ENG 5040 – Creative Research in English and the Media

MA in English and the Media

An analysis of sentence length and complexity in news articles

January 2014

Elaine GeradaGatt

83880M
Aim

The aim of this research is to point out stylistic similarities, differences and readability levels in three different articles written about the same news item and chosen from two different media delivery platforms. Since there are numerous stylistic features which could be assessed, the two stylistic features chosen for this research are sentence length and sentence complexity. The media delivery platforms are the print platform and the digital/web platform.

One article was selected from The Economist news magazine\(^1\), another article was chosen from the weekly printed edition of The Telegraph\(^2\) and the other article was selected from The Telegraph online newspaper.\(^3\) All of the articles focus on Typhoon Haiyan which devastated the Philippines between November 2 and November 11, 2013.

---

\(^1\) ‘Typhoon Haiyan Worse than Hell’, The Economist, 16 November 2013, p. 47.
Definitions and Research tools: Sentence and sentence length

A sentence is made up of a number of words which when placed together convey a particular statement, question, exclamation or command. A sentence usually consists of a main clause also referred to as an independent clause together with, one or more subordinate (dependent) clauses.

According to The Oxford Language Toolkit, there are delineated ways which should be adhered to when looking at sentence structure. A sentence should initiate with a capital letter and finish with a punctuation mark. Each of these punctuation marks will definitely give the sentence a different meaning, but this is not the purpose of this study.

For this particular research, the concept of a sentence will mean a set of words which create a statement, question, exclamation or command. The first word in the statement, question, exclamation or command will start off with a capital letter and the whole set of words will end with either a full stop, a question mark or an exclamation mark.

Coulthard claims that since written text is interactive, text analysis must be aware of the ‘purpose and process’ of its creation. Sentence type therefore can affect the clarity of the written piece. Sentences can be Simple, Compound, Complex and Compound Complex.

---

7 Refer to Appendix 2 for further information regarding sentence variety.
Thenumber of words for each sentence together with the number of sentences per paragraph will be recorded to conduct further analysis. Each sentence was then analysed in terms of the type of sentence used: simple, compound, complex and compound-complex. The Flesch Reading Ease test and the Flesch-Kincaid Grade Level Test will be applied to each of the texts. The results will be compared to the analysis of sentence complexity carried out for each article.
**Literature review**

Etymologically, the word 'news' is derived from the word 'new'. It is in fact the plural form and used to indicate something novel. Today the concept of news generally refers to anything that is happening currently and which is considered novel enough to be printed, broadcasted, shared on the Internet or shared during conversations. As Itule & Anderson point out news has always been of immense interest to the public and it is because of its interest that it is today shared on numerous media platforms.

The availability of media platforms to share news has led to the development of different news styles for each medium. A perspective shared by Cotter, Fang and Itule & Anderson, deliberately points to the fact that not only each medium has different requirements, but each medium is also consumed in a different manner at different points in time and by different audiences. Fowler, Cotter, and Fang perceive time as an important factor in that each medium lends itself to different styles and in turn each style adapts to each medium accordingly; all this takes place in a continuous evolutionary process.

Fowler importantly claims that news is ultimately a 'product', produced and shaped various institutions. Through the information provided and disseminated, news reflects and shapes the current values of a particular society at a particular point in time. Here Fowler draws on Halliday’s

---

9 Itule and Anderson p.62.
12 Itule & Anderson, p. 265.
notion of ‘function’ in a language. Halliday outlines three functions for language – language as a means to express content, language as an expression of the speaker’s attitudes and language as having a textual function. The third function which is more important for this research, is the function through which the text produced by the writer becomes recognisable by the reader. Language becomes essential in the creation of a text as it continuously strives to meet the new demands.\(^\text{14}\)

Fowler\(^\text{15}\) adds that printed news does construct the concept of ideal readers. The reader thus knowingly or unknowingly chooses a particular ideological position by the type of printed news s/he chooses to read. Similarly to Coulthard, Fowler believes that discourse is always meant for an implied addressee – ‘an imagined subject position which it requires the addressee to occupy’\(^\text{16}\) Coulthard shows that writing a particular piece of text intended for those who are able to read such a text is not an easy task. The audience is therefore imagined and the author’s decisions on what should be included in the text are based on this imagined audience.\(^\text{17}\) Consequently, once that text is published it starts to define an audience. So the text can be seen to be written for a specific reader but then the same text shapes the reader.

