the selected subcorpora reveals several important tendencies in the variation of approximators. First, extensive variation is observed among different discourse types. As will be argued further, the frequency of approximators differs considerably depending on the situational context of a discourse, i.e. leisure, business, political or academic situations. Second, variation is observable in how often different approximators are used. As will be demonstrated, some approximators are more extensively employed than others. In addition, different approximators are preferred in different types of discourse. Finally, approximators serve a variety of functions, which differ in written and spoken academic discourse.

5.1 Frequency of approximators in spoken discourse

The data have demonstrated that the frequency of approximators in spoken discourse amounts up to as many as 1,943.45 occurrences in all the subcorpora. As can be seen in Table 2, the frequency of approximators in leisure discourse is 761.56 occurrences per 1 million words, which makes up more than 39% of the total number of approximators. In business discourse approximators occur in 490.71 instances per 1 million words, which makes more than 25% of the total number. Approximators are even less frequent in political discourse; there are 357.77 occurrences per 1 million words in political discourse, which is 18.41% of the total number. The frequency of approximators in academic discourse is only slightly smaller than in political discourse. In spoken academic discourse approximators occur 333.41 times per 1 million words, which constitutes 17.16% of the total number. The varying distribution of approximators in different discourse types is especially evident in Figure 1. As could be expected, approximators are clearly more frequent in leisure discourse than in any other spoken discourse type. Differences in the frequency of
approximators among the other discourse types, however, are much less notable. The frequency of approximators in political and academic discourse differs by just one percent. In business situations approximators are more frequent than in political or academic discourse by just seven to eight percentage points.

<table>
<thead>
<tr>
<th></th>
<th>Leisure</th>
<th>Business</th>
<th>Political</th>
<th>Academic</th>
</tr>
</thead>
<tbody>
<tr>
<td>about</td>
<td>705.46</td>
<td>399.90</td>
<td>277.92</td>
<td>288.70</td>
</tr>
<tr>
<td>around</td>
<td>12.43</td>
<td>22.13</td>
<td>15.82</td>
<td>11.66</td>
</tr>
<tr>
<td>approximately</td>
<td>14.62</td>
<td>29.76</td>
<td>34.65</td>
<td>9.72</td>
</tr>
<tr>
<td>roughly</td>
<td>24.86</td>
<td>37.40</td>
<td>27.87</td>
<td>21.39</td>
</tr>
<tr>
<td>round</td>
<td>4.39</td>
<td>1.52</td>
<td>1.51</td>
<td>1.94</td>
</tr>
<tr>
<td>Total:</td>
<td>761.56</td>
<td>490.71</td>
<td>357.77</td>
<td>333.41</td>
</tr>
<tr>
<td></td>
<td>39.19%</td>
<td>25.25%</td>
<td>18.41%</td>
<td>17.16%</td>
</tr>
</tbody>
</table>

Total number of approximators: 1943.45

Table 2. Frequency of approximators in spoken discourse (per 1 million words)

Another interesting observation that can be made on the basis of the corpus data is that speakers of English evidently prefer for certain approximators. In fact, the approximator *about* is the one that clearly predominates in all types of discourse. The outstanding dominance of *about* is especially noticeable in Figure 1. The frequency difference between *about* and the other approximators is extremely drastic. For instance, in academic discourse *about* is almost 25 times as frequent as the approximator *around*. In leisure discourse this difference is even more dramatic, i.e. *about* is almost 60 times as frequent as *around*. In total, *about* is employed in 1671.98 instances per 1 million words, which constitutes 86% of the total number of approximators in all the four discourse types. It is interesting to notice that the frequency of *about* in academic discourse is slightly higher than in political discourse, which again supports the observation that approximators are used to a very similar extent in these two types of discourse.
So far the analysis has revealed two main tendencies about the use of approximators in different types of spoken discourse. First, approximators are most frequent in leisure discourse and least frequent in academic discourse. However, the difference in the frequency of approximators in business and especially political and academic discourse is not very big. Second, the data have demonstrated a clear preference for the approximator *about* in all the spoken discourse types.

