

A New Look at Teaching Comparisons – a Corpus-Based Study

Ute Knoch

University of Auckland, New Zealand

Abstract

Most current ESL textbooks discuss how to 'make comparisons' in English. This is typically accomplished by teaching the comparative form of the adjective generally directly followed by a 'than' clause (explicit basis of comparison). This study aims to look at other ways of making comparisons used by the native speaker and investigates the frequency of the explicit basis of comparison as it occurs in native speech. For this, the British National Corpus (BNC) was investigated. Results show that ways of comparing other than adjective comparatives are more frequently used and that the explicit basis of comparison is used only occasionally by native speakers. It is argued that this has clear implications for ESL teachers and textbook writers.

Introduction

The function of 'making comparisons' is usually realized by teaching lengthy rules about the formation of comparatives using adjectives and adverbs, about when to use the inflectional ending *-er/-est*, and when to use the periphrastic forms *more/most*. Some textbooks also teach the use of equatives or negative equatives. Several authors (Celce-Murcia et al., 1999, Quirk et al., 1985, Kennedy, 1996 and Mitchell, 1990) show that native speakers, in fact, use a variety of different constructions to compare and contrast in English. Amongst these are certain lexical items (e.g. *compared to/with*) and contrasting connectives (e.g. *but, yet*).

The aim of this study is to determine which comparative constructions in English are most commonly used by native speakers. This information was collected from the British National corpus, a database of more than 100 million words of modern British English. The role of corpora in ESL teaching is not to 'tell us what we should teach, but they can help us make better-informed decisions, and oblige us to motivate those decisions more carefully' (Gavioli and Aston, 2001, p.239).

Comparatives

Common structures taught in most ESL classrooms are comparatives of adjectives and adverbs. Quirk et al. (1985) point out that with gradable adjectives and adverbs three types of comparisons are possible:

a) comparisons to a higher degree, as expressed by the inflected form in *-er* and *-est* or their periphrastic equivalent with *more* and *most*

e.g. *Anna is cleverer/more clever than Susan.*
Anna is the cleverest/most clever student in the class

b) comparisons to the same degree, as expressed by *as...as*:

e.g. *Anna is as tall as Bill.*
Anna is not as (so) tall as John.

c) comparisons to a lower degree, as expressed by *less* and *least*:

e.g. *This problem is less difficult than the previous one.*
This is the least difficult problem of all.

The structures above can all be found in current common ESL textbooks but b) and c) appear less often. Furthermore, whilst all students are confronted with comparative forms involving adjectives (e.g. *John is taller than Peter*), not many textbooks show that other parts of speech can also be compared. Celce-Murcia et al. (1999) argue that this should be made clear to students, as not many languages have such a large range of comparative constructions. The following examples are listed to exemplify these constructions:

Adverb: *Bill runs faster than Peter*

Noun: *Jack has more money than Harry. Max has fewer books than I have.*

Verb: *Paul weighs more than Alex.*

Other ways of comparing:

Apart from the adjective, verb and noun comparisons described above, several authors have identified the role of certain lexical items and connectives in realising the function of making comparisons in English. Kennedy (1996) for example argues that English provides a number of ways, both morphosyntactic and lexical.

Mitchell (1990) draws up the following taxonomy of possible semantic relationships in making comparisons¹ (Table 1) in which he combines lexical items and comparisons involving adjectives, verbs and nouns into the categories earlier described by Quirk et al. (1983).

Table 1

SUPERIORITY (Comparisons to a higher degree)
There is/are more X than Y The amount/number of X exceeds the amount/number of Y There is a larger amount/number of X than Y The amount/number of X is greater than the amount/number of Y
INFERIORITY (Comparisons to a lower degree)
There is/are less/fewer X than Y The amount/number of Y is less than the amount/number of X There is a smaller amount/number of Y than X The amount/number of Y is smaller than the amount/number of X
EQUALITY (AND DENIAL OF EQUALITY) (Comparisons to the same degree – equatives)
There isn't/aren't as much/many Y as X The amount/number of Y isn't as large as the amount/number of X There is/are as much/many Y as Z The amount/number of Y is the same as the amount/number of Z

Whilst Mitchell's taxonomy groups the lexical items into clear categories (superiority, inferiority and equality), it leaves no room for constructions that fall outside these groups. These could include lexical items, like 'compared to' (e.g. *Paul is fast compared to his brother*) or certain connectives, which also realize the function of comparing and contrasting (e.g. *Paul is fast; however his brother isn't*). Mitchell's table also does not take into consideration which structures occur most frequently in speech and writing.

