Fishing for compliments: Precision and recall in corpus-linguistic compliment research

Andreas H. Jucker, Gerold Schneider, Irma Taavitsainen, Barb Breustedt

1   Introduction

Speech acts are not easily amenable to corpus linguistic investigations. They are defined through their illocutionary force or, more rarely, through their perlocutionary effect, and neither of these can be searched for directly. Therefore, speech acts can only be found in large corpora if they appear regularly with standard illocutionary force indicating devices (IFIDs) or in largely routinized forms. These IFIDs and routinized formulae can be used as search strings. In the case of questions, for instance, interrogative sentence constructions and, in writing, the presence of a question mark are typical IFIDs. In the case of apologies the presence of lexemes such as *sorry*, *excuse (me)* or *pardon* are typical features. But the presence of such features is never a guarantee for a particular speech act. The speech act may be carried out in other forms as well, and the forms may occur in other contexts.

Corpus linguistic methods for speech act analysis are discussed in several recent studies. Deutschmann (2003) argues on the basis of his investigation using the British National Corpus that apologies tend to occur in well-defined patterns. In a similar vein, Kohnen (2000a, 2000b, 2002, 2004) argues that directives can be located in historical corpora on the basis of their form. Alternatively, the researcher can retrieve speech act verbs from corpora, as Taavitsainen and Jucker (2007) did when studying speech acts of verbal aggression, but speech act verbs are used performatively only occasionally. If somebody says: "I apologize!" he or she uses the verb performatively and by saying it actually apologizes. Such performatively used verbs are IFIDs in their own right. More often, however, speech act verbs are used descriptively, e.g. the occurrence of a verb like *scorn* or *insult* usually does not indicate the presence of the speech act in its original form, but more often such words are used as labels of speech acts. In most cases they give an account of how somebody on an occasion insulted somebody or perhaps, more abstractly as a whole repertoire with several instances, they reveal an ethnographic view of what was considered insulting in a culture. Likewise, the presence of the word *apology* may indicate on what occasions apologies were needed or how they should be used in general.

Compliments have received a fair amount of attention from various scholars (see Taavitsainen and Jucker this volume and references there). They have been investigated from a cross-cultural perspective, and in particular the responses that people give when they receive compliments have been analyzed extensively. But so far there are no corpus-based investigations, even though there are claims in the relevant literature that compliments are highly routinized and formulaic. In their pioneering study, Manes and Wolfson (1981: 115) claim that "one of the most striking features of compliments in American English is their almost total lack of originality," and Holmes (1988: 452) supports this view: "Compliments are remarkably formulaic speech acts in that a very small number of lexical items and syntactic patterns account for the great majority of them." The formulaic nature of compliments is taken for granted, but a more precise description of the formulae and

their historical development with modern computer-aided methods remains to be done. [1]

2   A methodological challenge and aim of the paper

Empirical studies using authentic data have undergone a paradigm shift in English linguistics in recent decades. Corpus linguistic methods rely on processing large quantities of authentic data with statistical methods. The technical developments of the field have been rapid from the first pilot studies to present-day data-driven accounts of language use, e.g. the *Longman Grammar* (1999) is based on a 40-million-word corpus, and we have dictionaries based on frequency counts of corpora of hundreds of millions of words (for a brief history, see McEnery and Wilson 1996: 1–27). Software to investigate linguistic patterns has also been developed and has become more readily available. In general, corpus linguistics has shifted the emphasis of linguistic analysis to frequently occurring linguistic features and made comparisons of earlier assumptions of frequently occurring patterns possible. The research for this paper is based on the British National Corpus, which contains 100 million words of written and spoken language from various sources in Present-day British English.

It is a well-known fact that the applied methods as well as the corpus design influence the results. It is only reasonable to expect that an assessment of large multimillion databases with naturally occurring spoken language yield somewhat different results from materials picked up by qualitative reading, collected by elicitation or recorded by the diary method. The data on which the early studies on compliments were based were collected by several researchers in various situations in everyday interactions which the researchers observed or in which they participated (Manes and Wolfson 1981: 116). Holmes used the same method and the help of students in collecting her corpus (1988: 446). Manes and Wolfson claim specific frequencies for each of the patterns that they found in their corpus. Of the 686 compliments collected by Manes and Wolfson and their co-researchers, one pattern accounts for more than half of all the compliments. Their database was comprehensive by field linguistic criteria. Holmes used 484 compliment exchanges in her study, which is also a high number. Her observations were in accordance with the patterns established by Manes and Wolfson in whose study "three syntactic patterns accounted for 85% of the 686 compliments in their American corpus. This finding is replicated in the New Zealand data" (Holmes 1988: 452-453). The challenge that we undertake in this chapter is to assess the statements made on the basis of this ethnographic, field-collected material with the help of modern corpus-linguistic tools and a large electronic database of naturally occurring Present-day English, the British National Corpus. We set out to explore what results the patterns established by Manes and Wolfson yield in corpus-based searches and we wanted to test the accuracy of their claims. Our aim is to find out whether the same patterns can be observed in our corpus data, and whether these patterns occur with frequencies that are similar to

---

[1] A preliminary study tracing the formulae was conducted by Magnus Levin a few years ago and presented in a paper read at the ICAME Conference in Verona, 2004. He discovered 544 formulaic compliments of American and British English in spoken corpora (The Longman Spoken American Corpus and British National Corpus respectively). To our knowledge, the paper has not been published.

those in the earlier data. Another aim of our study is to learn more about the nature of compliments and speech acts in order to investigate speech acts more efficiently and improve our corpus-based methods. Thus the emphasis is on developing the methodology of corpus-based searches for pragmatic research tasks and apply the results to historical corpus studies. The test proved more difficult than expected, and two fundamental problems came out. First, the search for the relevant patterns may retrieve a large number of extracts that have the appropriate structure but which are not compliments. This is the problem of precision. In addition the searches may, for various reasons, fail to find all relevant compliments in the corpus. This is the problem of recall and it presents the second major flaw of corpus-based searches.