### 5.2 Frequency of approximators in written discourse

In written discourse approximators occur 1,554.33 times per 1 million words. As can be seen in Table 3, they are most frequent in academic discourse, where they occur in 575.93 instances per 1 million words, or slightly more than 37% of the total number of approximators in written discourse. Approximators in leisure discourse amount to 372.34 occurrences per 1 million words, which is almost 24% of the total number. In political discourse approximators are employed in 347.37 instances per 1 million words, or more than 22% of the total number. Thus in leisure and political discourse approximators are almost evenly distributed. Approximators are least frequent in written business discourse, where they make up 258.69 occurrences per 1 million words, or almost 17% of the total number.

Moreover, the difference in the frequency of approximators in written business, political and leisure discourse is no greater than that in spoken business, political and academic discourse (6–7%).
As for a preference for specific approximators, *about* is considerably more popular than the other approximators, as can be seen in Figure 2. However, in written discourse this difference is less dramatic than in spoken discourse (cf. Figure 1). In written discourse *about* occurs in 999.31 instances, which is 64% of the total number. In spoken discourse, as has been mentioned, it occurs in 86% of the total number of approximators (see Section 5.1). In Figure 2 it is also noticeable that *approximately* is used more frequently than *around* and *roughly* in academic discourse, but this difference is not very great.
Second, the approximator *about* is clearly preferred more than other approximators, but the difference between the frequency of various approximators in written discourse is not as dramatic here as in spoken discourse.

### 5.3 Comparison of the frequency of approximators in spoken and written discourse

Though the data obtained from spoken discourse might have suggested that approximators are least frequent in academic discourse, the data from written discourse have demonstrated the opposite results. As Figure 3 reveals, in written discourse approximators are more or less similar in frequency as in leisure, business and political discourse. In written academic discourse, meanwhile, approximators are considerably more often employed than in any other type of written discourse. Approximators in written academic discourse amount to 37% of the total number of approximators in spoken texts, which is almost identical to the frequency of approximators in spoken leisure discourse (39% of the total number). This striking similarity between the formal style of academic writing and informal communication in leisure events demonstrates that imprecision can be an equally dominant feature in very different discourse types. The specific purposes of approximators in written academic discourse will be discussed in Section 5.4.

Another observation that follows from the results presented in Figure 3 is that approximators are distributed relatively evenly in business and political discourse, both in written and spoken form. The frequency of approximators in these two types of discourse differs only by several percentage points.
Spoken and written academic discourse differs not only in the frequency of approximators, but also in the speakers’/writers’ preference for different approximators. As Figure 4 demonstrates, *about* is evidently preferred in both spoken and written academic discourse. However, in written discourse other approximators seem to have more significance than in the spoken one. In written discourse *about* constitutes 66% of the total number of approximators, whereas in spoken discourse the frequency of *about* reaches even 87%. Another clear tendency that can be observed in Figure 4 is that an approximator that is extremely rarely used in both written and spoken academic discourse is *round*. 
Thus the comparison of written and spoken discourse types has revealed several tendencies of the usage of approximators. First, there is a great difference in the frequency of approximators in spoken and written academic discourse. At first glance it seems strange that approximators predominate in written academic discourse, which is considered to be more formal than spoken discourse. This is especially true given that written discourse provides an author with the possibility to reconsider or check a figure and to make it more precise. However, approximated numbers do prevail, which suggests that such numbers serve specific purposes in written academic discourse. This tendency suggests that imprecision is highly recurrent in two very different discourse types: spoken leisure discourse and written academic discourse.

6. Usage of approximators in spoken and written academic discourse

Since the frequency of approximators is so evident a feature of written academic discourse, they can be expected to perform specific purposes in this type of discourse. The difference between the number of approximators in spoken and written discourse suggests that there must be a difference in the use of this category in the two modes. The analysis has revealed that approximators perform multiple functions in both spoken and written
academic discourse. Some of those functions are identical in spoken and written discourse, while others are typical of only one type of discourse. Besides, approximators occur in a different linguistic context in spoken and written academic discourse, as a great variety of hedging devices are employed alongside approximators in spoken academic discourse. This tendency, as will be argued later, is one of the factors that predetermines the higher frequency of approximators in written academic discourse.

