DOI: 10.1515/RJES-2019-0010

A CORPUS VIEW ON A COUPLE OF STRUCTURALLY AMBIGUOUS SENTENCES

NADINA CEHAN

Independent researcher

Abstract: *Ambiguity has played a central role in the formation of modern grammatical theory. The article revisits the topic of structural ambiguity in particular, by looking closely at the sentences ‘The shooting of the hunters is terrible’ and ‘Flying planes can be dangerous’, both taken from Chomsky’s works. The question posed is whether the data provided by corpora verify their ambiguity.*

Keywords: *ambiguity, corpus linguistics, lexical ambiguity, pragmatic ambiguity, structural ambiguity*

1. Introduction

(1) The shooting of the hunters is terrible. (Chomsky 2002: 88)

(2) Flying planes can be dangerous. (Chomsky 1965: 21)

These sentences embody structural ambiguity (sometimes referred to as syntactic or grammatical ambiguity, or amphiboly), a feature of the language that structural linguistics could not deal with (cf. Searle 1972). Structural ambiguity has supported both the surface / deep structure and the competence / performance distinctions that Chomsky introduced. “There is little question that ambiguity has played a central part in the formation of modern grammatical theory, though more as a catalyst than a direct field of enquiry” (Kess and Hoppe 1981: 13; also cf. Zwicky and Sadock 1975). The following discussion revisits the topic of structural ambiguity and closely inspects sentences (1) and (2) above by means of inquiries into available corpora.

2. Ambiguity: an Overview

Grice’s maxim (1989:27) about ambiguous discourse avoidance is frequently invoked alongside the expression of surprise at the existence of much ambiguous written and spoken language. Nevertheless, in many cases ambiguous sentences are presented in isolation. Thus, if one were to decontextualize the words written before here, one might observe that *sentences* can mean ‘judgements’ or ‘clauses’. Therefore, the second sentence in this paragraph could be argued to be ambiguous: is it judgements or clauses that are presented in isolation? If such a measure of ambiguity were applied all the time, one would surely reach the conclusion that ambiguity is everywhere. However, listeners and readers are usually primed for only one interpretation by the context in which they find themselves. Nevertheless, ambiguity is arguably more frequent and certainly more visible in ludic language, in newspaper headlines, in advertisements and on signs, where the intended (or not) word-play is part of a smaller, self-contained context. One could then contend that it is the added degree of isolation that allows for a potential ambiguity to be noticed in these cases.

Three basic types of ambiguity can be distinguished: lexical, structural and pragmatic. Lexical ambiguity is linked to the fact that words tend to have multiple meanings and that their semantic richness might make the text ambiguous. For instance, in

(3) I am waiting by the bank.

it is unclear whether one is waiting next to a financial institution or near a river.

Structural ambiguity allows a sentence or phrase to be syntactically or grammatically analyzed in at least two divergent ways. Sources of structural ambiguity, i.e. the grammatical

patterns that can lead to ambiguity, are discussed by Stageberg (1958, 1966, 1968a, 1968b) and Oaks (2010). A number of structurally ambiguous written sentences can be resolved in speech due to the placement of pitch and the intonational contour of the sentence (cf. Stageberg 1966, for a large number of examples). A sub-category of structural ambiguity is that involving scope, which often stems from the perceived divergent arrangements of the elements in the sentence. The example

(4) old men and women (Chomsky 2002:87)

is unclear as to whether it is intended to mean [*old men*] and women or old [*men and women*]. Ambiguity can also arise when considering the scope of coordinators:

(5) Susan and Matilda talked. (Zwicky and Sadock 1975:15)

One can view the sentence as *Susan [and Matilda] talked*, in which case one would understand that the two women talked one after the other, or as [*Susan and Matilda*] talked, in which case it is understood that the women talked together.

Pragmatic ambiguity is linked to speech acts. A sentence such as

(6) It's cold in here.

can be interpreted as either a statement, or, in the right context, as a request, or it can function as an expression of discontent. Thus, although the ambiguity is not signalled by its content, a pragmatically ambiguous sentence can be interpreted as at least two different speech acts.