No research to date has investigated which of the above-mentioned ways of making comparisons are used most frequently by native speakers in speech and writing.

¹ Mitchell's categories of superiority, inferiority and equality or denial of equality can be equated to Quirk et al.'s (1985) categories of comparisons to a higher, lower and same degree.

Explicit basis of comparison

The comparative form of adjectives and adverbs is used when we want to contrast one thing with another in order to point out some difference. For this purpose, as Leech and Svartvik (1994) point out, a subclause beginning with *than* can be added after the comparative word:

e.g. *His most recent book is more interesting **than his previous one.***

Leech and Svartvik (1994) further argue that elements of the subclause can be omitted if they repeat the information in the main clause. When adjectives are used for comparison, the basis of comparison is often left implicit, so that the addressee must infer the basis from the wider context:

e.g. *The houses on the other side of the river are more beautiful.*

Both Quirk et al. (1985) and Kennedy (1998) report that only 25% of the comparatives are accompanied by an explicit basis of comparison.

A brief look at textbooks suggests that almost without exception comparisons are taught in conjunction with the explicit basis of comparison. If the findings of Quirk et al. (1985) and Kennedy (1998) can be confirmed, this could have implications for textbook writers and ESL teachers.

Research Questions

For this study, the following research questions were investigated:

1. What are the most common structures used by native speakers when making comparisons?
2. Are most comparisons followed by an explicit basis of comparison?

Methodology

Data Source

The British National Corpus (BNC) was chosen for its size (over 100 million words) and for the fact that it represents a wide range of modern British English. It consists of a written part (90%) and a spoken part (10%). There are six and a quarter million sentence units in the whole corpus. To search the online-version of the BNC, the SARA software version 0,931 was used. Although the spoken part of the BNC is smaller than the written one, it is still of a very representable size which gives a good indication of spoken usage.

Data Collection

To search the BNC most effectively, the different constructions for making comparisons were identified. For this the following five categories were drawn up on the basis of the literature review (see table 2 below).

Table 2

Comparisons to a higher degree
Adjective + -er/-est <i>The gap between need and availability has grown even greater</i>
More/most + adjective <i>The less it happened, the more important it became</i>
Verb + more/most <i>People in sparsely populated rural areas have to pay the most for transport</i>
Comparisons to a lower degree
Less/least + adjective <i>The least expensive way to retain your freedom</i>
Verb + less/least <i>Other European car companies produce less than 30% of a car's value</i>
Less/least/fewer/fewest + noun <i>With less children, a woman can feel at peace</i>
Comparisons to the same degree
(Not) as adjective as <i>The wound was not as serious as it had at first appeared</i>
(Not) so adjective as <i>These images are so powerful as to demand an immediate response</i>
(Not) verb as much as <i>You may not get as much as you want</i>
As/so much/many noun as

<i>He has so much money as he could get</i>
Lexical Items: (examples)
Differ from/between/about/over/with <i>The wind differs in direction and speed from the true wind experienced when standing still</i>
Compare <i>Compared to the large, powerful bodies of the Orcs, Goblins look rather thin and scrawny, with gangly arms</i>
To verb like noun phrase <i>It seemed like the end of an era for licensed dealers</i>
Connectives
On the contrary <i>On the contrary, he welcomed it, befriended it</i>
In contrast <i>In contrast to much of western Europe, primogeniture was not practiced in Russia</i>
Whereas <i>Those blessed with a prophetic vision of soccer's future are often completely ignored or misunderstood, whereas charlatans who would not know one end of a football from another suck up to mediocrity and call it talent</i>

As not all the items identified occur frequently, a preliminary investigation of the corpus was carried out to identify the most common lexical items and connectives. Appendices 1 and 2 show only the items investigated in the main search. To ensure that the connectives investigated are actually used to make comparisons, a sample of 100 concordance lines was manually investigated and a percentage established. Furthermore, to compare search results of the written and spoken section of the corpus, the occurrences per million words were calculated.