6.1 Functions of approximators in spoken and written academic discourse

As has been mentioned, approximators perform a variety of functions, some of which are the same in written and spoken academic discourse. First of all, approximators are employed in both spoken and written academic discourse when more precision is unattainable. A higher degree of precision is impossible in generalisations, which refer to a variety of cases, so that using a range in numbers becomes unavoidable, as in (1)–(6).

(1) The average smoker at any time has \textit{approximately} ten percent of their blood out of action…(S)\textsuperscript{2}

(2) Typically sea water is alkaline \textit{around} about eight plus or minus nought point two PH units. (S)

(3) Each year, more than 15,000 women in Britain die from breast cancer and \textit{about} 2,000 from cancer of the cervix. (W)

(4) It is \textit{usual} to hire a new vehicle on a fully maintained basis for a period of \textit{about} three years. (W)

(5) Most birds take \textit{about} 20 minutes to lay an egg. (W)

(6) An elephant needs \textit{around} 600 lb (270 kg) of vegetation a \textit{day}. (W)

The underlined words and phrases in examples (1)–(6) indicate that these claims are generalisations, e.g. \textit{at any time}, \textit{typically}, \textit{each year}, \textit{usual}, a

\textsuperscript{2} The letters in brackets after each example indicate the discourse type. (S) stands for spoken academic discourse; (W) stands for written academic discourse.
day, which refer to the regularity of a phenomenon. Such premodifiers of the subjects as most in (5) and average in (1) show that these generalisations encompass not individual, but typical cases.

Paradoxically, approximators, as a category of deliberate ‘fuzzifiers’, can also specify an abstract notion. Most commonly, such abstract notions are adjectives, as in (7)–(8), where the specified adjectives are underlined.

(7) A few years ago a minute fossil only about 2 mm long was recovered from limestones of Ordovician age, and christened Janospira. (W)

(8) For nearly two weeks they used very low currents, about 0.05 amps, and discovered that the cell cooled down more than expected. (W)

(9) Although he died comparatively young, around five years old, he had sired some excellent dogs who were used extensively at stud. (W)

(10) It’s very short, it’s only about fifteen hundred words. (S)

However, not only adjectives are followed by approximated numbers, but also adverbs are specified by them, as can be seen in examples (11)–(13).

(11) After lunch (at about two o’clock) many of us feel tired and may take a short nap, even though body temperature does not nominally fall much at this time. (W)

(12) Asteroids often have five arms, but many species have more -- up to about forty. (W)

(13) ... stags start their reproductive life somewhat later, around about age three, and their reproductive life ends earlier, around about age fourteen. (S)

As can be seen in examples (7)–(13), approximated numbers follow such abstract notions as very low/short, minute, comparatively young, earlier, somewhat later, after lunch, more. The approximated numbers are provided as additional information, which specifies the preceding abstractions. In written academic discourse approximations appear then either in brackets (as in (11)) or are separated by commas (as in (12)), which suggests that they are additional, but not the main information.

Approximators are also employed when an exceptionally large or small number is referred to, e.g. millions or a millionth of something. In
such instances a higher degree of precision is impossible. In examples (14)–(17) references to such large numbers as *million* and *billion* are made.