It is assumed that it is odd for ambiguity to exist in language, since it may lead to less efficient communication. Ambiguity is thought to pose an additional demand on the listener's or reader's working memories. The question that arises is why ambiguity exists in language in the first place, since its existence involves more effort being exerted. Zipf (1949) offered a theory according to which the interaction of the different speaker's and listener's, writer's and reader's needs breeds ambiguity. In his view, language has an internal economy which reflects both the speaker and the listener taking the path of least effort in speaking and processing language (cf. *Chapter two: On the economy of words*, Zipf 1949: 19-55). Ambiguity therefore seems to be inherent to language.

3. Methods Used to Analyze the Data and Results

The two sentences under analysis, (1) and (2), are said to be ambiguous. The question posed is to what extent language data provided by corpora verifies their ambiguity. Chomsky (2002: 88) gives two readings for (1), according to which *the hunters* can be either the subject/agent or the object/affected of the sentence. In other words, in the first case, the hunters do the shooting, while in the second, they are the victims of the shooting. Accordingly, for (2) there are also two interpretations 'corresponding to "flying planes are dangerous", where planes is the subject, and "flying planes is dangerous"', where planes is the object (Chomsky 1965:21).

In order to verify whether there are indeed two possible readings for (1) and (2), inquiries were made in the British National Corpus (BNC), in the Corpus of Contemporary American English (COCA) and in the Corpus of Canadian English (Strathy). The three corpora were chosen in order to ensure language diversity and representativeness. Together, BNC, COCA and Strathy amass 670 million words. The corpora were accessed via corpus.byu.edu.

For (1), the inquiries took the form *the shooting of **, where * stands for whatever element follows *the shooting of*. The type, meaning and function of the word that followed was of interest. *The shooting of the hunters is terrible* as a whole sentence appeared only once, in the BNC, and the result was discarded because it was part of a book on linguistics. There were also three titles of literary works which were not taken into account. Other results which were ignored involved the noun *shooting* in combination with time expressions, such as *the shooting of last night* and *the shooting of a few days earlier*.

For (2), the first part of the search meant to discover whether *flying* in conjunction with *planes* could sometimes be an adjective and at other times a gerund participle. In other words, *flying planes* could either be a noun phrase with *planes* as the head noun, denoting objects that are

in the air, or it could be a verb phrase, meaning the activity ‘piloting planes’. Since the second part of the sentence might influence the interpretation of the first part, *can be dangerous* was also looked into. The question was what kind of words and phrases can come before *can be dangerous* and how frequently it combines with noun phrases versus gerund participial clauses. The inquiry took the form **can be dangerous*. As with *The shooting of the hunters*, the sentence *Flying planes can be dangerous* was found as such in the BNC and in COCA in linguistics texts and the two instances were not taken into account.

It should be noted that each returned result was inspected in its wider context so as to clearly understand the function of the words that were of interest. In addition, COCA turned out to be a tad repetitive because its content relies heavily on material from the mass media. For instance, the story of a young boy’s shooting became a national sensation and it was masticated by several media channels. This repetitiveness was overlooked, as the framing of the several shooting stories was somehow different each time.

3.1. *The Shooting of the Hunters*

The shooting of the hunters is terrible is said to have two interpretations: ‘The fact that the hunters were shot is terrible’ and ‘The hunters are terrible at shooting’. The argument put forward here is that, should (1) be ambiguous, then this ambiguity will be found in naturally occurring language. Specifically, one would expect after *the shooting of* both objects and subjects, both victims and shooters. Table 1 below shows the findings in the BNC, COCA and Strathy.

<i>the shooting of</i> *	human victims	filming or photo-graphing	birds or animals	objects shot	what is used to shoot	sport context	TOTAL
BNC	44	9	3	1	0	1	58
COCA	293	33	13	6	6	0	351
STRATHY	8	1	2	0	0	0	11
TOTAL	345	43	18	7	6	1	420

Table 1. Corpus results for *the shooting of* *

In answer to the main research question, of whether (1) is ambiguous or not, the answer provided by the corpus inquiries is that there is little evidence that *the shooting of* with the meaning of ‘to kill’ or ‘to injure’ can be followed by a subject. In 82 per cent of the cases, humans are the victims of shootings, as well as birds and animals:

(7) Stephen Shepkey, an art restorer and professed gun dealer, was the man who helped organize *the shooting of* David Wilson. (BNC K6F S_brdcast_news)

(8) The hunters defied European and French laws forbidding *the shooting of* the tourterelles on their annual nesting flight from the Sahara. (BNC J2W W_misc)