Despite the various media platforms available, the press still retains an important role in our society. People are still currently acquiring information through particular language used in printed news. Naturally, newspapers and news magazines have assumed a certain importance to linguists

\(^{15}\)Fowler, pp. 232-234.
\(^{16}\)Ibid., p.232.
\(^{17}\)Coulthard, p. 4. Speaking about the reader, Coulthard brings out the excellent example of the student who has to assume that the marker of his/her exam paper is intelligent and well-read but at the same time s/he has to assume that the Marker does not know about the situation to be assessed. This provides an excellent base from which one can look at the perspective of the author and his intended reader.
because of the way the information is written and presented to an audience.\textsuperscript{18} Tiono\textsuperscript{19} maintains that printed news categorically determines what news is to be placed within a newspaper and given to society. Similarly to Coulthard, Tiono claims that the way news is presented on any medium will also determine the audience that is going to read it. This introduces the concept of power highly emphasised Van Dijk\textsuperscript{20} as well as Fowler\textsuperscript{21} who claim that the grammar of a language becomes a product of the speaker’s and the reader’s mind. The collective meaning that the audience places on this linguistic aptitude is highly relevant here. Van Dijk\textsuperscript{22} believes that in understanding news media and the messages they convey, implies examining the structures and strategies of such discourses and their association with the audience as well as with institutional arrangements.

Structure and style therefore become essential in news writing. Although Crystal and Davy’s publication is relatively old in the sphere of linguistics, yet it still poses various interesting arguments, some of which are still applicable today. For instance, they do mention that there can be a number of journalistic styles, saying that there \textit{“is not one, but a number of journalese”}\textsuperscript{23}. News reporting may therefore have similar features across different media platforms but the overall styles are very different. Such styles may also differ because of the different audiences that each medium attracts. They claim that the different \textit{journalese} are basically made up of a varied blend of characteristics of journalistic material. In line with the varying journalistic styles explained by Crystal and Davy\textsuperscript{24} Cotter\textsuperscript{25} also speaks about the variation that takes place across the medium of delivery. Regardless whether such a medium is print, television, radio, or web, each channel endorses

\textsuperscript{19}Tiono, p.49.
\textsuperscript{21}Fowler, pp. 232-235.
\textsuperscript{22}Ibid., pp. 18-20.
\textsuperscript{24}Ibid., p. 175.
\textsuperscript{25}Cotter, p. 213.

Elaine GeradaGatt (83880M)
different elements of visuality, interactivity and comprehension. Printed texts are designed to attract
the reader’s attention by means of “a (syntactically) compressed but easily understandable rendering
of the information”. Similarly to Cotter however Crystal and Davydo acknowledge the need for
compression of information into a limited space which results in short sentences and paragraphs, the
need for clarity, precision and the avoidance of ambiguity. All factors are meant to increase ease of
readability in order to encourage more readers to read the particular text.

Cotter argues that reporters write with two important, linked goals in their mind. One goal is
the content of the news story and the other is the shape of the text or story design. News writing
text goals involve all the elements of story organisation as well as a good lead, appropriate sentence
structure, attribution and quotes. Both Fang and Itule & Anderson agree with Cotter that story
design remains an essential element in news writing.

Newspaper readership often depends on a relatively quick scan of the headlines and only if
those are effective enough to attract the reader’s attention will the reader go on to read the article.
Fang claims that a well-structured article will allow the reader to go through the text without
continuously going back to the previous paragraph to make sense of the story. If the reader has to
refer to the previous paragraph often, then it is more likely that the reader will give up on the article
and resort to other media to obtain news. Clarity and sentence length and structure are essential.
This is also commented upon by Cotter claiming that accuracy refers to both facts and content,
both of which must be correct. Since there are many claims on the readers’ and listeners’ time, the

26 Cotter, p. 28.
27 Cotter, p. 136.
29 Itule and Anderson, p. 60.
31 Fang, p. 10.
balance between content and participants is essential. Participants need to get to the news easily. This is further strengthened by McIntyre\(^{33}\) who states clearly that writing long complex-compound sentences might prove to be quite difficult for the reader to understand. Grabbing the reader’s attention and comprehension are paramount and therefore, headlines and paragraphs are short and highly informative, in the quest of being heard, read and understood by news consumers.\(^{34}\) Incidentally, McIntyre in fact comments on the way sentence length in English, has shrunk over time. He shows that research conducted by United Press International and the Associated Press has shown that readers find shorter sentences easier to understand than longer ones. McIntyre\(^{35}\) includes a table of results in order to sustain this argument. The table can be found below:

