



## THE NEW YORK TIMES AND CHINA DAILY ONLINE – LINGUISTIC COMPLEXITY IN NEWS HEADLINES

**Shujun Wan<sup>i</sup>**

Teaching Assistant,  
Department of English Studies,  
Xuhai College,  
China University of Mining and Technology,  
China

### **Abstract:**

News headlines play an important role in attracting readers' attention. By comparing 200 online news headlines collected from the New York Times and China Daily online, this paper aims at finding out the difference in linguistic complexity of English online news headlines in a native English speaking country and a non native English speaking country.

**Keywords:** online news headlines; lexical complexity; syntactic complexity

### **1. Introduction**

News plays an important role in people's daily life. There are multiple ways to access to news nowadays, traditional newspapers and magazines, radio, TV news programs, and of course, the Internet. The traditional printed newspaper, in the current digital age, however, seems insufficient in serving its original function of disseminating information. Therefore, online news, due to its easy access, convenience and mostly free of charge, has seen a significant rise in readership. Reuters Institute (Newman et al. 2018: 9) found out that the average number of people paying for online news has increased considerably in many countries, especially those in Nordic areas. As the case in China, with the economic development and the trend of globalization, several governmental news agencies have set up their own news websites in English version to publicize themselves in a better way. These news websites include *Xinhua.Net* (<http://www.xinhuanet.com/english/>), *China Daily*, (<http://www.chinadaily.com.cn/>), *People's Daily* (<http://en.people.cn/>) and so forth.

As one of the most researched areas of news discourse, the headline, as the eye of a news article, has received great discussion because it has been looked into from various aspects, e.g. historical development, linguistic structure and etc. In the digital era, analysis concerning news headline from electronically-mediated context could shed more light on relevant linguistic research.

---

<sup>i</sup> Correspondence: email [wansjcaptain@outlook.com](mailto:wansjcaptain@outlook.com)

The outline of the paper will be as follows: the first section will be a literature review of news, news headlines and linguistic complexity. Then research gaps are identified after the comprehensive review of previous research, which give rise to the research questions. Afterwards, I will introduce the research data and methodology in the present study. The next section will be a comprehensive analysis of the collected data using quantitative methods. Then there will be a summary and a discussion of the findings. Finally, I will discuss the outlook of future research.

## 2. Literature Review

### 2.1 News and Online News

According to the Cambridge Dictionary, the word “news” refers to information or reports about recent events, whether locally, domestically or worldwide. Therefore, concerning the study of news, “*news discourse is about past political, social or cultural events*” (Dijk 1988: 5).

Dijk (1988) categorized media news into three main categories. The first is new information about events, things or persons. Second is a (TV or radio) program type in which new items are presented. Last but not least, it implies a news item or news report, i.e., a text or discourse on the radio, on TV or in the newspaper, in which new information is given about recent events (Dijk 1988: 5).

Normally the printed newspaper would be the main way for people to access to news. According to Lewis (2003: 96), news in the mass media is defined by “*the reform and distribution of traditional print and broadcast*”. With the development of the Internet, online news has been playing an increasingly important role in spreading the latest information. Reasons for this are not hard to get: people can read online news anytime and anywhere as long as they have access to the Internet. Lewis (2003) has concluded why electronic communication can be easily relevant to news dissemination. Integrating writing, sounds and visuals, electronic mechanism contributes to the transmission of information in a considerable quantity and thus leads to multiple ways of interaction among Internet users.

Hence, online news gradually takes the place of traditional printed newspaper. As shown in Reuters Digital News Report 2016 (Newman et al: 2016), from 2012 to 2016, online news remains to be the top of people’s news source.

### 2.2 News Headline

Before tapping into previously conducted research regarding the theme topic, we need to look at news headline from a general overview. Headline is defined by the Cambridge Dictionary (<https://dictionary.cambridge.org/dictionary/english/headline>) as “*words printed in large letters at the top of a newspaper story that serve as its title*”.

With features of being at the top and printed large, a headline “*presents more or less detailed information of the content by providing a short summary*” (Schneider 2000: 48) and must draw readers’ attention. This basic function has been confirmed by Schneider (2000) who states that by describing the essence of a complex news article in several words, headline informs the readers in fast speed with accuracy and arouses the readers’ willingness to continue reading. To confirm this statement, Schneider (2000: 48) quotes White et al. (1993: 12) to further explain this point:

*“the headline should tell you in a glance what the story is, or else intrigue you enough so you want to find out what it is by reading on”.*

Headline’s function as information providing seems to have been widely accepted. As discussed above, headlines should *“encapsulate the story in a minimum number of words, attract the reader to the story”* (Reah 1998: 25). Most importantly, the headline itself is a source of information. Normally, elements including what, who, where and how should be included in a piece of information to cover things about what happened, who was involved, where it happened and what the situation was. Reah (1998) proposed the language structure of the headline may be non-standard. He stated that problems could possibly be encountered when analyzing the language of a headline from a what, who, where, how model because *“headlines may not contain a verb. The action is frequently nominalised (turned into a noun), which can distance the word or phrase from the actual action”* (Reah 1998: 25).

Visual function is another important function of headline proposed by Reah (1998: 23). First, as the Cambridge definition of headline implies, headlines are printed larger than the text of article, which attracts readers’ attention at first sight. Second, due to the reason that *“headlines work in conjunction with the other visual aspects of the newspaper text, in particular the pictures”* (Reah 1998: 23), a headline in the form of characters plus visual assistance can make the readers understand the news article in an easier way. Reah (1998: 28) further points out that headlines can act as *“opinion manipulators”* for reason that they can also be written to exert an impact on readers’ opinion.

There has been plenty of research concerning headlines in terms of linguistic features, gender representation, comparative analysis of headlines in different languages and etc. Schneider (2000) has done a research in the emergence and development of news headline from 1700 to the present days, using the Historical Rostock English Newspaper Corpus as the dataset. She found out that in the early days, headline functioned to act as the summary of news story by providing *“concrete information on the topic of the news report”* (Schneider 2000: 49). Beginning with an average of less than 3 words per headline, the average length of headline reached to more than 7 words in 2000. Also, headline writers have always been trying to express relations by adopting relational nouns, relational adjectives, relational prepositions, and verbs. What is more, popular papers play a more important role than quality papers in introducing new trends in journalism by using more headline words per text.