In addition, objects such as apples, effigies, a Malaysian Airline flight and other planes can be shot at:

(9) He had also supervised the making of the effigy that so resembled Kennedy. With *the shooting of* that effigy and the victory banquet afterward, David Jatney experienced a revulsion for his student life. (COCA 1990 FIC Bk:FourthK)

(10) Meanwhile, we’re still vulnerable to other, equally viable means of terrorism such as anthrax attacks on our water systems and *the shooting of* planes out of the sky from the ground. (COCA 2004 NEWS SanFranChron)

The shooting of can also be followed by a word that specifies the instrument used for shooting, namely arrows, guns, a revolver:

(11) One day shortly before I left Pirien, a man named Marco was acting out a story, walking and stalking and mimicking the stabbing of someone with a spear, *the shooting of* arrows, the cutting off of a head. (COCA 2014 MAG Smithsonian)

The second meaning of *shoot* in the phrase *the shooting of* is that of filming or taking photos:

(12) Lewis also had to contend with tabloid reports about an off-screen romance between her boyfriend and Geena Davis during *the shooting of* *Thelma & Louise*. (BNC CGB W_pop_lore)

The shooting of also appears in contexts dealing with sport events, since *to shoot* can also mean ‘to try to score’. It is the only context in which what follows is the subject, the doer/agent of the shooting:

(13) Surrey, with *the shooting of* Barbara Taylor and Sheila Edwards outstanding had started the last quarter 44-38 in front but Essex Met fought back to take a three-point lead. (BNC AHUW newsp_brdsht_nat_misc)

The corpus results suggest that the *the shooting of* displays only lexical ambiguity, stemming from the difference in the shooting instruments used: shooting with a weapon or shooting with a camera. It is then possible to construct an ambiguous sentence such as:

(14) The shooting of Marc was a horrible experience.

where *Marc* could be a human victim or the title of a movie, an affected or an effected. Stretching one’s imagination a bit, one could also interpret the sentence as meaning that Marc’s kicking the ball was a horrible experience.

Since *to shoot* also has a usage in sports, the corpus results suggest that the two interpretations of (1) could be ‘It is terrible that the hunters were victims of a shooting’, ‘The way the hunters were filmed is terrible’ and ‘The way the hunters shoot the ball is terrible’, with *shoot* having three completely divergent meanings. It is in the third interpretation that it is possible to parse (1) differently. One is dealing, therefore, with a lexical ambiguity which can lead to structural ambiguity.

3.2. *Flying Planes Can Be Dangerous*

(2) was analyzed by cutting it into two parts: *flying planes*, on the one hand, and *can be dangerous*, on the other. The interaction between the two parts is important in that they may conform to patterns which are harmonious, or to patterns that are conflicting. Specifically, supposing that *flying planes* is a common noun phrase and *can be dangerous* is generally preceded by a noun phrase, then this would suggest that *Flying planes can be dangerous* makes for a harmonious combination of patterns. However, if in *flying planes*, *flying* is usually a gerund participle, but *can be dangerous* is still routinely preceded by a noun phrase, then one may argue that there is a clash between these patterns.

The first question, then, is related to the frequency with which *flying*, in combination with *planes*, is either a gerund participle or a participial adjective. Thus, all the concordances in the BNC, COCA and Strathy containing *flying planes* were analyzed in their extended contexts, with the results shown below:

Flying	
<i>gerund participle</i> (denoting an activity carried out by somebody)	<i>participial adjective</i> (denoting an activity carried out by planes)
34	2

Table 2. The total frequency of *flying* as gerund participle and as participial adjective in collocation with *planes* in the BNC, COCA and Strathy

Generally, *flying* in *flying planes* is a gerund participle, as in the following examples:

(15) How did you go from *flying planes* to designing zip-line components? (COCA 2015 MAG PopScience)

(16) In this film the women were fixing planes, *flying planes*, and shooting other people’s planes out of the sky. (COCA 2011 FIC Bk:NudeWalker)

The only two sentences in which *flying planes* was a noun phrase with *planes* as noun and *flying* as a participial adjective were the following:

(17) Commuter carriers, defined for federal purposes as those *flying planes* with 30 seats or fewer, have averaged 6.6 accidents. (COCA 1994 MAG USNR)

(18) He did not watch and become exhilarated by the low *flying planes* or the rumbling earth. (COCA 1993 FIC Iris)

The very low co-occurrence of *planes* with *flying* as an adjective may be explained by the fact that a *flying plane* is more readily described as a *plane in the air* or *in the sky*. In addition, it is usually expected that a plane does not fly by itself, but is piloted. The word *plane* meaning aircraft does

not readily collocate with *flying* as a participial adjective. *Flying saucers, flying colours, flying machines* are among *flying*'s most frequent collocates.