<table>
<thead>
<tr>
<th>Sentence Length</th>
<th>Readability</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 words or less</td>
<td>Very easy to read</td>
</tr>
<tr>
<td>11 words</td>
<td>Easy to read</td>
</tr>
<tr>
<td>14 words</td>
<td>Fairly easy to read</td>
</tr>
<tr>
<td>17 words</td>
<td>Standard</td>
</tr>
<tr>
<td>21 words</td>
<td>Fairly difficult to read</td>
</tr>
<tr>
<td>25 words</td>
<td>Difficult to read</td>
</tr>
<tr>
<td>29 words</td>
<td>Fairly difficult to read</td>
</tr>
</tbody>
</table>

**Figure 1. Sentence Length and Readability table.**\(^{36}\)

Crystal and Davy however are not completely convinced by the concept of sentence length as a tool to analyse style and readability. They put emphasis on content rather than style, but then for

\(^{34}\)Cotter, p. 26.
\(^{35}\)McIntyre, p. 15.
\(^{36}\)McIntyre, p. 19.
Crystal and Davy\textsuperscript{37}, the type of sentence used is much more important. Sentence types and structures including the ways through which sentences link up with each other and within and between paragraphs are highly essential.

**The Flesch-Kincaid Grade Level Test and the Flesch Reading Ease Test**

These tests which rely on word length and sentence length are meant to indicate comprehension difficulty levels in texts. Although both systems essentially use the same core measures, they have different weighting. The equations used to work out grade levels and reading ease are as follows:

Flesch-Reading Ease Test:\textsuperscript{38}

$$206.835 - 1.015 \left( \frac{\text{totalwords}}{\text{totalsentences}} \right) - 84.6 \left( \frac{\text{totalsyllables}}{\text{totalwords}} \right)$$

Flesch-Kincaid Grade Level Test:\textsuperscript{39}

$$0.39 \left( \frac{\text{totalwords}}{\text{totalsentences}} \right) + 11.8 \left( \frac{\text{totalsyllables}}{\text{totalwords}} \right) - 15.59$$

The variables within both tests are sentence length (number of words per sentence), total number of sentences and total number of syllables. Both tests show different results. The results obtained should be interpreted as follows. The readability ease results will range from 0 to 100. If a high score is achieved, then the score should be easy to read. This test therefore equates difficulty levels of texts with shorter words and sentences. The table below provides a clear picture for the interpretation of the readability ease test results:

\textsuperscript{37}Crystal and Davy, p. 174.
\textsuperscript{38} The Flesch Reading Ease Readability formula (MyByline Media) <http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php> [accessed 8 December 2013].
\textsuperscript{39} The Flesch Grade Level Readability formula (MyByline Media) <http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php> [accessed 8 December 2013]
Through the Flesch-Kincaid grade level test, the values obtained in the readability tests are converted to grade levels in order to make it easier to equate the values obtained above with a school grade level. Therefore the results obtained in the Flesch-Kincaid grade level test correspond with a particular school year. Hence, if the score is 9.3 for example, that corresponds to Year 9 level (locally known as Form 3 level).

Thus it can be deduced that the higher the score achieved on the Flesch Reading Ease test, the lower the score on achieved in the Flesch-Kincaid Grade Level test. Conversely the lower the score the more difficult to read the passages are.

Undeniably, these tests have also been criticised for their pitfalls by various critics and bloggers. Loftus\(^41\) claims that such tests rely on the number of syllables per word, and therefore rely on inflexible standards. Since syllables are delineated by vowels, words which are made up of three letters or less would count as single syllables, whilst words which are made up of consecutive vowels would count as one syllable. According to Loftus\(^42\) then this leaves a loophole and reduces the

---

\(^{40}\)http://www.readabilityformulas.com/flesch-grade-level-readability-formula.php  
credibility of such tests. Other pitfalls include the use of punctuation marks such as full stops and semicolons which act as delimiters. For the purpose of this research however, these tests will be conducted on all the three articles. Bearing in mind the various limitations, they will still provide an adequate scale with which to compare texts.
Results

The article in the Economist\(^43\) has the highest number of sentences per paragraph and the least number of words per sentences. This is followed by the article in the Telegraph online newspaper\(^44\) both in terms of sentences per paragraph and in terms of words per sentence. The article in The Telegraph printed newspaper\(^45\) has the lowest number of sentences per paragraph and the lowest average of words per sentence.