Apart from the historical development, a research on headline concerning women demonstration was done by Kleinke (2000) by comparing English and German local newspaper. Viewing headline as a structural component of discourse, Kleinke (2000: 68) aimed at showing how headlines *“as one special aspect of newspaper policy, contribute to the systematic underrepresentation of women in local broadsheet newspapers”*. Kleinke’s (2000) research found out that topic and syntactic structure are the two most obvious aspects where the two newspapers bear differences. In the German newspaper, women are equally demonstrated in terms of the newspaper layout when it comes to the topic “Sports”. In the Scottish newspaper, however, “Crime” is the only topic choice where the marginalization of women is overcome at the layout level. Syntactically, in the Scottish newspaper, women take up lowest with 21% of all 119 female references in headlines in the syntactic position of a subject in SVO-constructions.

Napu (2018) contrasted English and Indonesian newspaper headlines in terms of lexical characteristics. He found that headlines in English and Indonesian share similarities in terms of nouns, active voice and verb omission while differ from each other concerning tense, conjunction, preposition and acronym. English headlines have more part of speech than the Indonesian ones. The number of nouns in Indonesian headlines slightly outnumbered that of English headlines (64.14% vs 61.33%), among which proper nouns and acronyms presented more in the Indonesian sample. Napu (2018: 115) also pointed out that English headlines use more conjunctions and prepositions than the Indonesian headlines and both samples prefer colon to the word “say”.

Gao (2016) conducted a comparative research on the conceptual metaphors in analyzing English and Chinese economic news headlines. Major finding was that the metaphorical expression can be classified into structural metaphor, ontological metaphor, and orientational metaphor. Gao (2016: 2633) pointed out that journey metaphor, game metaphor and war metaphor are the same subcategories of conceptual metaphors in both Chinese and English economic news headlines.

While the above reviewed studies have provided a contrastive view in headlines per se, research investigating news headlines from an English-native speaking newspaper or website and those from English as a Lingua Franca newspapers or websites has been adding more color to the study of headlines. Hu and Guo (2013) contrasted more than 200 online English sports news headlines from two English Websites, *Xinhua.net* and *Yahoo.com*. They found out that within the American online sports news headlines, short words, nouns and verbs are used, vogue words acted also as a preference. The conversion of parts of speech was an effective expression in sports news headlines. Non-finite verbs were frequently used. Moreover, the omission of words is a common linguistic phenomenon in sports headlines. When it comes to sentence structure, Hu & Guo (2013: 91) noticed that punctuation occupies least space than any other words to condense the headlines. Also, American online sports news headline employed common rhetoric devices to “*make expressions like word pictures*” (Hu & Guo 2013: 91).

Headlines collected from *Xinhua.net* tend to directly point out the results of sports events by using statistics and “*have fewer narratives about the news’ element of ‘how’ than that of the Yahoo*” (Hu & Guo 2013: 91). Lexically, abbreviations of federations, organizations and events are employed. There is no vogue word in the sample headlines and the conversion of parts of speech is rarely to be found. In terms of grammar, the most obvious difference is that the use of non-finite verbs as predicates, omitting functional words, or taking punctuation as conjunctions or verbs in the sentences, is rare. Also, using rhetorical devices can be a difficult task for Chinese news writers because “*they would have a big vocabulary, a good grasp of knowledge of the sports English, and proficient skills of English writing*” (Hu & Guo 2013: 93).

While Hu & Guo’s (2013) research lays more emphasis on the use of words in sports headlines, Lee (2012) focused on the structure pattern of online political news headlines by studying the similarities and differences between headlines of English in America and those in Korea. Lee (2012) concluded that both varieties of headlines employ similar types of constructions and both follow the conventional news headlines features. The English headlines from American (abbreviated as AE in Lee’s (2012) research) news websites use more sentential

structures, while those from Korean English (abbreviated as KE in Lee's (2012) research) news websites are shorter and prefer to use NP+ to V constructions and Verb-less constructions.

As for the AE headlines, Lee (2012) found out that simple sentences crouched in present tense ranks the top among all sentence constructions. Verb-less constructions, in which sentences that have deleted linking verbs including *be* or *become* were categorized, were the second most frequent structure pattern. Regarding to the KE headlines, similar result was that simple sentence also ranked first in the structure pattern. All simple sentences except two cases are crouched in the present tense. Compared to the AE sample, less varied types of tense were employed in simple sentences in KE headlines because neither progressive constructions nor past tense sentences appear. One apparent difference lies in the use of noun phrase. In AE samples, the NP position is occupied mainly by proper names. While in the sample of KE, not only proper names, but also NP referring to people and to objects were used.

### 2.3 Linguistic Complexity

In terms of news, the importance of writing headlines is needless to say. As mentioned in section 2.2, headlines function as the summary of the news article to attract readers' attention. Therefore, headlines need to be written well.

As Light (2001) (cited in McNamara et al 2010: 58) stated, the skills of writing are critical to professionals not only for their everyday work, but also for further promotion within their disciplines. However, to write English well for ESL and EFL learners could be a challenging task. Silva (1993) (cited in Wang 2016: 81) found that differences concerning fluency, accuracy and syntactic structure are significant between the written texts of native speakers and second language speakers. Hinkel (2003) (cited in Wang 2016: 81) also found that L2 writers tend to overuse simple sentence structures.

Studies have been done to investigate lexical richness and syntactic complexity concerning both writing and oral performance for L2 learners. Lu (2012) stated that lexical richness manifests L2 users' sophistication and range of productive vocabulary. Lu (2012) reviewed a variety of lexical richness measures proposed in the language acquisition literature. The most traditional one is type-token ratio (TTR), which is the ratio of the number of word types (T) to the number of words (N) in a text.

Syntactic complexity refers to the sophisticated linguistic structures in language use (Ortega 2003: 49). Wolfe-Quintero et al (1998: 69) defined it as "*a wide variety of both basic and sophisticated structures which are available and can be accessed quickly, whereas a lack of complexity means that only a narrow range of basic structures are available or can be accessed*".

The use of automated tools brings more sophistication to the research on linguistic complexity. Coh-Metrix is used to "*examine linguistic indices of cohesion, syntactic complexity, the diversity of words used by the writer, and characteristics of words*" (McNamara et al 2010: 57). Lexical Complexity Analyzer (Lu 2012) and Syntactic Complexity Analyzer (Lu 2010) are designed to examine lexical and syntactic indices. The former contains 25 lexical complexity measures chosen from measures adopted in Coh-Metrix and MLTD (measures of lexical and textual density), the latter selects 14 syntactic complexity measures from the large set of measures reviewed in Wolfe-Quintero et al (1998) and Ortega (2003).