The paper is structured as follows. In section 3, we discuss our point of departure, the definitions and methods of research, combining automatic searches with manual assessment. Classification problems and inter-annotator agreement issues receive special attention. In section 4, we introduce the main patterns suggested by Manes and Wolfson (1981) and present our search strings approximating the patterns. We discuss them in detail, with special focus on Pattern 1, which is the most frequent one, and we provide illustrative examples from the British National Corpus. We discuss the limitations of our approximations, report the numbers of hits and percentages of examples, and illustrate our precision and recall optimization strategies for this pattern. In section 5, we compare our findings and relative frequencies to Manes and Wolfson (1981), discuss differences and raise relevant questions about possible explanations. Finally, we conclude by bringing the discussion to a more general level and suggesting lines for future research.

3   Points of departure

3.1   Definitions of compliments

Manes and Wolfson (1981: 116) give a very general definition of compliments as "expressions of positive evaluation". They note that there was no need for further elaborations as "the students, naïve native speakers, did not ask that compliments be described or defined for them and indicated no confusion concerning what was expected of them. The data which they collected, with almost no exceptions, were unambiguously identifiable as compliments" (1981: 127). Thus they take a strong "folklinguistic" view on compliments, relying on naïve native-speaker intuition. A more technical definition of compliments is given by Holmes and it is also taken as the point of departure in the historical study of modern compliments in this volume:

> A compliment is a speech act which explicitly or implicitly attributes credit to someone other than the speaker, usually the person addressed, for some 'good' (possession, characteristic, skill etc.) which is positively valued by the speaker and the hearer. (Holmes 1988: 446; 1995: 117)

Although this definition gives more accurate guidelines for the recognition of compliments, the issue proved more complicated and showed that the borderlines are fuzzy.

3.2   Method: Combining quantitative and qualitative assessment

Our method of study consisted of several stages. In preparation for the electronic searches, we translated the patterns given by Manes and Wolfson (1981) into query language. We chose to use the CQP query language (Hoffmann and Evert 2006) for formulating the patterns. The CQP query language has a simple but powerful syntax which allows sophisticated searches for individual words and for lexico-grammatical patterns and supports regular expressions.

   Almost every query method fails to have complete precision and recall. Let us first consider recall. For example, a simple word query has incomplete recall since words containing typos remain unmatched. Queries relying on part-of-speech tags have incomplete recall as the BNC is not completely error-free. Syntactic query patterns based on part-of-speech tags have incomplete recall because it is virtually impossible to account for all possible sequences that can be generated from an inherently hierarchical system such as language.

   The more abstract the linguistic level of the query, the more frequent such errors become. Yet at all levels of language, versions of Zipf's law apply (see e.g. Baroni (2007) for an introduction to Zipf), which states that marked or rare phenomena are extremely infrequent; distributions tail off sharply. This entails that as long as the distribution of the phenomenon under investigation can be assumed to be independent of the distribution of the cases unmatched by a pattern, we get reliable results. Let us consider a search for noun phrases: also very elaborate patterns will fail to find all arbitrarily long and nested noun phrases. If our investigation aimed at finding out the maximum length of a noun phrase, such an approach would be totally inappropriate, there would be a very strong dependence between the investigation and the unmatched patterns. If our investigation aimed at describing the use of the definite versus the indefinite article, however, it is reasonable to assume that the few very long unmatched noun phrases will not show a behavior that is fundamentally different, and we can also expect that if we carefully write and test increasingly sophisticated patterns, the loss in recall will be very small, since the unmatched, very complex noun phrases, are extremely few.

   As for precision, a simple word query leads to errors if a token has a rare part-of-speech tag (for example the word *can* as a noun in *can of beer*), if it is an abbreviation (for example *can* as an abbreviation for *Canada*). Syntactic query patterns based on part-of-speech tags lead to many precision errors because of the lack of any parsing context and because of possible tagging errors in the BNC mentioned above. In a query looking for verb-object relations, for example, the sentence *Experts fear the virus will spread* a pattern-based approach inevitably returns a verb-object relation between *fear* and *virus*. In typical Corpus Linguistic methodology, results are filtered manually, so that precision errors are not a serious problem, until the number of hits exceeds what is possible to scan manually, and until precision falls below a certain threshold: one tends to overlook positive examples if precision is much lower than 1%.

   The statistics derived from the counts reported in BNC can thus be assumed to be reliable if we assume that they can be extrapolated to the few unmatched cases. The individual limitations of each pattern will be described separately.

   The results of Pattern 1 were overwhelming, and for the second phase we had to revert to a random sample of utterances to make the qualitative analysis possible. A representative sample of 300 examples was studied independently in order to screen the relevant examples from the "noise", i.e. the irrelevant examples.

3.3 Inter-annotator agreement

Classification problems are often faced in corpus-linguistic studies as there is always a subjective element in qualitative studies. The problems are often passed without recognition, but we wanted to refine the method by introducing a well-established practice used in other disciplines to our linguistic study. Two annotators independently annotated the hits returned by our patterns, which allows us to measure inter-annotator agreement and to asses the epistemic status of compliments as a linguistic category. We proceeded as follows for the annotation. After discussing the status of compliments in linguistics according to the definitions given above and considering some examples from the BNC corpus, two authors of this paper annotated the patterns individually, i.e. classified the material into "compliments" and "other". Annotator 1 found 237 compliments in pattern [1b], annotator 2 found 290 compliments. Annotator 1 found 26 compliments in pattern [1ab], annotator 2 established 28 compliments. Differing opinions on classification may lead to low inter-annotator agreement. Inter-annotator agreement was measured as the sum of the cases where both annotators regarded the match as a compliment divided by the sum of the cases where at least one annotator annotates the match as a compliment. For the fully manually annotated subpatterns of pattern [1], i.e. patterns [1b] and [1ab] together this is 249/339 = 73.5%. For all the fully manually annotated patterns (not only pattern 1) inter-annotator agreement was 76.6%. For our compliment count we use the conservative approach of only counting matching cases where *both* annotators classified the utterance as a compliment.