![Average Sentences Per Paragraph](image)

Figure 3. Average sentences per paragraph

\(^{43}\) ‘Typhoon Haiyan Worse than Hell’, *The Economist*, 16 November 2013, p. 47.


\(^{45}\) David Elmer, ‘Survivors of the typhoon roam streets with guns’, *The Telegraph*, 13-19 November 2013, p.15.
The article in the Economist obtained the highest score in the Flesch Reading Ease test with a value of 58.5%, followed by the article in The Telegraph (online) and the article in The Telegraph (print). When the values of the Flesch Reading Ease test are compared with the average sentences per paragraph for each article it follows that the article in The Economist has the highest score in terms of reading ease from all the three articles.

Consequently the average wordcount per sentence in each article has shown that The Economist has the least amount of words per sentence and it has the lowest value in the Flesch-Kincaid Grade level test. This is followed by the article in the Telegraph (online) which ranks second in terms of words per sentence and which has also ranked second among the three article in the Flesch-Kincaid Grade level test. Conversely the article in The Telegraph (print) has shown the highest average word count and the highest Flesch-Kincaid Grade level value of all three articles. The argument for brevity outlined by Cotter does not hold for The Telegraph (print).

---

These results can lead to the assumption therefore that the article in The Economist which has the highest number of sentences per paragraph, also shows a higher percentage value in the Flesch Reading Ease test. This can contribute to the fact that such an article proves to be less difficult to read. This is further substantiated by the Flesch-Kincaid Grade level value, wherein the same article obtained a value of 8.5, suggesting that it can be read by a grade 8 student (12 year old).

![Figure 5. Flesch reading ease](image)

![Figure 6. Flesch-Kincaid grade level](image)
When analysing sentence complexity, the results indicated that complex sentence type is the most widely used type of sentence in all the three articles. A complex sentence is made up of an independent clause which is joined by a single or multiple dependent clauses.\textsuperscript{47} This is followed by the compound sentence type which is a sentence made up of two independent clauses joined by a coordinator. This is followed by the simple sentence type which is a sentence made up of an independent clause and the compound-complex sentence type which is made up of two main clauses and a dependent clause.

The complex sentence was mostly used by the article in The Telegraph (online), followed by the Telegraph (print) and by The Economist. The compound sentence type was roughly used with the same frequency in all the three articles. The simple sentence category however indicated major discrepancies in overall use within the three articles. The article in The Economist showed the lowest frequency of simple sentence use while the article in The Telegraph (online) made use of the simple type of sentence slightly more. The article in The Economist however had over 30% of the sentences of the Simple category type showing a great discrepancy overall. The use of the compound-complex sentence fared the lowest in all the three articles, with the articles in The Telegraph (online) and The Economist making roughly the same use of this type of sentences. These were however superseded by the article in The Telegraph (print) which indicated the highest use of the compound-complex type of sentence from all the three articles.

\textsuperscript{47}Itule and Anderson, p. 33.
Following the above information, it can be argued that the article in The Economist, is therefore the easiest to read while the article in The Telegraph (print) is the most difficult to read out of all the three articles. This implies that Fang’s\textsuperscript{48} and Cotter’s\textsuperscript{49} concepts that news items have to be brief and to the point applies to the article in The Economist. Cotter further claims that readers need to get to the news easily, since there are increasingly more items trying to grab the readers’ attention at the same time. The article in The Economist, definitely strengthens this point. The article in The Telegraph (print) however indicates a notable discrepancy between what Cotter\textsuperscript{50} and Fang\textsuperscript{51} argue in terms of sentence brevity since the sentences in the article in The Telegraph (print) are relatively long when compared to the other two articles. There is also a notable discrepancy in sentence length

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{sentence_complexity.png}
\caption{Sentence complexity}
\end{figure}

\textsuperscript{48}Fang, p.10.
\textsuperscript{49}Cotter, p. 116-118, 136.
\textsuperscript{50}Ibid., p. 116-118, 136.
\textsuperscript{51}Fang, p.2-3.
between the article in The Telegraph (print) and The Telegraph (online), wherein the online article uses more shorter sentences than the printed version.