Considerable studies have been carried out to the examination of linguistic complexity in the EFL/ESL writing and oral performance. In Skehan's (1998) L2 research pattern, syntactic complexity is one of the three indicators in assessing L2 learners' writing proficiency. The other two indicators are accuracy and fluency. Casanave (1994) did examinations into development of writing performance of 16 Japanese English learners over a period of three semesters in terms of the measures of number of words per T-unit, T-unit complexity ratio, complex T-units ratio, error-free T-units ration and error-free T-units length. The study shows with the improvement of language proficiency, more complex and accurate sentences are produced.

By placing 169 L2 learners into 2 groups with high and low proficiency, Ferris (1994) studied the L2 learners' essays using 28 lexical and syntactic indices. Ferris (1994) found that the two groups of L2 learners vary from each other in 18 indicators and learners with high language proficiency produce more relative and adverbial clauses.

Aimed at figuring out the best measures to assess EFL and ESL learners' language proficiency, Wolfe-Quintero et al (1998) examined measures with regard to accuracy, fluency and complexity in 39 second language writing studies. They found that the best measures are clauses per T/unit and dependent clauses per independent clause.

What is worth noting is that factors including different learners, task difficulties and contexts could also influence the manifestation of syntactic complexity. Ortega (2003) investigated the influence of instructional context and proficiency sampling criterion on syntactic complexity of ESL and EFL learners. It turned out that compared to EFL learners, the syntactic complexity in writing produced by ESL learners is higher and studies using program level yield out wider range of complexity level than those using holistic rating as the proficiency sampling standards.

In China, research concerning linguistic complexity can be grouped into cross-sectional research and longitudinal research. A cross-sectional study was conducted by Bao (2009) to explore the changing patterns of syntactic complexity in EFL writings by college students across four grades. Using T-unit analysis, the syntactic complexity is specified into unit length (T-unit length and clause length) and clausal density or degree of embeddedness (measured by T-unit complexity ratio and dependent clause ratio). The result was that learners' length indices experience the most rapid growth from Grade 1 to Grade 2 while the density indices reach the peak values in Grade 2, displaying no significant increase.

Compared to Bao's (2009) study focusing on T-unit analysis, Xu et al. (2013) carried out a comprehensive research, analyzing the syntactic complexity in English essays of Chinese students and native speakers by adding some syntactic structures as indices to conduct a corpus-based study. Compared with native speakers' writings, unit length indices and clausal density indices of Chinese students are lower. Chinese students also adopt less passives and reduced structures although with the progress of the grades. However, this study lacks the reliability to reflect the gap between Chinese English majors and average educational native speakers because the writings of native speakers are selected from journals or teaching materials.

Lu (2012) investigated the relationship of lexical richness to the quality of oral narratives among ESL learners by using Lexical Complexity Analyzer. The system covers 3 dimensions of lexical richness, including lexical density, sophistication and variation using 25 measures. Conclusion was that while lexical variation may have the strongest effect for the quality of ESL

learners' oral narratives, there was no evidence of significant correlation between lexical density and sophistication with test takers' rankings.

Using Lu's (2010) Syntactic Complexity Analyzer as the tool, Wang (2016) looked into the writings of EFL Chinese students and more proficient users by analyzing PS (personal statement). Results are as follow: the mean sentence length and the mean clause length of EFL Chinese learners were found to be lower compared to proficient users. EFL Chinese use more in terms of the number of clauses, DC/C and DC/T but fewer coordinate structures. Although there is no significant difference in the use of verb phrases, EFL Chinese learners use much fewer complex nominals than their language proficient counterparts.

### 3. Research Gap and Questions

The review conducted above reveals the potential of analyzing news headlines from different perspectives, whether in its historical development, syntactic or lexical features, the use of metaphor, or cross-cultural analysis and etc. However, there is still space for comparative research on online English news headlines between those in an English-native speaking environment and those in English as foreign language or English as Lingua Franca context.

In terms of English headline analysis in China, most research focuses on the use of metaphor or translation theory. Seldom have people studied linguistic complexity by comparing online political news headlines from English news websites. Therefore, this study aims at filling this gap and figuring out the similarities and differences of news headlines from a linguistic perspective.

In light of the research gap, the present study would address the following research questions:

- 1) What are the differences and similarities between American English headlines and Chinese English headlines in terms of lexical features?
- 2) What are the differences and similarities between American English headlines and Chinese English headlines in terms of structure pattern?
- 3) Can the findings of the present study be in line with previous analysis? If differences are attested, what are the differences?

Based on the results from the previous research reviewed in the above section, following are the research hypotheses:

First, lexically, headlines collected from the American news website (abbreviated as AE in the following sections) would use more nouns and prepositions. However, the presentation of statistics could be more apparent in the sample headlines from the Chinese English news website (abbreviated as CE in the following sections).

Second, concerning structure pattern, AE headlines would prefer longer sentences and use more tenses while CE headlines would be shorter and present tense would be more applied.

Thirdly, the results of the present study would be generally in line with previous research, but new findings would definitely be attested.

## 4. Data Collection

### 4.1 Instrument

For this study, to obtain authentic language material, 200 sample headlines were chosen from two quality national English newspaper websites: *The New York Times* (<https://www.nytimes.com/>) and *China Daily* (<http://www.chinadaily.com.cn/>). The reason for choosing quality English newspapers is that they “especially appear to benefit from the combination of high-profile investigative reports and a ‘digital first’ strategy” (Newman et al 2018: 111).

As one of the leading newspapers in America, *The New York Times* “saw digital subscription revenues rise by a fifth over 2016, itself a record year” (Newman et al 2018: 111). The online news website of *The New York Times* ranked the third after *Huffpost* and *Yahoo! News*. (Newman et al 2018: 112). In the brand trust section, *New York Times* ranked the 7th after local television news, *Wall Street Journal*, *ABC News*, *CBS News*, *NPR News* and *Washington Post*.

Target readers of *China Daily online* are non-Chinese native speakers who are interested in China, EFL and ESL learners who want to know more about news both home and abroad in English or teachers who use news articles published in *China Daily* as authentic materials for teaching aims such as news writing in English, Chinese-English translation and so forth.