Results and Discussion

What are the most common structures used by native speakers when making comparisons?

Table 3 following shows a summary of the results from the search.

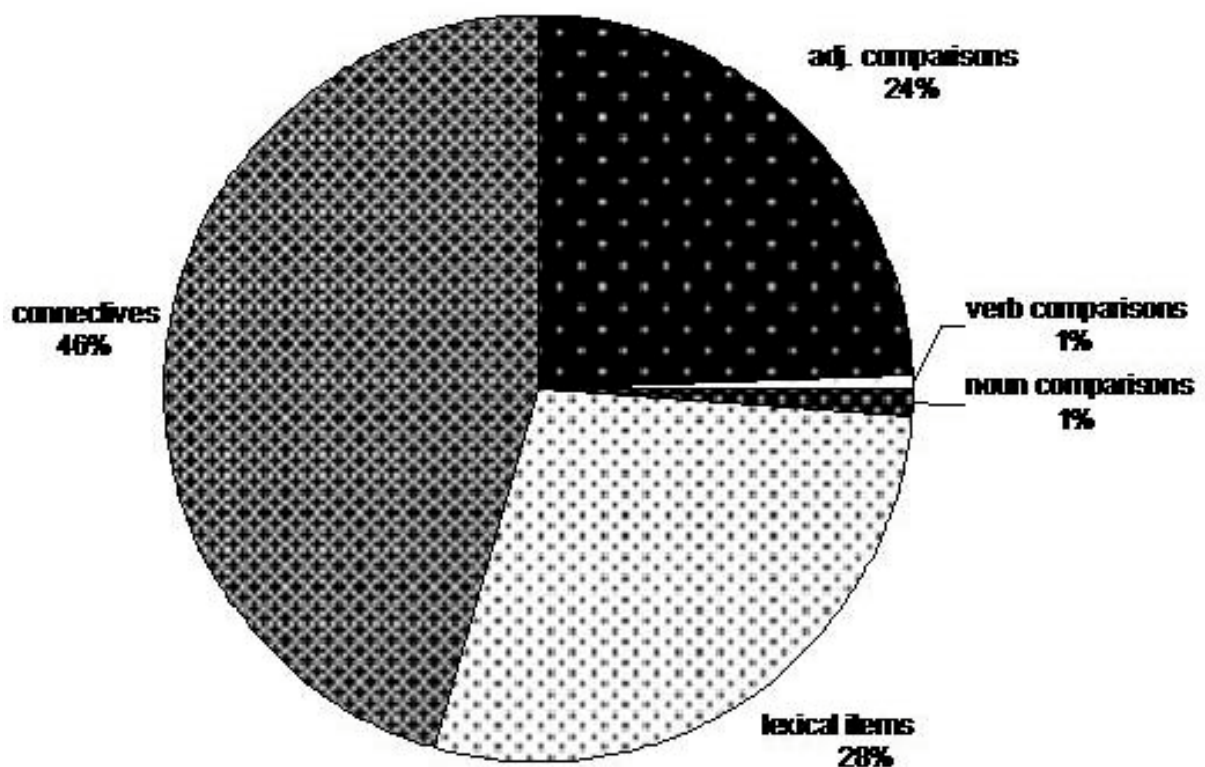
Table 3

Comparisons to a higher degree	
– adjectives	3299
- verbs	105
- nouns	135
Comparisons to a lower degree	
- adjectives	30

- verbs	10
- nouns	10
Comparisons to the same degree	
- adjectives	222
- verbs	2
- nouns	19
Lexical items	4126
Connectives	6643

Figure 1 below clearly shows that the largest category of comparisons made by the native speaker of British English involve connectives (46%), followed by lexical items (28%) and adjective comparisons (24%). Comparisons involving nouns and verbs only make up a very small part, that is 1% each.

Figure 1: Ways of making comparisons in English



This shows how important a part is played by connectives and lexical items in the function of making comparisons in English. The limited set of lexical items discussed earlier also plays an important part, which has probably been underestimated by most ESL textbooks in the

past. Closely following the lexical items is the group of adjective comparisons usually taught the most to ESL students. It can therefore be argued that on top of the cognitively relatively demanding and confusing rules of adjective comparisons, the groups of connectives and vocabulary words need to be added to help students master the function of making comparisons in their own production and understanding. It can further be contended that comparisons involving verbs and nouns do not contribute much to using comparisons productively.