(14) *...it seems that dolphins and porpoises achieved their modern enlarged brains about 15–20 million years ago, whereas the evolution of the human brain is a phenomenon of the past few million years at the most. (W)*

(15) *British pigs produce about seven million gallons of liquid muck a day, posing a big disposal problem for farmers and a water quality headache. (W)*

(16) *The centre of the Sun where fusion occurs is at a temperature of around fifteen million degrees but practical fusion needs a much higher temperature than this -- hundreds of millions of degrees. (W)*

(17) *Our planet formed about 4.5–5 billion years ago. (W)*

It has to be noted that large numbers are especially frequently approximated with *about* in written academic discourse. For instance, in written academic discourse *million* is preceded by *about* in 46 instances (or 3.2% of the total number of occurrences). In spoken discourse the frequency of the pattern *about* + number + *million* is considerably lower, i.e. 4 instances (or 1.3% of the total number of its occurrences). *Around*, meanwhile, is employed with *million* in written academic discourse in only 6 instances (or 2.9% of the total number of its occurrences). Similarly, *approximately* co-occurs with *million* in 8 instances (or 2.7% of the total number of its occurrences); *roughly* co-occurs with *million* only once, which is only 0.5% of the total number of this approximator.

Small numbers are approximated less frequently than big numbers. The use of approximators with numbers smaller that one is demonstrated in (18)–(20).

(18) *If you examined Dover’s chalk cliffs with an electron microscope you would see their myriad forms, about one millionth of a metre across, the remains of trillions of these microscopic plants. (W)*

(19) *The lesion is made by dissecting away, using a tungsten needle, a square patch of embryonic skin measuring roughly 0.5x0.5mm and about 0.1mm deep (Fig. 1 a). (W)*
(20) *For the Christian community [roughly one per cent of Pakistan’s population] the day holds significance, and TV broadcasts special programmes for their benefit.* (W)

The approximated numbers in examples (18)–(20) are sufficient to indicate that the quantities under discussion are exceptionally small, and thus a higher degree of precision is not necessary. In such cases the smallness of the quantity is highlighted, and the exact number made less important (cf. Channell 1990).

Broadly speaking, approximators are used when uncertainty arises. Uncertainty is especially clearly expressed in spoken academic discourse, as in example (22). Here the multiple means of expressing uncertainty are underlined. In written academic discourse uncertainty is also sometimes indicated outright as the reason for a lack of precision (see (21)).

(21) *The lowest average Reynolds number at which this provides an instability mechanism is again uncertain but is around 5 × 10³.* (W)

(22) *Erm I’m not sure what time yet but erm it’s more likely to be sort of just before lunch or just sort of or maybe just after lunch or round four o’clock...* (S)

In such instances as (21)–(22) the highest possible degree of precision is provided; the approximators suggest the impossibility of more precision and a lack of certainty.

In spoken and especially in written academic discourse approximators are frequently employed when the speaker or writer refers to a proportion. In such cases approximators precede percentages, as in (23)–(25), or fractions, as in (26)–(29).

**Approximator + fraction**

(23) *... cover the pad and as you spiral up the arm all you need to do is to cover about two thirds of what you’ve just done before...* (S)

(24) *Turnover rose 3.1 per cent to DM 32 520m, with acquisitions accounting for about two thirds of this increase.* (W)

(25) *For three of the shapes the appropriate name was provided by around three-quarters of the bottom third pupils and by nearly all the top third.* (W)
Approximator + percentage

(26) So we know that so far about fifty percent our anthropogenetic CO two has been locked away in this system of the ocean. (S)

(27) By 1988 this had dropped to about 15 per cent of spruce, and other species such as pine and fir showed similar improvements. (W)

(28) Although all males try to get a second female, only about 15 percent are successful. (W)

(29) Northern offshore spotted dolphins were around 40–50 per cent of their original numbers, and common dolphins about 60 per cent. (W)

The data have revealed that the approximator about precedes percentages in 55 instances in written discourse (or 3.8% of the total number of its occurrences) and in 17 instances in spoken discourse (or 5.7% of the total number of its occurrences). Around in written academic discourse approximates percentages even more frequently, with 23 instances (or 11.3% of the total number of occurrences). Approximately precedes percentages in as many as 32 instances (or 10.7% of the total number of its occurrences). In spoken academic discourse, both around and approximately precede percentages just once. Roughly very rarely approximates percentages in spoken and written discourse, with only 4 instances in written discourse and 1 instance in the spoken one.

Hypothetical situations inevitably involve a high degree of uncertainty and speculations, so that approximators are frequently employed in utterances describing such situations, as in (30)–(31).