4   Assessment of the patterns

We now take up the formulae discussed in Manes and Wolfson (1981: 120-121), examples 29 to 37 in their text, and give illustrative examples.

       The patterns as described in the following are all approximations, as we have mentioned in section 3.2. They share the general limitation that they fail to cover some embedded forms (Manes and Wolfson 1981: 121). They give three examples of embedded forms. Their examples (38) and (40) (ibid.) match a simple, intuitive pattern: a simple noun phrase followed by a verb (often a copula) and a positive evaluation centered on an adjective. (38) is repeated here as (1), (40) as (2), italics added.

(1) I think *your hair looks good* this way.

(2) Why don't you just accept the fact that *you did a good job*

But (39), repeated here as (3), involves a surface word order alteration which entails that the corresponding pattern fails to match.

(3) By the way, I have to tell you how *professional* I thought *your magazine looked.*

If we are ready to assume that examples such as (3) are not fundamentally different with respect to the phenomenon of compliments from (1) and (2), then we can obtain reliable descriptions using these patterns. We will now discuss the individual patterns.

4.1 Pattern 1

The first pattern is represented as follows in Manes and Wolfson (1981). The number in brackets represents the percentage of compliments in their corpus of elicited examples which match the pattern given in [1].

$$[1] \quad \text{NP} \left\{ \frac{is}{looks} \right\} (\textit{really}) \text{ ADJ } (53.6\%)$$

In this and all the other patterns, NP stands for a noun phrase, which typically includes a second person possessive determiner or a demonstrative determiner. It may also stand for a personal or demonstrative pronoun. The curly brackets signify an option, and the round brackets an optional element. Verbs are cited in the present tense, but may occur in other forms. *Look* stands for any linking verb other than *be* (*look*, *seem*, *smell*, *feel*, ...). *Really* stands for any intensifier (*really*, *very*, *so*, *such*, ...). And ADJ stands for any semantically positive adjective. Examples reported by Manes and Wolfson (1981: 121) include *Your hair looks nice* or *This is really good*.

In CQP pattern [1] can be approximated as

```
[1a] _NN* (is|'re|are|were|look*|seem*) (really|very|such|so) _AJ0
```

It returns 7690 matches from the BNC. Pattern [1a] is an approximation. It both overgenerates and undergenerates with respect to pattern [1] of Manes and Wolfson. Overgeneration leads to precision errors, undergeneration to recall errors. As mentioned above, moderate overgeneration is no problem because manual filtering is used.

i. It overgenerates, because the final adjective is unrestricted. The list of 72 adjectives collected by Manes and Wolfson is too long to expect that it is complete. In order to keep recall levels acceptable, we have thus decided not to restrict the adjective.

ii. It undergenerates, because the list of linking verbs is open in Manes and Wolfson (1981: footnote 5). They do not supply information on other linking verbs. They implicitly suggest that they can be assumed to be very rare. It is generally known that the set of copular verbs is closed and dominated by *be*, and that frequencies sharply tail off.

iii. It undergenerates, because the list of intensifiers is also open in Manes and Wolfson (1981: footnote 5). Again, we can assume that the list of intensifiers is closed and tails off.

iv. It seriously undergenerates because it only reports cases that include an intensifier. Manes and Wolfson (1981: 118-119) point out that intensifiers occur in over a third of the data. A modification of pattern [1a] with an optional intensifier reports 114252 matches, a number that is too big to allow manual filtering, and that will have extremely low recall (see the precision of pattern [1a] below). We suggest the working assumption that the distribution of compliments with and without intensifiers is similar.

v. It undergenerates, because the search is restricted to NPs that end with a noun. This point is problematic. The discrepancy is greater here than in ii. and iii. , especially as Manes and Wolfson (1981: 119) observe that 75% of all compliments in their data include second person pronouns or demonstratives.

While it can be expected that the undergeneration mentioned in the points *ii*. and *iii*. (closed list of linking verbs and intensifiers) is minimal, this cannot be expected for point *v*. Manes and Wolfson (1981: 119) observe that 75% of all compliments in their data include second person pronouns or demonstratives, which remain unmatched with pattern [1a]. As a remedy, we have thus formulated patterns for second person pronouns and demonstratives, namely patterns [1b], [1c], [1aa] and [1ab]. In pattern [1b], the initial NP is the second person pronoun *you*.

```
[1b] you ('re|are|were|look*|smell*|seem*) (really|very|such|so) _AJ0
```

It returns 1226 matches. A compliment example is given in extract (4).

(4) Mrs Browning was hardly less excited than Ferdinando. Paying her first visit to the Casa Guidi Wilson was moved to exclaim, "Why, ma'am, **you look so well!**" Mrs Browning laughed and made a gesture of dismissal. "Oh, I am tired of being told so, Wilson."
(BNC ADS 763-765)

In pattern [1c] the initial NP is a demonstrative pronoun. We have observed that [1c] overlaps with pattern [3]. In order to exclude this overlap, an extended pattern was actually employed, given as [1cX], which excludes cases where the final adjective is followed by a noun. This extended pattern could only be formulated in extended CQP syntax. The queries and the number of returned matches are listed in the following.

```
[1c] _DT0 (is|'re|are|were|look*| smell*|seem*) (really|very|such|so)
_AJ0
 [1cX] [pos = "DT0" %c] ( [word = "is" %c] | [word = "'re" %c] |
[word = "are" %c] | [word = "were" %c] | [word = "look.*" %c] | [word
= "smell.*" %c] | [word = "seem.*" %c] ) ( [word = "really" %c] |
[word = "very" %c] | [word = "so" %c] | [word = "such" %c] ) [pos =
"AJ0" %c] [pos != "NN.*"]
```

Pattern [1c] returns 820 matches, pattern [1cX] reduced this to 721 matches. A compliment example is given in extract (5). It is noteworthy that this compliment is immediately followed by a second one. The second compliment corresponds to pattern 7 discussed below.

