English adjectives of similarity: a corpus study of their determiner/modifier status

Abstract. This paper investigates the determiner/modifier distinction issue. As reference grammars tend to define determiners and modifiers very vaguely especially in neighbouring cases there is a need for more formal rule together with obligatory semantic constraint. The left-hand periphery of English noun phrases is of strict order: determination zone precedes the modification zone (from left to right), so determiners go before the adjectival string. As the determination zone precedes the modification one it is assumed that only modifiers can be preceded by adjectives. The adjectives of similarity are tested in this way. The corpus-based analysis results in fixing possible adjectives accompanying the adjectives of similarity on their left. It is claimed that no adjective of similarity can be considered a determiner.

1 Introduction

This paper is concerned with adjectives of similarity and their functioning in English noun phrases. Adjectives of similarity alongside adjectives of identity and difference belong to adjectives of comparison which were discussed as a group by M. Halliday and R. Hasan (1976) [1], T. Breban (2011) [2], P. Matthews (2014) [3], K. Davidse (2022) [4]. All the authors posed a crucial issue of differentiating determiner/modifier status for these adjectives.

Determiners and modifiers are integral parts of English noun phrases. Noun phrase (NP) consists of a head and dependents. The dependents are determiners and modifiers, pre-head and post-head modifiers. This paper is limited to pre-head elements.

Pre-head modifiers consist of chiefly adjectives, non-finite clauses, nouns. In this paper only adjectives are of interest. Determiners are seen as functional elements limiting the potential referent of a noun phrase [5, 6, 7]. Determiners are divided [e.g. 6] into predeterminers, central determiners and postdeterminers due to their mutual layout in NP. The usual list of them includes: for central determiners - definite and indefinite articles, possessives, demonstratives, quantifiers some, any, no, every, each, either, neither, certain, enough, the relative determiners whose, which, the wh-determiners in – ever, the interrogative determiners what, which, whose; for predeterminers include universal quantifiers all, both, half; fractions; the multipliers double, twice, three times, etc; such, what. The following classes have been associated with the postdeterminer function: same; ordinal numerals; general ordinals next, last, past, (an)other, further; cardinals; quantifiers many, few, little, several. Other researches add to the above list: numerous, etc. [8, 9]; different, identical; complete, entire, whole; above, aforementioned; customary, expected, famous, given, habitual, necessary, normal, notorious, obvious, odd, ordinary, or original, particular, possible, probable, regular, respective, special, ...
categorization: one first identifies an object, singles it out, in other words, “binds” it the communicative situation, “grounds” it and only after that describes it and gives some more

classification.

2 Materials and methods

I will analyze synchronic data extracted from the BNC, COCA, NOW corpora for the three adjectives: similar, comparable and analogous. They describe the likeness, resemblance or correspondence between two or more things. All these three words can be used both attributively and predicatively, but only attributive usage is exploited in this paper.

I go for The British National Corpus as it is the most conservative corpus, the American Corpus of Contemporary English is chosen for its world-wide nature. The News on the Web corpus is the most diverse and extensive corpus. It represents so-called global English and it is useful precisely because of its diversity and the possibility to trace emerging trends and tendencies. I do not compare the corpora, there is no preference, and all the data have the same value and add to the results.

In order to investigate more or less established varieties I selected only that part of the overall data originated from the written portions of the corpora. This was to avoid the risk that the analysis would be affected by the possible slips, mistakes or inaccuracies usual for spoken speech. The written speech is preferable here, as it needs some time for thinking over, so it presents a kind of conscious choice. The inquiry was the possible existence of adjectives before “similar, comparable or analogous” premodifying a noun. The raw data were cleared up so the total extraction from the corpora yielded 4589 tokens. I studied the examples using quantitative and qualitative methods, as it is a regular procedure for the corpus-based studies: counting, semantic decomposition, descriptive method, interpretative analysis, semantic, syntactic, contextual analyses.

3 Results and discussion
Cambridge dictionary's definitions for these words run as follows: similar-looking or being almost, but not exactly, the same; comparable-similar in size, amount, or quality to something else; analogous-having similar features to another thing and therefore able to be compared with it.

BNC, COCA and NOW give the following numbers of attribute use: for similar – 692, comparable – 3897, analogous – 99. It seems reasonable to start the description of findings common for all these three words.

There are the following semantic blocks of adjectives accompanying these words on their left. Words, expressing row, order of appearance or succession: previous, preceding, recent, latest, past, prior, earlier, nearest, near, neighboring, adjacent, closest, actual, current, living, existing, future, on-going, up-coming, subsequent, final.

1. House sales over the six months went up from 357 units for the previous comparable period to 409. (BNC)
2. To be conscious of any change in experience there must be a comparison with previous similar situations (COCA).
3. Images made by functional magnetic resonance imaging technology show that remembering and imagining send blood to identical parts of the brain. What does this say about the goal of living in the present? But for most of us, the phenomena of the present connect in our mind with previous analogous experiences. Recognition involves memory: comparing what is seen with what was seen (COCA).
4. Numerous similar studies prove infection prevention is cost effective, but hospital administrators still claim they can’t afford it (COCA).
5. In April alone, DNS reported 58 homeless deaths from COVID-19, 1.6 times higher than the overall city rate. While there is no reliable analogous data for other cities, what happens in New York can be a lesson for other (NOW).
6. And yet, is there a single comparable counter-example from any time in history of a similar economic collapse and widespread hunger under any form of extreme free market capitalism? (NOW)
7. Texas ends its driest year on record, with an average of only 14.8 inches of rain. The only comparable drought occurred during the drought of record during the 1950s (COCA).