According to the introduction from its official website ([http://www.chinadaily.com.cn/e/static\\_e/about](http://www.chinadaily.com.cn/e/static_e/about)), as China’s leading English-language newspaper, *China Daily* aims to introduce China to the entire world. And it is quoted more often by overseas media than any other Chinese publication. Its website, *Chinadaily.com.cn*, is China’s most influential English-language multimedia platform. Besides analyzing the collected headlines using my own method, the computational systems, namely Lexical Complexity Analyzer and Syntactic Complexity Analyzer will also be adopted to test whether my own method is applicable in terms of headline analysis and to examine the linguistic indices of the sample headlines in a more sophisticated manner, and most importantly, to add more accuracy and reliability to my methodology. The Syntactic Complexity Analyzer (Lu: 2010) and the Lexical Complexity Analyzer (Lu: 2012) are open to public by accessing it at <http://www.personal.psu.edu/xx113/downloads/l2sca.html>. The software analyzes data using Stanford Parser and also Treegex (Lu 2010: 475).

The web-based interface to the Lexical Complexity Analyzer allows users to analyze 25 measurements of lexical complexity, covering lexical density, lexical sophistication and lexical variation or range. The web-based interface to the L2 Syntactic Complexity Analyzer allows users to automatically analyze 14 different measures of syntactic complexity, covering length of production units, amounts of coordination, amounts of subordination, degree of phrasal sophistication and overall sentence complexity.



## 4.2 Data Retrieval and Filtering

**Table 1:** A comparison of text categories of the two websites

	The New York Times	China Daily
Text Category	World, U.S., Politics, N.Y., Business, Opinion, Tech, Science, Health, Sports, Arts, Books, Style, Food, Travel, Magazine, Magazine, T Magazine, Real Estate, Video	China, World, Business, Lifestyle, Culture, Travel, Watch this, Sports, Opinion, Regional, Service, Forum, Newspaper

The above table illustrated the different text categories from two websites. The total number of headlines of all text categories from the two news websites yielded more than 1000 within a five-day's period (18.02.2019 - 22.02.2019), each news website has over 100 headlines per day. To make the present study comparable to the one done by Lee (2012) (see Section 2. Literature Review) on the analysis in American English and Korean English online political news headlines, the scope of data was narrowed down on political news. Hence, 20 headlines were chosen each day randomly from each news website. The political headlines chosen from American news website (abbreviated as AE in the following sections) were sampled from sections that contain political news, including *World*, *U.S.*, *Politics* and *N.Y.* The Chinese English (abbreviated as CE in the following sections) counterparts come from two main sections, namely *China* and *World*. Therefore, there are 200 sample headlines in total with 100 from *The New York Times* and another 100 from *China daily*.

To analyze the sample data, each headline is assigned a unique identifier. For example, CD.01 stands for the first chosen headline from *China Daily*. In the same way, NYT.01 refers to the first chosen headline from *The New York Times*.

## 5. Data Analysis

### 5.1 General Findings

Before going deeper into the detailed analysis of the headlines from the two samples, we need to have a brief overview of the length of the headlines. The length of headlines was calculated in terms of the number of words used in the headline.

**Table 2:** An overview of headlines from both samples

	AE	CE
Whole number of words	1049	843
Number of headlines	100	100
Average number of words in all headlines	10.49	8.43
Number of the shortest headline	6	4
Number of the longest headline	14	14

As shown in Table 2, the whole number of words in the AE sample outnumbered that in the CE sample. In terms of the length of the longest and shortest headlines, although there appears not much difference, the AE sample is slightly longer than the CE counterpart.

## 5.2 Lexical Findings

The concept of “word” needs to be figured out before looking at headlines lexically. As stated by Quirk et al (1973: 8): the units of language have a meaning and a structure which relate them not only to the world outside language but to other words within the language. The study of words is the business of Lexicology but the regularities in their formation are similar in kind to the regularities of grammar and are closely connected to them.

By defining word as a “lexical item” (Halliday & Yallop, 2007: 3), Halliday & Yallop (2007: 3) explained that every language has a vocabulary, or ‘lexicon’, which forms one part of its grammar- or, to use a more accurate term, one part of lexicogrammar.

In this section, the usage of word class in the two samples would first be illustrated to determine the frequency of occurrence. Examples and tables would be provided respectively. Then a comprehensive comparison regarding to lexical features of the two kinds of headlines would be given so as to make a clear insight into this category.

The following linguistic elements would be included to analyze the two sets of samples:

- Nouns that are usually used to name things, places and persons. For example, *apple, boy*. This category would include proper nouns that are applied to refer to specific names of people, places, dates, and etc. For example, *Putin, Berlin, September*; Abbreviations and acronyms which are words formed from the initial letters of a fixed phrase or title such as *WTO, UN, Brexit* (as according to Cambridge Dictionary, *Brexit* is short for “British exit”. Hence, the word *Brexit* would be counted under this category). And definitely, common nouns such as *computer, tourist* and etc. (*Brexit* is counted not only as an abbreviation, but also as a proper noun)
- Verbs that usually denote processes, actions, or states (Verspoor & Sauter, 2000: 91). Examples include: *be, get, dance, sign* and etc.
- Conjunction: coordinate conjunctions *like and, but, or, not, for, so, yet*; correlative conjunctions such as *not only...but also, either...or..., both...and*; and conjunctive adverbs including *moreover, however, otherwise* and etc.
- Preposition: Prepositions are usually short words (Verspoor and Sauter, 2000: 109) like *in, on, at, about, with, of, to, by, before, and after*.

In terms of the category of verb, only the number of verbs would be analyzed in this section. Grammar features concerning tense would be analyzed in section 5.3.

**Table 3:** The use of part of speech in the 200 headlines

Frequency of different parts of speech in two sample headlines				
Part of speech	AE		CE	
	N	%	N	%
Noun	453	56,63	424	60,66
Verb	184	23	137	19,54
Preposition	142	17,75	136	19,46
Conjunction	21	2,63	4	0,5
Total	800	100	701	100

As can be seen from the above table, the difference in the number of nouns and prepositions is not that obvious. One apparent distinction lies in the use of conjunction. Although in both

samples, the percentage of conjunction is less than 3%, conjunction is employed far more frequently in AE headlines than in CE headlines. In the following section, the use in nouns, verbs, prepositions and conjunctions in the two sets of data will be elaborated in detail.