(5) "Let's have a look at your book then." Quickly she would flick through his exercise books, glancing from page to page as they flew by. "Oh, **this is very good**, Alan. What a brainbox you are!"
(BNC HJH 546-549)

A version of pattern [1a] where the noun head of the initial NP is preceded by the second person pronoun *your* is [1aa].

```
[1aa] your _NN* (is|'re|are|were|look*| smell*|seem*)
(really|very|such|so) _AJ0
```

It returns 716 matches. A compliment example is given in extract (6).

(6) "What about you, Megan. You're lookin' grand. **Your hair is so long!**" "Aye, I'm
   a real woman now.
   (BNC HGL 1326-29)

Without context the sentence *Your hair is so long* cannot be identified as a
compliment but with the preceding compliment, *You're lookin' grand* (pattern 1a) and
in particular with the following compliment response it is clear that the two speakers
involved in the example treat *Your hair is so long* as a compliment.

   A version of [1a] where the noun head of the initial NP is preceded by a
demonstrative pronoun is [1ab].

```
[1ab] _DT _NN* (is|'re|are|were|look*| smell*|seem*)
(really|very|such|so) _AJ0
```

It returns 94 matches. Extract (7) contains two relevant examples.

(7) Tracy: **Those curtains look really nice**.
   Annette: Do they look nice? <pause> Yeah that's nice.
   Teresa: It's got a huge bobble on it.
   Annette:Yeah, Tracy said **your curtains look really nice**.
   (BNC KB9 111-115)

The 7690 matches of pattern [1a] are too many for complete manual inspection. We
have selected a random subset of 300 matches for manual inspection. The manual
inspection revealed that out of the 300 examples, only one was a compliment, defined
by the criteria of the definition. This finding corresponds to a precision of 0.33%.
This indicates that precision for this subset is generally very low, most likely below
1%. If we extrapolate to the entire BNC we can expect only about 25 compliments to
correspond to pattern [1a].

   We have mentioned that Manes and Wolfson (1981: 119) observe that 75% of
all compliments in their data include second person pronouns or demonstratives, cases
which pattern [1a] largely fails to include. The patterns that we have formulated for
these cases, [1b], [1cX], [1aa] and [1ab] can thus be expected to deliver the bulk of
compliments following pattern 1. Since the matches returned by these patterns are
considerably fewer, we have manually inspected most of them, and we can expect
precision to be considerably higher. The qualitative assessment is necessary as it is
the context of utterance and the response to it that defines whether the phrase can be
classified as a compliment.

   Precision of pattern [1b] indeed turned out to be much higher, about 20%. It
would be higher still if only a closed list of positive adjectives were used, but since
the list of adjectives compiled by Manes and Wolfson (1981) seems to be too large to
be closed we left the adjective unrestricted. In addition to the "expected" adjectives
that Manes and Wolfson list, such as *good*, *nice*, we also found many adjectives that
they do not list, for example *talented*, *sexy*, *friendly*, *wise*, *kind*, *lovely*, *fit*, *sensible*,
etc. Examples include the following.

(8) He poured the wine and lit a cigarette for himself. "I won't offer you one. I'm sure
   you don't smoke. **You look so fit**." – "I am fit. I swim thirty lengths twice a week.
   I work out with weights for two hours on Saturdays.
   (BNC A0R 1231-1237)

(9) "I do not see that anything else will do, not for the moment. The situation is too
   far gone. This is what she wants and so perhaps it is what she needs. We can only
   carry out her wishes, we can only try." – "**You are so sensible**, you have always

seen things in a clearer light." Florence Ames shook her head. "It does not make me happy. I have seen too much of this."
(BNC AD1 2842-2849)

We will now discuss some of the compliments that we have found using pattern [1b] and classification. We shall pay special attention to problem cases of classification. The comments we make also apply to all other patterns. It is always very important to look at the context carefully. Even seemingly very positive evaluations can be very remote from compliments. In the following example, a typical compliment phrase occurs as a response to a scaring gesture; obviously the utterance is not a compliment at all but a collaborative second part to the posed question, playing along in the game.

(10)   As Estella was leading me along the dark passages, she stopped suddenly and put her face close to mine. "Look at me, boy! Am I pretty?" – "Yes, I think **you're very pretty**." – "Am I rude to you?" – "Not as much as last time." – She hit my face as hard as she could. – "Now, you coarse little boy, what do you think of me?" – "I won't tell you."
(BNC FPU 542-550)

An example that we have judged to be a compliment, although in very grave circumstances, is the following.

(11)   "You see" -- his voice trembled slightly, his blue eyes became haunted, his bloom of good looks seemed to collapse inwards -- "I have cancer. I'm told I have six months to live. If you could write something -- anything -- I'd appreciate it so much." –"My God, that's terrible!" I said. "And **you look so well**."
(BNC AE0 1951-1956)

Some compliments are forced, the receiver of the compliment clearly expects to be complimented. We have decided to classify them as compliments, even if they are not deliberate compliments, but concrete examples of "fishing for compliments".

(12)   "Still, the blouse and skirt don't look too bad, d'you think?" – "**You look very nice**, Dolly." – "Like to take me out, would you?"
(BNC CKE 2304-2306)

An especially problematic set of examples is provided by ironical or playful compliments. They are particularly frequent in fiction (see Jucker and Taavitsainen this volume). The following example represents ironical compliments. The negative evaluations and more subtle meanings have to be assessed in context of the unfolding discourse. We do not count them as compliments if the meanings are clearly ironical, turning the positive surface utterance into negative evaluation (cf. the definition above).

(13)   "Oh, Squadron Leader Latimer, **you're so brave**. Marry me, and make me happy ever after." She poured out her scorn, and with it her jealousy and frustrated rage. Johnny remained silent for a time, and then said: "Shut up Bella. Hold your tongue."
(BNC G1S 2961-2966)

Playful compliments presented problems, and examples with playful meanings provide many of the cases in which the two annotators often had differing opinions (for inter-annotator agreement, see above). Especially the spoken part of the BNC contains many such playful compliments. Problems were presented by cases where a positive evaluation is potentially outweighed by a negative connotation, as in the

following example, which was annotated as compliment by one annotator, but as no compliment by the second annotator.