(30) If the annual quantity of net deployed within all North Pacific fisheries in a 6-month season is taken to be 3–5 million k (about 2–3 million miles), the total possible deaths of marine mammals would be about 100,000–181,500. (W)

(31) After about 100,000 years or so, the needle of the compass would abruptly become unstable and then swing round to face in the opposite direction, so that what was once compass north would become compass south, and vice versa. (W)

In examples (30) and (31) imaginary situations are modeled. Therefore, exact numbers would be less appropriate than approximations, since precision is impossible in such speculations.
Similarly, precision is impossible when future situations are predicted. In such cases approximators are frequently used to precede a number, as in (32)–(36).

(32) Bidwell believes that a group of about 50 consultancies will make up the final membership. (W)

(33) The jacket will weight around 820 tonnes and the topsides will have an operating weight of approximately 1,200 tonnes. (W)

(34) The suggested list will fill an area roughly 15ft x 5ft. (W)

(35) At some later time they could, by chance, all be in the right half or back in the left half, but it is overwhelmingly more probable that there will be roughly equal numbers in the two halves. (W)

(36) But that’s gonna take about five or six years to come to fruition… (S)

Such future references always involve uncertainty, which reduces the possibility of precision.

It is also common to employ approximators when references to the past are made, such as when a date is provided. Though various approximators are used in such contexts, most typically the approximator about is used to approximate a year. Such uses of approximators are demonstrated in examples (37)–(39).

(37) About 1970 a number of young adult women were found to have a rare cancer of the vagina: the appearance of several cases close together prompted enquiry about possible causes. (W)

(38) It is one of three seedlings of a hybrid raised in Switzerland around 1920 and consequently named after Swiss mountains. (W)

(39) Now in fact, before patch clamp came along in in around nineteen eighty there were there was another technique… (S)

If the date refers to the era B.C., the possibility of precision becomes even more reduced, as in (40)–(42).
The eruption which probably had the most far reaching effect on civilization was that of Santorin (Thera) in the Aegean, in about 1470 B.C. (W)

It begins with the atomic theory developed by Democritus around 460 BC. (W)

Then, around 1700 BC, the Minoans developed from the hieroglyphic script a new syllabic form, today called Linear A. (W)

When the author refers to events that took place almost four thousand years ago, as in (40) and (42), a precise date is neither important nor possible.

References to the past are made not only by providing an approximated year, but also by indicating how long ago something happened.

I experienced the problem about five years ago, when I had a bout of wheezing in the late summer. (W)

In (43) the approximated time period is sufficient to position the event in the past.

Approximators are employed not only when precision is unattainable, but also when precision is unnecessary. Without knowing the real intentions of the speaker/writer, it is very difficult to identify this function. However, some authors indicate outright that absolute precision is not necessary in certain contexts, as in (44)–(45).

The exact strength of solution does not matter -- about 1 dessertspoonful of cooking salt in 1 litre of tap-water is suitable. (W)

Its eigenvalues are 9.805, 2.411 and, very approximately (great accuracy is not necessary), &formula which yields, again approximately, &formula Proceeding in this way, we find at the end of the fifth step &formula and in the sixth step this gives &formula. (W)

The examples above suggest that academics are so conscious of the principle of academic precision that they sometimes feel obliged to justify themselves for a lack of precision.

The corpus data demonstrate that in written academic discourse approximately is frequently used when shapes are described. Another approximator used to refer vaguely to a shape is roughly, though it is less
often used in such contexts than *approximately*. The use of approximated references to various shapes is demonstrated in (46)–(50).

(46) *The ventral arm plates are **approximately** pentagonal with an obtuse proximal indented in the middle and are separated from one another.* (W)

(47) *The oral shield is large and is **approximately** triangular in shape.* (W)

(48) *The oral shield is **approximately** rhombic with the centre of the plate slightly depressed.* (W)

(49) *Disturbances first appear as **approximately** sinusoidal fluctuations, indicating a selective amplification process like that in a boundary layer.* (W)

(50) *The sand and silt particles, which are **roughly** spherical in shape range from 2,000 microns (in) to 50 microns in diameter (sands) and from 50 to 2 microns(silts)…* (W)

In such instances the principle of academic accuracy is upheld by employing approximators to point out that the indication of the shape cannot be exact.