### 5.2.1 The Use of Nouns

**Figure 1:** The Frequency of Nouns in Two Samples

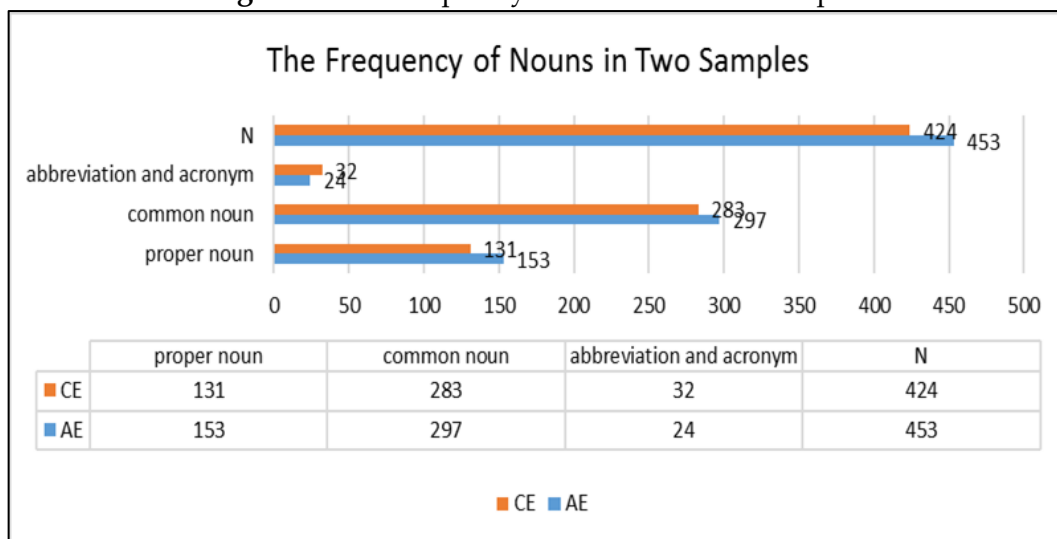


Figure 1 is a combination of a table and a bar chart illustrating the number of different categories of nouns used in the two samples and a comparison between the two samples. There are 424 nouns in the CE sample in total. The largest number of nouns that appears in one single headline is 9. Here is the case:

CD.96: *EU to end trade talks if US imposes car tariffs: EU Trade Commissioner*

In this example, the 9 nouns are *EU, trade, talk, US, car, tariff, EU, trade, and commissioner*.

Among all the 424 nouns, there are 131 proper nouns, 282 common nouns and 32 abbreviations and acronyms. In the 33 abbreviations and acronyms, *U.S.* appeared seven times; *PM* four times; *EU* three times; *5G, PLA, DPRK, FM, pct* twice each; *Brexit, ASUS, mln, VJ, MPs, CPC, IPR, Govt, Dept* once respectively.

In the AE sample headlines, there are 453 nouns used in total. There are 6 headlines that have 7 nouns, one of them would be explained here :

NYT.46: *Huawei Risks to Britain Can Be Blunted, U.K. Official Says, in a Rebuff to U.S.*

In the example above, the 7 nouns are *Huawei, risk, Britain, U.K., official, rebuff, and U.S.* Interestingly, one headline, which is an interrogative, has no noun. This headline is also listed here:

NYT.55: *Just How Endangered Is Watchmaking?*

153 proper nouns are used to refer to specific persons, place, date and etc. 297 common nouns can be counted. Among all the 100 headlines in the AE sample, there are 24 abbreviations and acronyms. *U.S.* appeared seven times; *ISIS* four times; *U.K.* and *Brexit* three times each; *U.N.* twice; *NATO*, *DNA*, *K.K.K.*, *N.H.*, and *CNN* appeared once respectively.

As can be seen in the above figure, there is not much difference in terms of the number of common noun or proper noun used in the two samples, but abbreviation and acronym is the only category that the CE sample outnumbers that of the AE sample. Both the CE and AE samples applied abbreviations and acronyms such as *UK*, *US*, *ISIS*, *PM* and etc. One kind of abbreviation and acronym that is employed particularly by the CE data is the use of unit of measurement. For example:

CD.30: *Hong Kong population grows by 0.9 pct to over 7.48 mln*

In this headline, the two units of measurement *pct* and *mln* were used. The former refers to “percent/percentage”, and the latter is the abbreviation of “million”. The reason for using abbreviation instead of unit of measurement can be to save more space and to make the number instead of the unit more apparent. Another feature that needs to be noted in the abbreviation and acronym category is that in the CE sample, some abbreviations and acronyms that are regarded as with “Chinese characteristics” appear quite frequently. Examples:

CD.15: *PLA welcomes launches of private rockets*

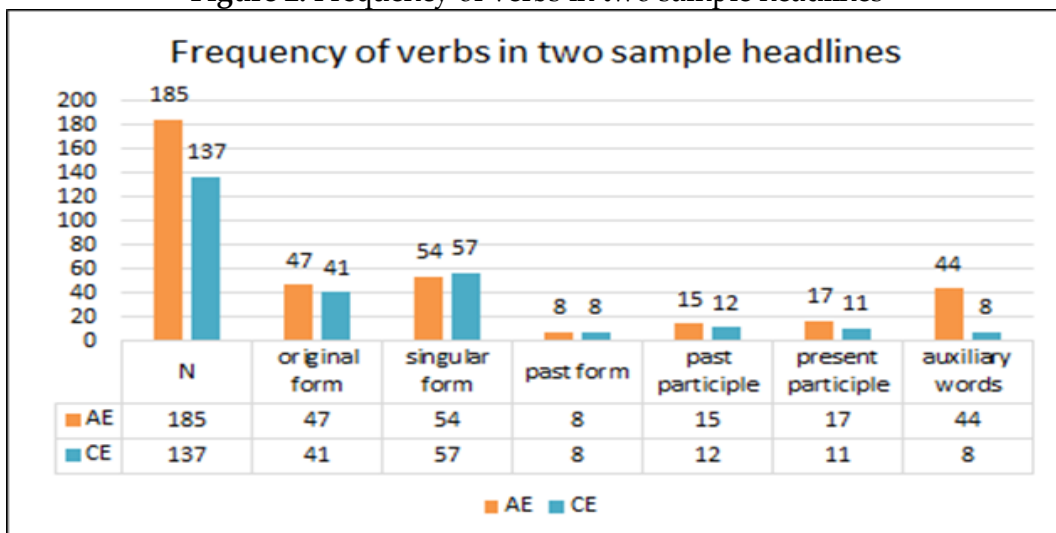
CD.84: *CPC meeting discusses government work report*

In headline CD.15, *PLA* stands for People’s Liberation Army and *CPC* in CD.84 refers to Communist Party of China.

### 5.2.2 The Use of Verbs

Figure 2 illustrates the frequency of different forms of verbs in two sample headlines.