A further analysis of the adjective, verb and noun comparisons in Figure 1 was carried out to discover which of the three categories drawn up by Quirk et al. (1985) (i.e. comparisons to the higher, lower and same degree) is the most commonly used by native speakers. This shows that 92% of the comparisons involving adjectives, verbs and nouns are in fact comparisons to a higher degree, whilst only 6% are comparisons to the same degree and 2% are comparisons to a lower degree.

Are most comparisons followed by an explicit basis of comparison?

This section of the study attempts to investigate what percentage of comparisons involving adjectives, adverbs, verbs and nouns are followed by the explicit basis of comparison '*than*'. Due to limitations in searching with the SARA software, only occurrences of '*than*' directly after or with a one word gap after the structure in question were included in the results. A quick scan of the concordance lines produced as out put by SARA, shows that only a very small number of comparisons have been missed due to these limitation.

In the table below (Table 4), the first number shows the occurrences of the explicit basis of comparison per million words in the whole corpus whilst the figure in brackets indicates the percentage of occurrences of the explicit basis of comparisons for each structure.

Table 4

	whole BNC	written	spoken
Comparisons to a higher degree			
More + adj + than	20.99 (13%)	21.99 (13%)	12.25 (15%)
Adj + -er + than	261.60 (12%)	262.74 (12%)	251.78 (15%)
Verb + more + than	10.63 (11%)	10.12 (10%)	15.04 (15%)
More + noun + than	5.98 (7%)	5.85 (8%)	7.14 (7%)

Comparisons to a lower degree			
Less + adj	4.24 (15%)	4.61 (15%)	0.96 (11%)
Verb + less	1.39 (20%)	1.47 (21%)	0.77 (12%)
Less + noun	0.85 (14%)	0.9 (15%)	0.39 (6%)
Fewer + noun	0.5 (16%)	0.55 (14%)	0.1 (5%)

From Table 4 above it becomes clear that only a fraction of all the structures are followed directly by an explicit basis of comparison. The percentages are very similar for all the structures, usually ranging between 7% and 20%. For comparisons to a higher degree it can be argued that generally the explicit basis of comparison can be found more commonly in spoken than in written language. This can be attributed to two factors. Firstly, it is possible that the speaker moves the explicit basis of comparison closer to the part of speech compared, in order to make it less cognitively demanding for the interlocutor to follow the message being communicated. If this is the case, then the research presented here picked these forms up more frequently, as it was not possible to record occurrences of *'than'* further away than one word from the part of speech compared. Secondly, it is also possible that there are actually more occurrences of *'than'* as an explicit basis of comparison in speech. It is possible that speakers use this structure to clarify their speech to the interlocutor.

The figures found for comparisons to a lower degree show that here the explicit basis of comparisons is more common in writing. The figures are too low though to draw any conclusions. They are a little lower than Kennedy's (1998) findings. He found that about a quarter of all instances of *'more/less + adjective/adverb/noun'* are followed by an explicit basis of comparison.

If the results of this research do in fact reflect the real native speakers usage of the explicit basis of comparison *'than'*, this can have implications for the ESL classroom. It needs to be made clear to students of all levels that native speakers only rarely signal the basis of comparisons in their speech and writing as a lack of understanding of this concept could impede the comprehension of ESL students when reading or listening to native discourse. Furthermore, ESL textbooks usually do not make it clear to students that the explicit basis of comparison does not always necessarily occur directly after the part of speech compared, or that *'than'* does not occur at all. Constructions where *'than'* only occurs much later in an utterance or sentence, or does not occur at all, are cognitively more demanding for students and should be specifically discussed in class.