Some functions of approximators in spoken academic discourse have not been encountered in written academic discourse. One such function is making impromptu calculations, as in (51).

(51) *So therefore, if this term I’ve got one, two, three, four, five weeks **roughly** left and I know that sometimes <pause> ermm <pause> what’s that?* (S)

Calculations at the time of speaking can occur only in spoken discourse, since the written one is usually carefully planned and based on accurate calculations.

Another function which is typical of spoken academic discourse, but has not been encountered in the written one, is the use of approximators for encouragement. Most commonly, the approximator *roughly* is used by the teacher to encourage the student to make a contribution. The use of approximators for encouragement is demonstrated in examples (52)–(53).

(52) *And what’s E to the point three? **Roughly**? What’s what’s E as a number? Just **approximately**?* (S)
What’s $Y$ equal to, roughly? *Approximately.* (S)

Such a use of approximators makes classroom communication more interactive and fosters student participation.

Finally, approximators are used for self-correction in spoken academic discourse, as in (54).

...there were two million, *about* two million names... (S)

Self-correction is also possible only in spoken discourse, which is much more spontaneously constructed than written discourse.

Thus this investigation has demonstrated that approximators are not only a frequent category in academic discourse, but they also serve multiple purposes in it. The data have revealed that approximators in both spoken and written discourse are mainly used when precision is either unattainable or unnecessary. Precision is unattainable when generalisations are made, proportions and percentages are referred to, and exceptionally large or small numbers are provided. Precise quantities are also impossible when the speaker/writer refers to hypothetical, future or past situations. In written academic discourse approximators tend to precede references to the shape of an object, whereas in spoken discourse such a function was not encountered. Its absence can be explained by the considerably lower frequency of approximators in spoken academic discourse. Thus it can be accidental that this use of approximators does not occur in such a small number of their occurrences. The three functions that occur only in spoken academic discourse are (a) making impromptu calculations, (b) encouraging listeners to participate, and (c) self-correction. These uses of approximators can be characteristic only of spoken discourse, which is a considerably more spontaneous and interactive type of discourse than written academic discourse.

To generalize, approximators do not differ drastically in the two modes of academic discourse, though some functions are performed in only one of the discourse types. Considerably more dramatic differences can be observed when the use of multiple self-distancing alongside approximators is analysed in the two modes. These differences are discussed below.
6.2 Multiple self-distancing devices in spoken and written academic discourse

An evident difference between spoken and written academic discourse is the tendency that, in spoken academic discourse, approximators are used with a variety of self-distancing devices in the same utterance. In written academic discourse, meanwhile, multiple self-distancing is considerably less frequent. When multiple self-distancing does occur, approximated numbers co-occur with modals indicating the degree of probability, as in (55)–(56).

(55) At the end of November, when the shoots should be about 1 -- 1 - 1/4in high, the bulbs can be gradually acclimatised to living room conditions. (W)

(56) It may vary from just around one tooth, to a whole section of your mouth and is not usually painful. (W)

Approximated numbers are also followed by general extenders, as can be seen in (57), where the general extender is underlined.

(57) After about 100,000 years or so, the needle of the compass would abruptly become unstable and then swing round to face in the opposite direction, so that what was once compass north would become compass south, and vice versa. (W)

As demonstrated in (58)–(59), another category that tends to co-occur with approximators in written academic discourse is such hedges as is thought or is believed.