(14)   "My precious White Rose!" murmured the queen-dowager. "**You are so young**, so tender -- you know not the wickedness of the world, of devious and ambitious rascals. How could you know! …"
(BNC CCD 1401)

Many examples may have a compliment component, but it is unclear, or they only have very little compliment force, often using conventionalized phrases. While we in principle agreed not to count them as compliments inter-annotator agreement was also relatively low on these examples.

(15)   I know immediately that I wouldn't like to go; I would hate to be a servant in a posh house, but I find it difficult to say this. Eventually, at tea time, I tell Nicola's mother that I can't go to India because I'm starting a course at university. She says in her calm, posh voice that this is fine, and continues to stir the curry she is making for their evening meal. Do I like curry? she wants to know. Feeling guilty again, for not liking curry and for not wanting to go to India and letting her down, I say no, I will just have a boiled egg and toast, I'm not very hungry, that will be fine. "**Your needs are very small**," she remarks. I think it is an odd comment. What does she mean? Obviously, she is talking about the boiled egg and toast, but it is a funny way of putting things, talking about "needs" and not appetite or eating habits. It sticks in my mind. For some reason, it disturbs me.
(BNC ADG 178-188)

Expressions that look like compliments are often used as a conventionalized phrase to introduce a request or kindly reject an offer or indeed a compliment. We have excluded obvious cases from the compliment class, but again inter-annotator agreement is relatively low.

(16)   Kee looked at Conway. He said, "I am a houngan, and I understand voodoo. I know you come from America, and you Americans do not believe in things like that. But I can do many things that you do not understand. I will help you if you are good to me." – "**You're very kind**," said Conway, laughing at the old man. "But I really don't think your voodoo can give me all the things I want in life."
(BNC GWA 406-412)

(17)   After the meal, Dolores had cut up a huge melon, and dished out chunks on their plates. Then she brought them all strong black coffee. "It is an honour to have you here," she told Shelley. – "**You're very kind** to me."
(BNC JYA 3223-3226)

Some compliments appear as a part of a prayer or a ritual ceremony. We have decided to view them as conventionalized or ritual phrases outside the scope of normal interaction of speakers and hearers in the everyday world, and therefore we do not annotate them as compliments. In the examples below, the use of the second person is more generic and ritualistic (example 16). An address to God (example 17) does not count as a compliment either.

(18)   Meanwhile Panna, despite her bulk, was putting on a fine display. She wobbled her head one way, wobbled her bottom the other, all the while singing an Urdu verse which Zakir translated as follows: God bless you, **You are very sweet**,

You are very lovely, God will give you long life.
(BNC H89 652-657)

(19)   Songs such as "Father God I wonder ... .", "You are here ... .", "Lord **you are so precious** to me ... ." are appropriate here, with lines that express this intimacy.
(BNC C8L 1451-1453)

There are a number of meta-compliments in the BNC, comments about compliments. One of the annotators excluded obvious cases, such as the following.

(20)   Once upon a time (I said, and he stared bitterly bitterly at the floor) there was a very ugly monster who captured a princess and put her in a dungeon in his castle. Every evening he made her sit with him and ordered her to say to him, "**You are very handsome**, my lord." And every evening she said, "You are very ugly, you monster." And then the monster looked very hurt and sad and stared at the floor.
(BNC G07 2338-2341)

Pattern [1b] delivers 226 cases which both annotators have marked as compliment, pattern [1ab] 23 cases. Pattern [1cX] returns 721 matches, on which manual inspection of a 100 random sample showed that precision is very low (about 1 %). Pattern [1aa] returns 716 matches, on which manual inspection of a 100 random sample showed that precision is very low (also about 1 %).

4.2 Pattern 2

Manes and Wolfson's (1981) second pattern is represented as follows.

$$[2] \qquad \text{I (really)} \begin{Bmatrix} like \\ \hline love \end{Bmatrix} \text{NP (16.1\%)}$$

In this pattern, *like* and *love* stand for any verb of liking (*like*, *love*, *admire*, *enjoy*, ..). Examples reported by Manes and Wolfson (1981) are *I love your hair* and *I really like those shoes*. Examples that we found in the BNC include the following.

(21)   "**I really admire you**, bringing up four from the time the youngest was only five *and* working full time." "No bravery. Circumstances dictated it."
(BNC ABW 2307-2309)

(22)   As a newcomer to the sport, **I really enjoy your Saturday golf pages**.
(BNC CEK 1877)

Pattern 2 was approximated as

```
I (really|very|so|such) (like|love|admire|enjoy)
(_AT0|_DT0|_NN*|_DPS|_PNP)
```

It delivered 94 matches. This approximation over- and undergenerates with respect to pattern 2, the reasoning is similar as in pattern 1.

11 of the 94 matches were classified as compliments by both raters; precision is thus about 12%. The distribution is very irregular. In 6 of the cases which were not classified as a compliment, the final NP is the second person pronoun *you*, 10 are direct speech or from spoken parts of the BNC. There are 3 cases where the final NP

starts with the pronoun *your*. Compared to pattern 1, which delivered far above a hundred compliments, pattern 2 delivers very few.

4.3 Pattern 3

The third pattern is represented as follows.

> [3]   PRO is (really) (a) ADJ NP (14.9%)

As above, ADJ stands for any semantically positive adverb. PRO stands for a personal or demonstrative pronoun, *you*, *this*, *that*, *these* or *those*. Examples reported by Manes and Wolfson (1981) are *That is a nice piece of work* and *This was really a great meal*. An example that we found in the BNC is the following.