Figure 2: Frequency of verbs in two sample headlines



In the CE sample, there are 137 verbs, in which 8 auxiliary words in total. The most predominant form of verb is singular form. Details are shown in the above figure. Three headlines containing auxiliary words would be given as examples here:

CD.34: *Grading will aid choices of nursing homes*

CD.48: *US confirms its envoy has met with Venezuelan officials*

CD.94: *Maduro says he is considering closing Venezuela border with Colombia*

In headline CD.34, the modal auxiliary verb ‘will’ refer to possible future event. The ‘has’ in CD. 48 refers to present perfect tense and the ‘is’ in CD. 94 represents present progressive tense.

Three headlines containing no verb would be listed here:

CD.73: *Trump optimistic about upcoming summit with Kim.*

CD.90: *Former official with China’s top court under investigation*

CD.96: *Canadian provinces still eager for China’s business, tourism*

In the AE sample, there are 185 verbs in total. Same as the CE sample headlines, verbs in singular form rank the top, the number of which slightly outweighs that of verbs in original form. The use of auxiliary words needs to be noted because quite a number of AE headlines use auxiliary words, either to indicate tense, or to start an interrogative.

Two example headlines of using auxiliary words are listed as follow:

NYT.16: *When Does Kicking Black People Off Juries Cross a Constitutional Line?*

NTY.38: *Trump Has Publicly Attacked the Russia Investigation More Than 1,100 Times*

In headline NYT.16, auxiliary verb “does” is used to form a question and in NYT.38, auxiliary verb “has” is the symbol of perfect aspect.

There are also two headlines that have no verb, whatever form it is. The two headlines would be given here.

NYT.56: *In Beijing, a Communist Funeral for an Inconvenient Critic*

NYT.72: *A Victory for the Civil War ‘Cyclorama’*

As illustrated in the above figure concerning frequency of verbs in two sample headlines, several differences can be attested: first, the total number of verbs used in AE headlines far outnumbers that in the CE headlines. Second, as the most predominant verb form in both

samples, verb in singular form in CE headlines is the only category whose number outweighs the AE counterparts. Thirdly, in terms of the use of auxiliary words, the number in AE sample is almost 5 times that of the CE sample. Possible reason can be that the CE headline writers try to avoid using modal verbs such as “could” or “may” to describe possibility. Instead, they are more willing to use more concrete verbs to describe events.

### 5.2.3 The Use of Prepositions

Among all the 100 CE sample headlines, there are totally 136 prepositions with an average number of 1.36 preposition in each headline. However, this does not mean that there is a preposition in every headline because 14 out of 100 headlines have no preposition. The largest number of prepositions that appears in one headline is 4, and there are two cases:

CD.08: *Interest in air travel perks up in response to super moon*

CD.56: *Putin says preliminary results of national projects to be summed up in 2020*

In example CD.08, the 4 prepositions are *in*, *up*, *in* and *to*. In example 56, the 4 prepositions are *of*, *to*, *up* and *in*.

Among all the 100 AE sample headlines, there are totally 143 prepositions, meaning that the average number of prepositions in each headline is 1.43. 16 headlines employ no preposition. As the case in the CE data, the largest number of prepositions in one headline is 4, and there are also 2 cases:

NYT.18: *Anthony Weiner Is Out of Prison and In a Halfway House In New York.*

NYT.39: *Florida Student, 11, Arrested After Dispute Over His Refusal to Say Pledge of Allegiance*

In NYT.18, the prepositions are *out*, *of*, *in* and *in*. The prepositions in NYT.39 are *after*, *over*, *to* and *of*.

### 5.2.4 The Use of Conjunctions

Another apparent distinction in terms of lexicology in the two samples of data lies in the use of conjunction. Although there are totally 21 conjunctions in the AE sample, the frequency of conjunction is only 4 in the CE counterpart. Two examples would be listed as follows:

CD.07: *As Tianjin develops, it preserves old styles*

CD.10: *Domestic workers look for better status and conditions*

As shown in CD.07, *as* acts as the conjunction, indicating the meaning “during the time that...”. In CD.10, *and* is obviously the conjunction which is used to connect the two words, *status* and *conditions*.

Compared to using conjunctions, punctuations such as comma present more in CE sample headlines. As here in the following examples:

CD.77: *Peru hotel robbed, no Chinese injury*

CD.81: *State Council appoints, removes officials*

In the AE sample, there is more multiple use of conjunction. Two examples are listed here to better illustrate the situation:

NYT.17: *Obama Quietly Gives Advice to 2020 Democrats, but No Endorsement*

NYT.37: *Is 2019 the Year When New York Passes Congestion Pricing?*

Therefore, in the AE sample, not only the coordinate conjunction such as *but* is used, but also the subordinators including *when* is applied to connect the sentence.

### 5.3 Structure Pattern

In this section, a revised version of Lee's (2012) structure pattern would be adopted here to categorize the sentence structure in the current data. Here are the reasons for employing Lee's (2012) patterns:

*“First, the current study and Lee’s study contain the same amount of headlines. Second, both studies aim at headlines of online political news. Third, the two studies share similar research settings. Lee’s (2012) sample data come from two parts, one is the news headlines collected from American news websites including Washington Post, and etc. The other is the English headlines collected in South Korea online English news websites. In the present research, two sets of sample headlines were also chosen. 50% of the headlines are collected from the American quality English newspaper online The New York Times and the rest 50% from the Chinese leading English news platform China Daily.*

*Here is Lee’s (2012) structure pattern of categorization:*

- a. NP+ to V
- b. Vless
- c. NP
- d. Simple S
- e. Complex S
- f. *Interrogative*” (extracted from (Lee 2012: 334))

NP refers to noun phrase, NP+to V refers to a noun phrase plus an infinite verb or verb phrase. Vless means that there is no predicate in the sentence. To look into the linguistic complexity of a text, it is necessary to look at whether the sentence in the text is simple or complex because as Verspoor & Sauter (2000: 35) noted, *“sentences may have different degrees of complexity”*. Verspoor & Sauter (2000: 35) define simple sentence as a sentence consists of *“one main clause*

only". Complex sentence, on the other hand, can be defined as "a sentence that contains at least one full dependent clause with its own subject and predicate" (Verspoor & Sauter 2000: 37). They further noted that "A dependent clause is a clause that starts with a subordinator, a word like because, although, if, who, where, when, that and so on" (Verspoor & Sauter 2000: 37).