(58) The eruption of the volcano buried a town, Akrotiri, which is thought to have had a population originally of about 30,000. (W)

(59) This is what seems to have happened early in the history of the Earth, because the oldest known fossil remains of proteinoid globules are believed to date back about 4 billion years. (W)

A greater multiplicity of self-distancing devices can be seen in (60), which is closer to the spoken mode because of the variety of such devices used.
(60) For a coarse to medium grained mature sandstone, with only a small fraction of clay, porosity ranges from 15 pu or more at around 2200 m to 1 pu or less below 6500 m. (W)

Thus in written academic discourse multiple self-distancing devices do co-occur with approximated numbers to indicate the degree of certainty. However, such co-occurrences are not frequent.

To see the contrast between spoken and written academic discourse more clearly, we should compare the examples provided above (in (55)–(61)) with the following examples in (62)–(64) which have been obtained from spoken academic discourse.

(61) ... it was appropriate for kids to be introduced to sex when they were around y’know sort of seven or eight or something in some cases. (S)

(62) ... right okay em are you free a bit later on this afternoon, probably around about threeish possibly? (S)

(63) Now in fact, before patch clamp came along in a in around nineteen eighty there were there was another technique which was available for looking at ...

It can be seen that utterances with approximators contain a variety of self-distancing devices such as discourse markers (y’know), hedges (sort of, probably), general extenders (or something), repetition (in in a in), and the suffix –ish in threeish. In fact, approximators rarely occur without multiple self-distancing devices in spoken academic discourse.

Hence the data suggest that in spoken academic discourse considerably more self-distancing devices are available than in the written one. Therefore, the self-distancing effect, which is typical of approximators, can be achieved through various other linguistic categories. It can be assumed that approximators in spoken discourse are not as necessary as in the written one. The formality of written academic discourse restricts the author’s choice of self-distancing devices. Besides, the possible number of such devices in an utterance is considerably lower than in spoken discourse. Thus approximators appear to be sufficiently formal to be used in highly technical contexts. Due to their formality, they are a highly recurrent self-distancing category in written academic discourse.
7. Conclusions

The present corpus-based account of approximators in various discourse types has yielded several observations about imprecision in academic discourse. It can be stated that the hypotheses raised at the beginning of the investigation have been only partly corroborated. One of the initial hypotheses was that approximators, whose frequency was expected to be discourse-specific, are less frequently employed in academic discourse than in other types of discourse. However, the data have revealed that approximators are an especially important feature of written academic discourse. Though it has to be admitted that approximators appear much less often in spoken academic discourse, they are almost as frequent here as in business and political discourse. Moreover, the only type of discourse where approximators are more frequent than in written academic discourse is spoken leisure discourse. Besides, this difference in frequency is very slight.

It was also expected that there should be a preference for certain approximators in written and spoken academic discourse (see Hypothesis 2). It has been revealed that speakers of English prefer different approximators in different discourse types. Generally, the approximator about predominates in all discourse types, whereas round is hardly ever used in any type of discourse. However, in academic discourse no outstanding differences in the use of different approximators have been observed. Though spoken and written academic discourse differs by the frequency of approximately, this difference is not great.

As was hypothesised, approximators perform multiple functions in spoken and written academic discourse. Those functions appear to depend to some extent on the discourse mode, spoken or written. Generally, approximators are used when more precision is impossible or unnecessary. They are used to generalize certain tendencies, as well as to refer to proportions, percentages, and exceptionally large or small numbers. Approximators are also employed when the speaker/writer refers to hypothetical, future or past situations. The three functions that have been encountered only in spoken academic discourse include the functions of making impromptu calculations, encouraging and self-correction.

Considerable differences can be observed in the frequency of multiple self-distancing devices used alongside approximators in written and spoken academic discourse. Approximators are considerably more frequent in
written than in spoken academic discourse, since in written academic discourse fewer devices for expressing a lack of precision are available. It can be assumed that approximators, as they are sufficiently formal, are one of the main categories used for self-distancing in written discourse. In spoken academic discourse, meanwhile, speakers can rely on a greater variety of such devices. Therefore, here approximators become less important, since other devices can approximate numbers.

References


Contact information:

Jurate Ruzaite
Department of English Philology
Vytautas Magnus University
Donelaicio 58
Kaunas
Lithuania
e-mail: j.ruzaite@hmf.vdu.lt