Can be dangerous is analyzed in terms of what precedes it. Of particular interest here is the number of noun phrases in comparison with the number of gerund participial clauses that collocate with *can be dangerous*. The results are shown in Table 3 below:

* <i>can be dangerous</i>	noun phrase	gerund participial clause
BNC	35	10
COCA	215	83
STRATHY	20	3
TOTAL	270	96

Table 3. The number of concordances showing noun phrases and gerund participial clauses or phrases preceding *can be dangerous* in the BNC, COCA and Strathy

There are around three times as many noun phrases preceding *can be dangerous* than there are gerund participial clauses. Most noun phrases encountered do not have the structure of *visiting parents*, but are either simpler, as in (19), or much more complex, as in (20):

(19) Shell-fish *can be dangerous* in their own right. (BNC A0D W_fict_prose)

(20) Home-made enguli and similar home-made drinks in other East African countries *can be dangerous* since the second fraction is often contaminated with the toxic first and third fractions. (BNC HSD W_non_ac_nat_science)

Similarly, gerund participials can vary from one word, as in (21), to complex phrases, as in (22):

(21) Now, modeling *can be dangerous*. (COCA 2015 SPOK NPR)

(22) Talking to people who don't understand you *can be dangerous* and dull as ditchwater. (COCA 2015 FIC MassachRev)

Some of the returned results were very similar in form to sentence (2) and could be analysed as participial adjective plus noun preceding *can be dangerous*:

(23) *Sweeping generalizations can be dangerous*. (COCA 1992 NEWS USAToday)

and gerund participle plus its object preceding *can be dangerous*, as in:

(24) *Stopping piracy can be dangerous*. (COCA 1993 NEWS USAToday)

By comparing Table 2 with Table 3 one may conclude that there appears to be a conflict between *flying planes* and *can be dangerous*. Most of the times, *flying* in *flying planes* is a gerund participle with *planes* as its object, while *can be dangerous* is mostly preceded by noun phrases. This is what could be called a clash in expectations.

3.3. Discussion

The corpus data for *The shooting of the hunters is terrible* have indicated that this sentence can be ambiguous, but not in the Chomskyan way. Unless used with the very infrequent sportive meaning of 'shooting to score', *the shooting of* is usually followed by a theme or patient, not an agent. Nevertheless, as agents, we expect hunters to shoot their guns, not play football. In other words, people's world knowledge impacts on sentence interpretation. Moreover, as Oaks (2010:17) explains, "the preposition *of* can relate to agentive (roughly corresponding to subject) as well as patient (roughly corresponding to direct object) roles – semantic relationships that are syntactically different when they are paraphrased". It is then the presence of the preposition *of* that makes it unclear whether the hunters are victims, shooters or players.

In the case of *Flying planes can be dangerous*, one may first notice the use of the modal verb *can*, whose lack of inflection for number engenders the ambiguity. "This pattern of a gerundive phrase with a modal in [a subject-verb-complement] structure [...] is commonly seen as a linguistic example of structural ambiguity" (Oaks 2010:85). In addition, the corpus analysis indicates that there may be a clash in hearer / reader expectations due to the most frequent collocates of the two parts of the sentence. It may be assumed that the frequency effects which tug and pull – *flying planes* being usually an *-ing* phrase, while *can be dangerous* being normally preceded by a noun phrase – render the sentence visibly ambiguous. Yet the forms that the noun phrases take range from one word to highly complex phrases, which make it difficult if not

impossible to distinguish patterns in the noun phrases returned by the corpus inquiry. At the same time, phrases similar to *flying planes* where the *-ing* word is a participial adjective are very rare (see example 24). On the assumption that frequency influences understanding, this suggests that the meaning of the sentence will be given by the way *flying planes* is most frequently analyzed, namely as gerund participle plus object.