The structure pattern used in the current research is revised from Lee's (2012) research, with certain changes to the old patterns and new category added. The "Vless" in Lee's (2012) is here replaced by "None predicate". This category would include headlines that have no verbs as predicate, hence, the Noun Phrase in Lee's (2012) would be put under "None predicate" as there are not much occurrence of purely noun phrase as headlines in two sets of data. Although Verspoor & Sauter (2000: 37) explained that in a compound sentence, both parts are simple and independent sentences, while in a complex sentence the dependent clause cannot stand on its own and functions as a constituent of the main clause, Lee (2012) categorized compound under complex sentence. Therefore, in the current study, compound sentences are also analyzed as a subcategory of complex sentence.

What is more, in the present study, some headlines contain not only one sentence. One headline can contain two simple sentences or one simple sentence and an interrogative (in this case there is no conjunction between two sentences but punctuations including full stop and question mark). Hence, a new category "Combination" will be put into the structure patterns. Also, when the headline contains a colon, it will also be categorized into this pattern. Here are two examples:

NYT.92: *The Vatican Is Talking About Clerical Abuse, but Italy isn't. Here's why.*

NYT.93: *ISIS Cases Raise a Question: What Does It Mean to Be Stateless?*

The newly revised structure patterns used to categorize the 200 headlines would be:

- a. NP+ to V.
- b. None predicate
- c. Simple S
- d. Complex S
- e. Interrogative
- f. Combination

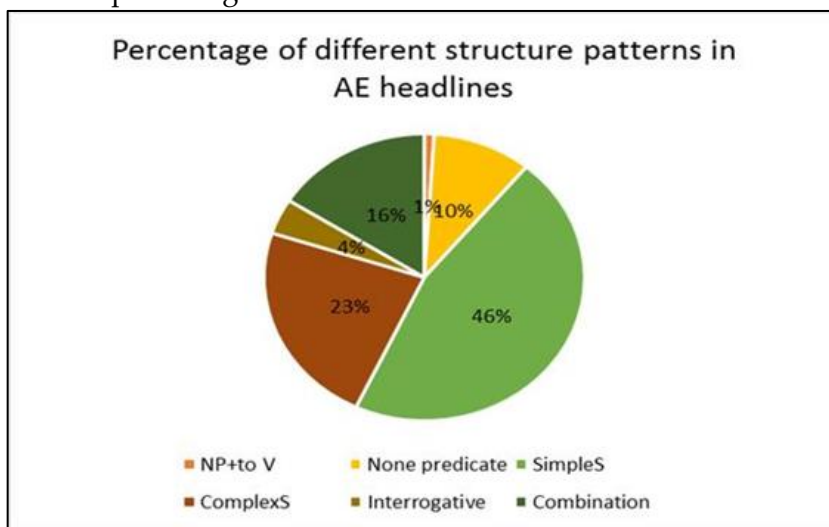
### 5.3.1 Structure Patterns in Sample Headlines

**Table 4:** The number of different sentence constructions in AE headlines

NP+to V		None predicate	SimpleS	ComplexS	Interrogative	Combination
AE	1	10	46	23	4	16



**Figure 3:** The percentage of different sentence constructions in AE headlines



As shown in the above table and figure, it is apparent that simple sentence takes the first place concerning structure pattern in the AE sample. With a proportion of 23%, complex sentence ranks the second. 16 headlines are a combination of two sentences. Also, there are 4 cases of interrogatives in the 100 AE headlines. Here one example would be listed:

NYT.16: *When Does Kicking Black People Off Juries Cross a Constitutional Line?*

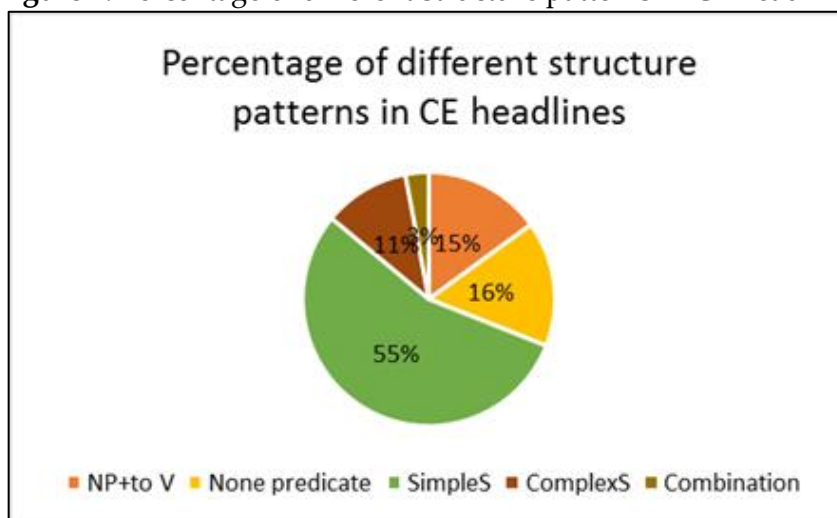
There is only one case of using “NP+ to V” pattern in AE sample:

NYT.88: *Indonesian Art Collective to Curate Next Documenta Exhibition*

**Table 5:** The number of different sentence constructions in CE headlines

	NP+to V	None predicate	SimpleS	ComplexS	Interrogative	Combination
CE	15	16	55	11	0	3

**Figure 4:** Percentage of different structure patterns in CE headlines

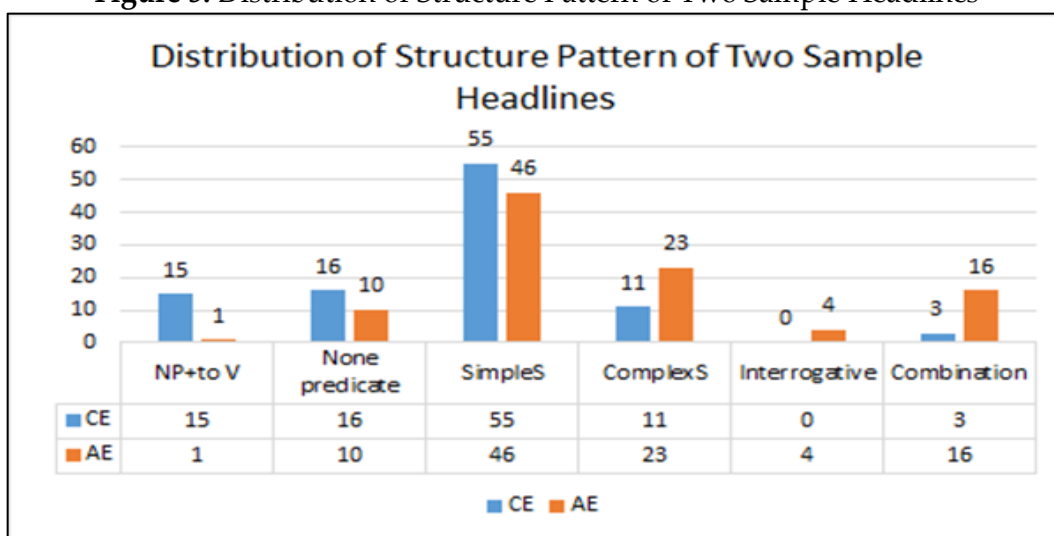


The above table and figure show the distribution of different structure patterns in CE headlines. Simple Sentence still ranks the top, followed by None Predicate which occupies 16% and NP + to V takes the third place. Example of None Predicate would be CD.45 and NP+ to V CD.05.