(23)   On Wednesday the tea party started very well. "**These are very good cakes**, Miss Cuthbert," Mrs Allan said to Marilla.
(BNC FPT 309-310)

Pattern 3 was approximated as follows.

```
[3a](_DT0|you) (is|are|'re|were) (really|very|such|so) _AJ0 _NN*
```

```
[3b](_DT0|you) (is|are|'re|were) (really|very|such|so) a _AJ0 _NN*
```

Pattern [3a] has 115 matches, [3b] 46 matches. [3a] contains 3 compliments. [3b] contains 5 compliments. Contrary to the results in Manes and Wolfson (1981), pattern 3 delivers very few compliments. Most of the matches of [3b] are clearly negative assessments and therefore not compliments. Again, there is a marked difference to the results of Manes and Wolfson (1981).

4.4 Pattern 4

The first three patterns account for 85% of all compliments in Manes and Wolfson's (1981) corpus. In addition, six other patterns, patterns 4 to 9, emerged (Manes and Wolfson 1981: 120-121 and footnote 5, p.132). These nine patterns together cover 97.2% of the data.
    The fourth pattern is represented as follows.

> [4]   You V (a) (really) ADJ NP (3.3%)

An example from Manes and Wolfson is *You did a good job*. An example that we found in the BNC is given in extract (24).

(24)   "I'm here on a two-year contract. **You speak very good English**." A chink is temporarily exposed. He smiles to himself in the mirror. "It's for the job," he says proudly, "I get promotion if I speak good English."
(BP8 394-398)

Pattern 4 was approximated as

```
[4a] you _V* (really|very|so|such) _AJ0 _NN*
[4b] you _V* (really|very|so|such) a _AJ0 _NN*
```

Pattern [4a] has 75 matches, [4b] has 46 matches. [4a] contains 10 compliments, precision is thus quite high, about 13%. [4b] contains 7 compliments, precision is high, about 15%.

4.5 Pattern 5

The fifth pattern is represented as follows.

[5]   You V (NP) (really) ADV (2.7%)

An example from Manes and Wolfson is *You really handled that situation well*. An example that we found in the BNC is given in extract (25).

(25)   "It was Bach?" "Telemann." "**You play very well**." "Once, I *could* play. Never mind. "
(BNC G13 357-361)

Pattern 5 was approximated as

```
[5a] you _V* (really|very|so|such) _AV0
[5b] you _V* _NN* (really|very|so|such) _AV0
[5c] you _V* _AT0 _NN* (really|very|so|such) _AV0
```

Pattern [5a] has 409 matches. A random 100 sample of [5a] contained 5 compliments. By linear extrapolation we can expect about 20 compliments from [5a] in the BNC. Pattern [5b] has 12 matches, and [5c] has 10 matches. The approximation of the optional NP to a noun or a determiner and noun is very crude, but the very low numbers suggest that these 2 patterns deliver only few compliments. [5b] contains 2 compliments, 5c none.

4.6 Pattern 6

The sixth pattern is represented as follows.

[6]   You have (a) (really) ADJ NP (2.4%)

An example from Manes and Wolfson is *You have such beautiful hair*. An example that we found in the BNC is given in extract (26).

(26)   He raised her hand to his lips and kissed it, then continued to hold it against his lips. "Why didn't you tell me **you had such a beautiful name**?"
(BNC JXT 3146-3147)

Pattern 6 is a subset of pattern 4 in our approximation. Manes and Wolfson (1981) explicitly mention that they treat *have* separately. We will thus report joint results for pattern 4 and 6 when we compare our findings to Manes and Wolfson in chapter 5.

4.7 Pattern 7

The seventh pattern is represented as follows.

[7]   What (a) ADJ NP! (1.6%)

An example from Manes and Wolfson is *What a lovely baby you have!* An example that we found in the BNC is the following.

(27)   "Hot soup is so restoring," she said. " *So restoring* !" cried Madame Maillot, or whatever her name is. "**What a perfect expression!** Who but you could think of it?
(CA6 1401-1404)

Pattern 7 was approximated as

```
[7a] what a _AJ0 _NN* !
[7b] what _AJ0 _NN* !
```

Pattern [7a] has 134 matches, [7b] has 13 matches. Pattern [7a] contains 10 compliments. Pattern [7b] contains no compliment.

4.8 Pattern 8

The eighth pattern is represented as follows.

[8]   ADJ NP! (1.6%)

An example from Manes and Wolfson is *Nice game!* Two examples that we found in the BNC are the following.

(28)   There was very little inconvenience in leaving out the butter and salad cream, and I have enjoyed the diet even more whilst watching the inches disappear, and enjoying being complimented on how much slimmer I look. Oh, and how lovely to be able to open the wardrobe doors and say "I haven't got anything to wear -- they are all *too big*. "**Absolute Heaven!** Very many thanks.
(BNC BNS 173-176)

(29)   "… little support <pause> are most affected <pause> and we need to er, go back to what Kathleen was saying about education, we need to <pause> help young girls get self-confidence, more se--, coping skills to deal with these pressures, but also, as other people have said, we need to get the government to look at the fact as, that <pause> encourage women to keep smoking <pause> and address those issues as well. " -- "**Good discussion!** Thank you all very much indeed!"
(BNC FLM 423-425)

Pattern 8 was approximated in CQP as

```
[8] . _AJ0 _NN* !
```

It has 388 matches in the BNC. We found 3 compliments in a random 100 matches subset, which extrapolates to about 11 compliments in all the matches.

4.9 Pattern 9

The ninth pattern is represented as follows.

[9]   Isn't NP ADJ! (1.0%)

We have approximated pattern 9 in CQP as

```
[9] is {not} _NN* _AJ0 !
```

An example from Manes and Wolfson is *Isn't your ring beautiful!* The pattern does not return any match from the BNC. Modified versions of the pattern, for example,

```
[9a] is {not} _NN* _AJ0 ?
```

return few matches (63 in the case of [9a]) but none of them is a compliment. We have therefore not found any compliment of this type in the BNC.