Analysis of some ESL textbooks

Table 5

Textbook/Structure	adj + -er/-est	more/most + adj	irregular adj	less/least + adj	more/most + noun	fewer/less + noun	(not) as adj as	explicit basis of comparison	explicit basis of comparison	like	the same as	similar to	different	comparable to	equivalent to	enough, sufficiently, too	connectives
Headway Advanced							+			+							
New Cambridge Advanced English	+	+					+	+	+	+	+	+	+	+	+		+
Advanced Murphy	+	+					+	+								+	
New Headway Upper Intermediate																	
Reward Upper Intermediate																	
Upper Intermediate Matters																	
Objective First Certificate	+	+					+	+									
Landmark Upper Intermediate																	+
Cutting Edge Intermediate	+	+		+		+	+	+		+	+						
Reward Intermediate	+	+		+	+	+	+	+									
Essential Grammar in Use Intermediate	+	+	+	+		+	+	+			+	+					
New Headway Pre Intermediate	+	+	+					+									
Changes 2	+	+						+									
Reward Elementary	+	+	+					+									
Essential Grammar in Use Elementary	+	+						+									
New Headway Beginner																	

An analysis of some current ESL textbooks clearly shows that most textbooks concentrate on presenting adjective comparisons directly followed by the explicit basis of comparison 'than'. The minority also teach connectives or lexical items in the context of making comparisons. Surprisingly, quite a number of textbooks do not look at the comparisons at all.

Implications for ESL textbook writers

It can be argued that textbooks greatly influence the way certain structures and functions are presented to students in class. The present study shows up several implications for textbook writers, as it identifies the range and frequency of structures used for comparing and contrasting by the native speaker.

It was found that connectives and lexical items are more common than the structures of adjective comparisons traditionally taught in textbooks. Course book writers could introduce structures to students in order of frequency, so that highly frequent structures are taught at lower levels, whereas infrequent structures are only introduced at higher levels.

Secondly, these findings suggest that students should be exposed to and made aware of the fact that explicit basis of comparison is not always used by native speakers. Knowledge of this could both aid comprehension and production.

Generally, it seems more sensible to teach comparisons rather as a function and less as a grammar point to allow room for vocabulary words like '*compared to*' or '*to differ from*' and connectives like '*yet*' or '*however*' to be taught in the context.

Recommendations for ESL teachers

This study has shown that most textbooks do not present the full range of structures used by native speakers to compare or contrast. Many teachers of English might not themselves be aware of this, or might have limited resources to supplement the textbooks used in their teaching contexts. The frequency information collected for this study provides teachers with a reliable source of information about how comparisons are realised in native language use.

Most recommendations made to textbook writers also apply for teachers in the ESL/EFL classroom. Teachers should be well-informed about the findings of this study so as to evaluate the textbook they use. As Celce-Murcia et al. (1999) suggest, teachers should be recommended to introduce making comparisons as a function covering a range of structures, rather than simply as a grammar point covering comparative and superlative adjectives.

Merely teaching the rules of forming adjective comparisons does not do justice to the function of comparing, as can be seen in Figures 1 in the results section. Teachers should further encourage students to use the whole range of structures identified in this study. They could possibly accomplish this through awareness-raising techniques, like collecting a small learner corpus of the students' own output (see Seidlhofer, 2002) and comparing the data with native speaker performance of a similar task or making corpora or concordances available to learners as suggested by Gavioli and Aston (2001) and Meunier (2002).

About the Author

Ute Knoch is a research assistant and PhD candidate at the University of Auckland. Her research interests are corpus linguistics, learner corpus linguistics and language testing.

Email: u.knoch@auckland.ac.nz

References

- Bell, J. and Gower, R. (1992). *Upper-Intermediate Matters*. Students' Book. Essex: Longman.
- Capel, A. and Sharp, W. (2000). *Objective First Certificate*. Students' Book. Cambridge: Cambridge University Press.
- Celce-Murcia, M. and Larsen-Freeman, D. (1999). *The Grammar Book. An ESL/EFL Teacher's Course*. Boston: Heinle & Heinle.
- Cunningham, S. and Moor, P. (1999). *Cutting Edge Intermediate*. Students' Book. Essex: Pearson Education Limited.
- Gavioli, L. and Aston, G. (2001). 'Enriching reality: language corpora in language pedagogy'. *ELT Journal* 55, 238-246.
- Greenall, S. (1997). *Reward Elementary*. Students' Book. Oxford: Macmillan Heinemann.
- Greenall, S. (1995). *Reward Intermediate*. Students' Book. Oxford: Macmillan Heinemann.
- Greenall, S. (1996). *Reward Upper-Intermediate*. Students' Book. Oxford: Heinemann.
- Haines, S. and Steward, B. (2000). *Landmark. Upper-Intermediate*. Students' Book. Oxford: Oxford University Press.
- Hewings, M. (1999). *Advanced Grammar in Use*. Cambridge: Cambridge University Press.