The antecedents for these phrases are nearby, in the same or neighboring sentences. Thus, the antecedent of “only comparable drought” is in the same sentence – “the drought of record during the 1950s”, “numerous similar studies” has an antecedent expressed verbally in exact numbers. So, phoric nature of these phrases is observed.
“Similar” and “analogous” are not as rich in left-hand adjectives as “comparable.” Above-mentioned examples form the exhaustive list of such adjectives for “similar” and “analogous.” “Comparable” presents a greater variety of adjectives most of which describe “sales, revenue, data, numbers, financials, statistics, etc.” “comparable” and “entire, full.”

The company’s annual comparable sales... solid comparable sales... strong comparable sales... “Comparable” presents a single, inseparable unit, a kind of term (as it is the case with “comparable sales” and “similar” and “analogous”). “Comparable” means that all elements of this network are engaged with no part left out.

11. Each and every Google product has at least one excellent competitor I could switch to (except for YouTube), but there’s no rival company that has an entire comparable ecosystem.

12. ... the first year for which full comparable data... two full comparable periods... the entire comparable periods... “Full comparable data” and “full/entire comparable periods” show that there are no omissions from either a dataset or given timespans.

13. They did not generate comparable revenues for two full comparable periods.

14. They were not in operation during the entire comparable periods a year ago.

“Full comparable data” and “full/entire comparable periods” show that there are no omissions from either a dataset or given timespans. The adjectives emphasize the idea of maximal capacity here.

15. Dollars based upon the foreign currency exchange rates for the applicable comparable period.
On average, South African services industries are better attuned to the needs of women. The sectors have a high percentage of women employees. Mining and other heavy industries lag behind in terms of gender pay equity. Salaries in government are, on average, better for both men and women than similar comparable jobs in the private sector (COCA).

The Platinum Card offers several perks that are unmatched by similar comparable cards, such as the Chase Sapphire Reserve Card (COCA).

I led four national surveys of Canadian workers. First, I established pre-pandemic data points with a survey in September 2019. Then, throughout the pandemic, I fielded similar comparable surveys in September 2020, 2021 and 2022 to track trends—roughly 13,500 study participants in total (COCA).

After all, that is what the public employee system does: firefighters here want the same salary as firefighters in Toronto, police want similar comparable wages with Toronto or the OPP or the RCMP or whoever currently pays the highest. Tax clerks want the same wages as their counterparts in Barrie or Hamilton; a municipal truck driver is a truck driver; a custodian is a custodian, a City Solicitor is a City Solicitor or so goes the song (COCA).

In all these examples “comparable” transfers the generalized likeliness meaning, thus speakers feel the need to describe these “comparable N” units as “similar”. Or, it can also be assumed that this tautology reflects the speakers’ need for emphasizing the similarity. The word “similar” acts as an intensifier in this case. The innate reasons for such double similarity remain unknown. Again, it can also be a mistake but the sources (written newspaper articles, reviews, reports) cast doubt on such a suggestion.

4 Conclusion

Several blocks of accompanying adjectives are distinguished. The most numerous one is the group of adjectives expressing row, order of appearance or succession: previous, preceding, recent, latest, past, prior, earlier, nearest, near, neighboring, adjacent, closest, actual, current, living, existing, future, on-going, up-coming, subsequent, final.

Other groups consist of adjectives expressing indefinite quantity (various, numerous, multiple, countless, sundry, innumerable); relevance (relevant, unnecessary, worthy, basic, additional; appropriate); size (big, little, large, great); frequency (frequent, rare, perpetual); complexity (simple, complex), availability (available, possible, potential); positive opinion (good, cool, interesting, significant, encouraging, effective, excellent, impactful, positive, reliable, fascinating); level of identity (common, different); exclusiveness (only, single, solitary). The first block of adjectives and adjectives of indefinite quantity (various, numerous, multiple, countless, sundry, innumerable) are considered to be postdeterminers (as mentioned earlier in the introduction). This seems reasonable, they can serve as synonyms to corresponding determiners but they still have distinct lexical meaning as opposed to the grammaticalized determiners. I assume that in this case, the two zones (determination zone and modification zone) mix up, and their limits get fuzzy. Even if these adjectives can be considered postdeterminers and not modifiers, there are still other adjectives of not so fuzzy status (namely, the above mentioned adjectives of size, frequency, relevance, availability, positive).
All these variants are common for the three given adjectives of similarity. The adjective “comparable” stands out of the row with a greater variety of adjectives accompanying it. The data show many examples of left-hand adjectives describing location, source, period, size and intensity of the notions “comparable sales/revenue/data/numbers/financials/statistics/figures”, the idea of ability to be used in some way: applicable, suitable; the actions applied to corresponding entities: published, predicted, listed, indexed. Besides, “comparable” is accompanied by “similar” as in “similar comparable deals/yield, etc.

Taking into account the determiner/modification zone distinction and the vast range of adjectives accompanying these adjectives of comparison, I assume “similar, comparable and analogous” are not the participles of the determination zone. They seem to act as modifiers, not determiners.

References

1. M. Halliday, R. Hasan, Cohesion in English (1976)
23. D. A. Axelrood, M.S. Brodskaya, Nauchnyi dialog, 1, 9 (2020)
24. O. V. Dolmatova, Nauchnyi dialog, 10, 110 (2019)