Contrary to Crystal and Davy\textsuperscript{52}, sentence length might actually be a relevant tool to indicate difficulty levels of texts. They do emphasise however that sentence type is a relevant tool in text readability.

If tests such as the Flesch-Kincaid Grade Level and the Flesch Readability Ease indicate comprehension difficulty in texts and if sentence complexity can substantiate the information obtained through such tests then it can be argued that Coulthard's viewpoint that texts are actually designed for particular audiences still holds. This can be taken further arguing similarly to Coulthard\textsuperscript{53} that text can also define and shape the reader.

\textsuperscript{52} Crystal and Davy, p. 174-5.
\textsuperscript{53} Coulthard, p. 4.
Conclusion

As previously argued, new technologies, more particularly Internet, are playing an extensively significant role in altering discourse forms. Print media whether on paper or online have also implemented new modifications over time to keep up with the ever increasing challenges. Stylistic features such as sentence length and complexity are continuously adapting not only to the types media, but also to continuously compete with all the other factors which are trying to grab the readers’ attention.

This research study is by no means a comprehensive study of sentence analysis in print and online media, conversely it is merely the beginning of a much larger undertaking. Indeed Carter and Nash do claim that analysis of style in language requires examining a relationship between the different levels of organisation in a language. Following Cotter, the linguistic study of media is still pretty much in its early stages. The media message has often been the central concern of numerous scholars, however detailed studies about the language in the media are still in their infancy. It is hoped that through this study, the interest in stylistic features in the media will continue to grow.

54 Carter and Nash, p.15.
55 Cotter, p. 20.
Appendix 1

Background to The Economist news magazine and The Telegraph print and online newspapers

The Economist, despite the fact that it is currently issued in the format of a magazine, can be considered as a newspaper primarily because historically it incorporated two British newspapers, these being the ‘Bankers Gazette’ and the ‘Railway Monitor’ between 1845 and 1932. It has since then covered news of an economical and political agenda. In the About us section within the Economist website, the quote “Many hands write The Economist, but it speaks with a collective voice.” defines clearly why none of the articles published in The Economist show their author. Its circulation currently stands at 1.5 million.56

The Telegraph which was founded in June 1855 started as The Daily Telegraph and Courier. Its current circulation is 552,065 (as at 2013) The Telegraph also includes a number of sister publications such as: The Sunday Telegraph, The Young Telegraph as well the website.

The website for The Telegraph was launched in 1994 and has since then provided readers with the current news items in line with the daily publication.57

Appendix 2

The simple sentence

The simple sentence consists of a single subject and a single predicate forming one independent clause. The following example shows a simple sentence - “Early estimates of the economic cost are about $15 billion.”58

The compound sentence

The compound sentence comprises two independent clauses. These are joined together in one sentence through a coordinator. The following is an example of a compound sentence - “At least 300 are already reported dead on Samar Island, with 2,000 missing, but much of the island remains completely cut off from the outside world with no power or communications.”59

The complex sentence

The complex sentence is made up of an independent clause which is joined by a single or multiple dependent clauses as shown by the following example - “Mr Aquino said his government had deployed 300 soldiers and police to the city and promised more were on their way "to show the government's resolve and stop this looting," he said.” 60

---

58 Typhoon Haiyan Worse than Hell’, The Economist, 16 November 2013, p.47.
The compound complex sentence

The compound complex sentence includes aspects from both the complex and the compound sentences already described above. It comprises two main clauses and a dependent clause. The following is an example of a compound complex sentence - “Television footage shot from the air revealed entire streets flattened and roads strewn with trees and debris, raising fears that those of its 40,000-odd residents who did not flee before the storm hit may not have survived Haiyan’s impact.”61

---

Bibliography


Newsom, Doug, and James Alvin Wollert, Media Writing: News for the Mass Media (Wadsworth Publishing Company, 1985)


Van Dijk, Teun A., “Power and the News Media,” Political communication in action, 1996
http://www.discourses.org/OldArticles/Power and the news media.pdf> [accessed 2 February 2014]

Wardrip-Fruin, Noah, and Nick Montfort, eds., The NewMediaReader, (Massachusetts: MIT Press, 2003)