5   Discussion

5.1 Query problems

As shown above, it is possible to search modern tagged corpora with compliment formulae like "NP is/looks [intensifier] ADJ"; "I [intensifier] like/love NP"; "PRO is [intensifier] ADJ NP". Preliminary tests on the tagged version of BNC showed that the patterns are approximations that either over- or undergenerate for various reasons as the patterns that we have used are approximations. The main shortcoming is that they require an intensifier in order to alleviate the filtering task. Either intensifiers are distributed extremely unhomogenously across the patterns, or explanations for the striking facts that while pattern 1 occurs far more frequently, and patterns 2 and 3 occur far less frequently than expected need to be found. The comparison between expected contributions from Manes and Wolfson versus our BNC data is shown textually in table 1, and graphically in figure 1. The BNC counts and percentages reported are based on complete manual rating for patterns 2, 3, 4 & 6, 7, and 9, and on linear extrapolation from random subsets for parts of pattern 1 – as explained in detail in section 5.2 – and for patterns 5 and 8.

Approximation to pattern	BNC compliment counts	BNC compliment %	Manes and Wolfson compliment %
1	262	76.4%	53.6%
2	11	3.2%	16.1%
3	8	2.3%	14.9%
4 & 6	17	5.0%	5.7%
5	22	6.4%	2.7%
7	12	3.5%	1.6%
8	11	3.2%	1.6%
9	0		1.0%
TOTAL	**343**	**100%**	**97.2%**

Table 1: Compliment pattern frequencies in Manes and Wolfson's data and in the BNC.

There are a number of reasons why the surface patterns used are crude. For example they cannot catch repairs, hesitations, marked constituent order etc.; approximations to higher level constituents such as NPs are often crude, and they depend on intensifiers. The differences are big enough to warrant closer investigation, however.
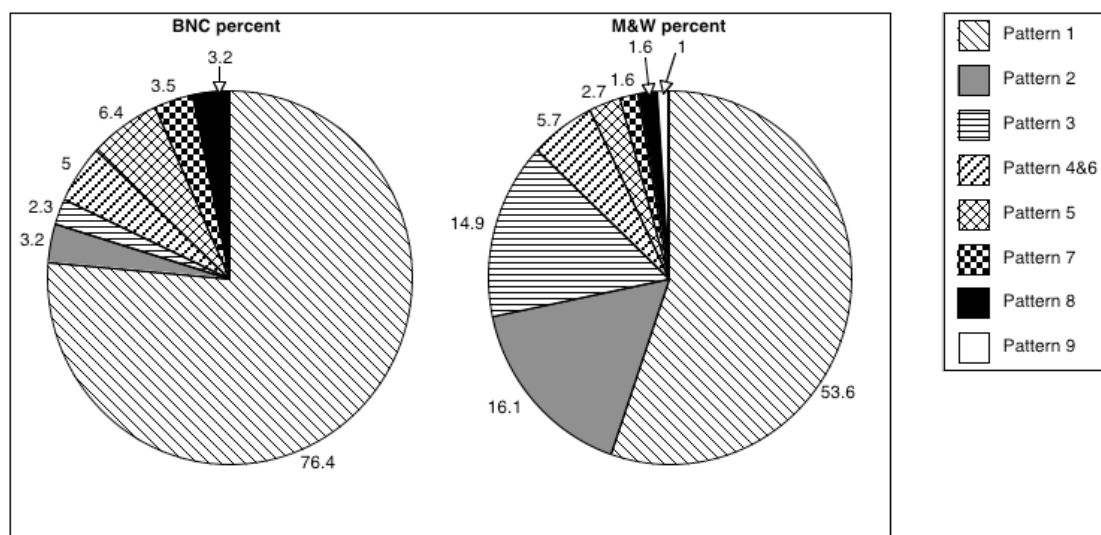
Figure 1: Compliment pattern frequencies in the BNC graphically compared to Manes and Wolfson's (M&W) data.

5.2 Precision and recall revisited

As mentioned in the introduction, the two major problems revealed by our study were those of precision and recall. We have described in section 4.1 how we approximated Manes and Wolfson's pattern 1 by means of several subpatterns. Both precision and recall of pattern [1b] are much higher than precision and recall of pattern [1a]. Recall of [1a] was 0.33% in a random subset of 300 matches, recall of [1b] on the whole BNC is about 20%. Since the total number of compliments expected to be found by pattern [1a] is quite small, and since a manual inspection of all the matches is prohibitive, we have allowed ourselves to assume that the random sample is representative of all the matches of pattern [1a]. We now address the question of precision and recall of our other patterns corresponding to Manes and Wolfson's pattern 1.

Precision of pattern [1cX] turns out to be low, only about 1%. A random 100 sample contained one example. By linear extrapolation we can expect about 7 compliments coming from this pattern in the entire BNC. Precision of pattern [1aa] is also only about 1%. A random 100 sample contained one example. By linear extrapolation we can expect about 6 examples in the BNC. Precision of pattern [1ab] is higher again, about 25%, the matches contained 24 compliments.

Patterns [1cX] and [1b] correct a recall error of [1a] (they are extensions of [1a]), patterns [1aa] and [1ab] are versions of [1a] with higher precision, they are in fact specific subpatterns of [1a], which entails that they must have lower recall than [1a]. It is difficult to assess the amount of loss of recall precisely, but as a rough indication we can compare the linear extrapolation of [1aa] plus the manual count of [1ab] (29 cases) to the linear extrapolation of [1a] (25 cases). If the linear extrapolation of [1a] was considerably higher than the linear extrapolation of [1aa] plus the manual count of [1ab] this would indicate poor recall.

If we add the complete results of the manual annotation of patterns [1b] and [1ab] to the linear extrapolation counts for patterns [1cX] and [1aa], we can assess the

number of compliments following pattern 1 in the BNC. We can expect slightly above 250 compliments from pattern 1, as summarized in table 2.

Pattern	Count	Method
[1b]	226	Complete manual annotation, annotators agree
[1ab]	23	Complete manual annotation, annotators agree
[1cX]	7	Extrapolation from random sample
[1aa]	6	Extrapolation from random sample
TOTAL	**262**	

Table 2. Frequencies of compliment patterns in BNC, manual assessment.