- Jones, L. (1991). *New Cambridge Advanced English*. Cambridge: Cambridge University Press.
- Kennedy, G. (1996). 'The Corpus as a Research Domain' in S. Greenbaum. (ed.) *Comparing English Worldwide*. Oxford: Clarendon Press.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman.
- Leech, G. and Svartvik, J..1994. *A Communicative Grammar of English*. London: Longman.
- Meunier, F. (2002). 'The pedagogical value of native and learner corpora in EFL grammar teaching' in S. Granger, J. Hung and S. Petch-Tyson (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Mitchell, K.(1990). 'On Comparisons in a Notional Grammar'. *Applied Linguistics 11*, 52-72.
- Murphy, R. (1990). *Essential Grammar in Use. Elementary*. Cambridge: Cambridge University Press.
- Murphy, R. (1985). *English Grammar in Use. Intermediate*. Cambridge: Cambridge University Press.
- Quirk, R., Greenbaum, S., Leech, G. and Svartvik, J.(1985). *A Comprehensive Grammar of the English Language*. New York: Longman.
- Richards, J. C. (1995). *Changes 2. English for International Communication. Students' Book*. Cambridge: Cambridge University Press.
- Seidlhofer, B. (2002). 'Pedagogy and local learner corpora: working with learning-driven data' in S. Granger, J. Hung and S. Petch-Tyson (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Soars, J. and Soars, L. (2002). *New Headway English Course Beginners. Students' Book*. Oxford: Oxford University Press.
- Soars, J., Soars, L. and Sayer, M. (2000). *New Headway English Course Pre-Intermediate. Students' Book*. Oxford: Oxford University Press.
- Soars, J. and Soars, L. (1998). *New Headway English Course Upper-Intermediate. Students' Book*. Oxford: Oxford University Press.
- Soars, J. and Soars, L. (1989). *Headway Advanced. Student's Book*. Oxford: Oxford University Press.

Appendix 1

The following table shows the lexical items investigated in this study. A pre-investigation of the corpus showed that other items identified for this category occurred too infrequently and were therefore not further investigated.

	whole BNC	written	spoken
senior to	24	23	1
junior to	23	23	0
superior	2169	2128	41
inferior	842	811	31
prior	4765	4578	187
major	23371	22332	1039
minor	4456	4219	237
anterior	201	201	0
posterior	152	152	0
too	26800	24198	2602
enough	23286	20797	2489
sufficiently	2506	2440	66
to differ from/between/about/over/with	4147	4044	103
difference	11164	9759	1405
different	47607	42442	5165
differently	1518	1374	144
to v like np	114329	87200	27129
similar	18295	17575	720
equal	7455	6574	881
equivalent	2130	2058	72
identical	2148	2099	49
to be alike	180	167	13
to be the same as np	60274	53537	6737
to be unlike np	4553	4472	81
exceeds	3046	2964	82
diminishes	1800	1752	48
increase/s	31712	30401	1311
compared	12866	12284	582
in comparison to	152	142	10
comparison with	1054	1016	38

Appendix 2

The following table shows the connectives investigated for the purpose of this study.

	whole BNC	written	spoken
alternatively	1729	1698	31
on the other hand	5311	5064	247
contrariwise	15	15	0
instead	7262	7031	231
conversely	812	798	14
oppositely	10	10	0
on the contrary	799	786	13
in contrast	2238	2233	5
by contrast	1433	1420	13
by comparison	535	517	18
but	443164	377293	65871
though	9618	5366	4252
however	54673	54011	662
whereas	6169	5508	661
on the other side	15	15	0
by way of comparison	6	5	1
by way of contrast	37	37	0
contrastingly	4	4	0
yet	32515	29278	3237
nevertheless	7045	6776	269
despite that	92	79	13
in spite of that	25	23	2
although	42758	41101	1657
while	48804	47195	1609