The remarks above lead to a possible prediction: should the ambiguity of the two sentences be tested in a different manner, they will not necessarily be found ambiguous. For instance, if native speakers were asked to rephrase in their own words *The shooting of the hunters is terrible* and *Flying planes can be dangerous*, they would not necessarily discover two structurally divergent reinterpretations for each sentence. Given the very high frequency of *the shooting of* being followed by human victims, it is likely that *the hunters* will be interpreted as victims and subjects will not imagine that there are other ways to understand the sentence. Likewise, the very low incidence of *flying planes* as participial adjective plus noun is likely to lead to a unique interpretation of *flying* as *piloting*.

4. Conclusion

Structural ambiguity, where a sentence can be parsed in two divergent ways, has supported the rise of modern grammatical theory. The sentences on which this article has focused were proposed decades ago by Chomsky to sustain his arguments. Sentence ambiguity is best observed out of context. The increased degree of language isolation in certain text types, such as news headlines, favours it being noticed. Otherwise, potential ambiguity is very likely overlooked by listeners and readers, despite its ubiquity.

The present article has taken a fresh look at the famous structurally ambiguous sentences *The shooting of the hunters is terrible* and *Flying planes can be dangerous* by conducting a corpus investigation on them. The elements with which the different parts of sentences collocate have come into focus. For *The shooting of the hunters* research has indicated that an agentive interpretation for *the hunters* is not possible unless they are doing sport. Nevertheless, it is suggested that people's knowledge that hunters must have some kind of weapon and the polysemous nature of *of* may lead towards the two interpretations.

For *Flying planes can be dangerous*, the ambiguity builds on the gerundial or adjectival nature of *flying*, as well as on the use of the modal *can*, unmarked for person or number. In addition, the corpus results indicate that if one were to cut the sentence in half, one would find that *flying* in *flying planes* is usually a gerund participle clause, while noun phrases usually combine with *can be dangerous*. These frequency effects may make the sentence obviously ambiguous for the listeners or readers, but it is on the first part of the sentence that the onus of decyphering rests.

Ambiguity cannot be verified objectively, without the intrusion of a person's intuition and knowledge of the world and language. It is people who identify ambiguity and propose different readings at the outset and it is the human brain that identifies possible connections which sets of raw data cannot retrieve on their own. While corpus data can provide unexpected results, it is the linguist who draws out its relevance and fills in the picture. One way of verifying the interpretation of the corpus data has been proposed by asking subjects to paraphrase the ambiguous sentences. Should the results from such experiments corroborate the interpretation of the corpus data, it would prove that frequency has a role to play in ambiguity resolution.

References:

- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Massachusetts: The MIT Press.
1993. *Lectures on Government and Binding: The Pisa Lectures*. Berlin: Mouton de Gruyter.
2002 [1957]. *Syntactic Structures*. 2nd edition. Berlin, New York: Mouton de Gruyter.
Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge: Harvard University Press.
Kess, Joseph F. and Ronald A. Hoppe. 1981. *Ambiguity in Psycholinguistics*. Amsterdam: John Benjamins B.V.
Oaks, Dallin D. 2010. *Structural Ambiguity in English*. London and New York: Continuum.

- Searle, John R. 1972. "Chomsky's Revolution in Linguistics" in *The New York Review of Books* (June 29, 1972 Issue).
- Stageberg, Norman C. 1952. "Obstacle-words in Group Conferences" in *The Journal of Communication* 2(1): 82-87.
1958. "Some Structural Ambiguities" in *The English Journal* 47(8): 479-486.
1966. "Structural Ambiguity: Some Sources" in *The English Journal* 55(5): 558-563.
- 1968a. "Structural Ambiguity for English Teachers" in *Selected Addresses Delivered at the Conference on English Education* (6): 29-34.
- 1968b. "Structural Ambiguity in the Noun Phrase" in *TESOL Quarterly* 2(4): 232-239.
- Zipf, George Kingsley. 1949. *Human Behaviour and the Principle of Least Effort. An Introduction to Human Ecology*. Cambridge: Addison-Wesley Press.
- Zwicky, Arnold M. and Jerrold M. Sadock. 1975. "Ambiguity tests and how to fail them" in *Syntax and Semantics* 4. John Kimball (Ed.). New York: Academic Press, pp 1-36.

Note on the author

Nadia CEHAN is a graduate of Babeş- Bolyai University of Cluj. Her research focus includes Corpus Linguistics and Language Learning.