If the observation in Manes and Wolfson (1981: 120) that pattern 1 delivers slightly above half of all cases carries over from their diary collection method to our corpus search, then we can only expect about 500 compliments with intensifiers in the entire BNC, fewer even if we consider that our data seems to be more dominated by pattern 1 (see table 1). In Manes and Wolfson's data over a third of the compliments contain an intensifier (1981: 118-119). The total number of compliments would then hardly reach 2000 cases.

6   Conclusions

The qualitative assessment with two independent annotators reinforced our previous view of speech acts as fuzzy notions (Jucker and Taavitsainen 2000). There is always a subjective element in interpretation, and it is the context that decides. It provides the clues for interpretations as meanings are negotiated. Both illocutions and perlocutions count. Computerized searches are capable of locating locutions but qualitative assessments are needed in pragmatic research to reveal local meanings of the utterances.

Manes and Wolfson argue that only the ethnographic method is a reliable method for studying compliments. By ethnographic method they understand what might be called the diary method or the participant observation method. They argue that other types of data, such as novels or plays, are unsuitable because they conform to artistic requirements and they do not "reflect exactly the complexity of actual speech use" (1981: 115). The complexities and difficulties of interpretation came out clearly in our corpus study. Ironical utterances, even with opposite pejorative meanings, may have the same surface structure as compliments proper. Ritual and religious uses provide further cases in point in which the utterances cannot be taken at their face value. Genre restrictions have to be taken into account, but once they are recognized and qualitative assessments of utterance meaning carried out, a wider range of material is perfectly acceptable for speech act studies and pragmatic research in general.

The frequencies of individual patterns that we found in our data differ considerably from those reported by Manes and Wolfson (1981). But is must be remembered that they collected their examples through the diary method. It is possible that some patterns were more salient to the collectors and therefore were more likely to be picked up and to be included in their collections. Our own investigation is heavily indebted to Manes and Wolfson because we rely on the patterns that they established on the basis of their data. It is very likely that there are

other compliments hidden in the BNC that do not conform to any of the patterns established by Manes and Wolfson. However, in order to find out how many compliments we missed, a manual search of a substantial corpus would be needed, i.e. a bottom up approach in the sense of Kohnen ("Tracing directives through text and time"; this volume). Ultimately we would need large pragmatically tagged corpora. Such corpora are not yet available even though some steps have already been taken in that direction (Culpeper and Archer, this volume). On a large scale pragmatic tagging cannot be carried out manually. If it is to be carried out automatically, we will presumably have to rely on (improved versions of) search algorithms like the ones that we developed for this paper.

No big corpora with sufficient material for a study like the present one were available at the time Manes and Wolfson (1981) or Holmes (1988) conducted their studies. The option of using corpora, or of arguing for or against their use did not pose itself then, but the time has come to consider new options for retrieving material for pragmatic research tasks. For example, it is perhaps possible to develop the patterns for lexical searches and thus improve the precision and recall of computerized searches.

Several new lines of study emerge from the present one. Since the BNC contains varied genres, including fiction, a closer investigation of the dispersion across genres, as far as the low counts of the relatively rare phenomenon of compliments allows, would merit further consideration. Another major research line opens up with tagged historical corpora, but steps towards that direction remain to be taken in a later study.

References

Baroni, Marco. (2007) Distributions in text. In Anke Lüdeling and Merja Kytö (eds.), Corpus linguistics: An international handbook, Berlin: Mouton de Gruyter, to appear.

Deutschmann, Mats. (2003) Apologising in British English. (Skrifter från moderna språk 10). Umeå: Institutionen för moderna språk, Umeå University.

Hoffmann, Sebastian, and Stefan Evert. (2006) BNCweb (CQP-Edition): The marriage of two corpus tools. In: Sabine Braun, Kurt Kohn and Joybrato Mukherjee (eds.). Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods. Frankfurt am Main: Peter Lang, 177-195.

Holmes, Janet. (1988) Paying compliments: A sex preferential politeness strategy. Journal of Pragmatics 12.4, 445-465.

Jucker, Andreas H., and Irma Taavitsainen. (2000) Diachronic speech act analysis: Insults from flyting to flaming. Journal of Historical Pragmatics 1.1, 67-95.

Kohnen, Thomas. (2000a) Explicit performatives in Old English: A corpus-based study of directives. Journal of Historical Pragmatics 1.2, 301-321.

Kohnen, Thomas. (2000b) Corpora and speech acts: The study of performatives. In: Christian Mair and Marianne Hundt (eds.). Corpus Linguistics and Linguistic Theory. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20). Freiburg im Breisgau 1999. Amsterdam: Rodopi, 177-186.

Kohnen, Thomas. (2002) Towards a history of English directives. In: Andreas Fischer, Gunnel Tottie and Hans Martin Lehmann (eds.). Text Types and Corpora. Studies in Honour of Udo Fries. Tübingen: Gunter Narr, 165-175.

Kohnen, Thomas. (2004) "Let mee bee so bold to request you to tell mee": Constructions with 'let me' and the history of English directives. Journal of Historical Pragmatics 5.1, 159-173.

Levin, Magnus. (2004) A corpus-based approach to compliments in American and British English: Regional variation and sex-based differences. Paper delivered at ICAME25 Conference in Verona.

McEnery, Tony, and Andrew Wilson. (1996) Corpus Linguistics. (Edinburgh Textbooks in Empirical Linguistics). Edinburgh: Edinburgh University Press.

Manes, Joan, and Nessa Wolfson. (1981) The compliment formula. In: Florian Coulmas (ed.). Conversational Routine. Explorations in Standardized Communication Situations and Prepatterned Speech. The Hague: Mouton, 115-132.

Taavitsainen, Irma, and Andreas H. Jucker. (2007) Speech acts and speech act verbs in the history of English. In: Susan M. Fitzmaurice and Irma Taavitsainen (eds.). Methodological Issues in Historical Pragmatics. Berlin: Mouton de Gruyter, 107-138.