CD.45: *Japan set to host Trump a second time*

CD.05: *Supermoon to light up China's Lantern Festival*

**Figure 5:** Distribution of Structure Pattern of Two Sample Headlines



As illustrated in the above figure, one similarity is that simple sentence ranks the top in terms of structure patterns in CE and AE headlines, with 55 and 46 cases respectively. There is also not much difference in None Predicate headlines. However, the distinctions with regard to other structure patterns cannot be neglected. The CE sample uses far more NP + to V. structure than the AE counterpart, certain reason can be that the CE headline writers hope that this structure can make the headlines short and brief. There is no interrogative in CE headlines. Reason can be that instead of using an interrogative with the question mark at the end of a headline to raise a question, CE headline writers prefer to tell the readers directly about past or upcoming events. The number of complex sentences in AE sample is double that of the CE sample. Interrogative does not appear in the CE sample but there are 4 headlines of this structure in the AE sample. When it comes to the Combination, although there are 3 headlines of this structure in the CE sample, all of them appear to be the use of colon. For example:

CD.61: *Cuba, China have much to do together for a better world: Cuban diplomat*

However, in the AE sample, apart from using colon, 7 of the Combination structure headlines contain one declarative and one interrogative. Example:

NYT.83: *They left Canada for ISIS. Should They Be Allowed Home?*

### 5.3.2 The Use of Quotation

The use of quotation in both samples is worth noting because ways in which quoting is done vary in the two samples apparently.

Brendel, Meibauer and Steinbach (2011: 163) categorized quotation into pure quotation, direct quotation and indirect quotation. For example:

- *“Berlin” has six letters.* (pure quotation)
- *Quine said: “Quotation has a certain anomalous feature”.* (direct quotation)
- *Quine said that quotation has a certain anomalous feature.* (indirect quotation) (extracted from Brendel, Meibauer and Steinbach 2011:163)

In the CE sample, three headlines used quoting. 4 other headlines used colon instead of using a word “say”. Now let us take a deeper look at them. CD.09: *Snow-making catalyst is safe, experts say.*

CD.53: *Trump aides pushed to share nuclear tech with Saudis, report says* CD.94: *Maduro says he is considering closing Venezuela border with Colombia* CD.58: *Three lawmakers quite UK PM May’s Conservatives over ‘disastrous’ Brexit*

In the above four headlines, CD.58 is the only headline that contains a quotation mark, which, according to Brendel, Meibauer and Steinbach (2011), is a pure quotation. The rest three headlines are all indirect quotations, employing the word “say”. The difference is that in CD.94, the word “say” is at the second position of the sentence while in the other two cases, “say” is at the end of the sentence. The reason can be that in CD.94, the headline writer hopes the use of the proper noun “Maduro”, as the subject at the beginning of the sentence can attract readers’ attention.

One headline among the 4 using colon would also be listed here:

CD. 96: *EU to end trade talks if US imposes car tariffs: EU Trade Commissioner*

Here the word “say” is replaced by a colon, which indicates the information source. The frequency of quotation in AE sample headlines is more than that in the CE data. In AE headlines, quotation mark appears in 12 headlines, the word “say” in 9 headlines and 6 headlines employ the punctuation colon. What is more, 3 headlines use both quotation mark and colon. They would be listed as follow:

NYT.15: *Becoming Greta: ‘Invisible Girl’ to Global Climate Activist, With Bumps Along the Way*

NYT. 36: *‘Insult to China’: A Model’s Freckles Spark an Online Storm*

NYT. 77: *Scores Dead in Bangladesh Fire: ‘This Isn’t About Poverty, It’s About Greed’*

Another special headline that uses a colon is an interrogative sentence.

NYT 99: *ISIS Cases Raise a Question: What Does It Mean to Be Stateless?*

Two examples with the word “say” will be listed here:

NYT.47: *Oil Companies May Be Complicit in Atrocities in South Sudan, U.N. Panel Says*

NYT.45: *Alabama Woman Who Joined ISIS Can’t Return Home, U.S. Says*

As shown in NYT.45 and NYT.47, the word “say” is at the end of the sentence. In one case this word is at the second position of the headline:

NYT.98: *Kamala Harris Says Jussie Smollett Developments Are Disappointing*

There are also other headlines containing the word “say”. But instead of acting as the symbol of quoting, “say” refers means to express a thought or opinion. Example is shown in NYT.07:

NYT.07: *Shinzo Abe Won’t Say If He Nominated Trump for a Nobel Prize*

### 5.3.3 Tense

**Table 6:** Distribution of the tense of CE and AE Simple Sentence Headlines

	Present	Progressive	Past	Auxiliary	Total
CE	53	0	1	1	55
AE	40	0	0	6	46

As shown in the above two tables, present tense ranks the top in simple sentence of both sets of headlines. In the CE data, there is one headline using past tense. But this situation does not occur in the AE counterpart.

CD.60: *15 people injured after ship collision in Baltic sea*

Also, the use of auxiliary words indicating different tenses is presented more in the AE data. Among the 6 AE simple sentence headlines using auxiliary words, 4 of them are in present tense. One in the past perfect tense and another in the future tense.

When it comes to the complex sentences, in the CE sample, all complex sentence headlines are crouched in the present tense. In the AE sample, however, one headline is crouched in the past tense and another in the future tense.

## 5.4 Findings from Lexical and Syntactic Complexity Analyzer

### 5.4.1 Lexical Complexity Analyzer

The Lexical Complexity Analyzer designed by Lu (2012) contains 25 measures covering lexical features including lexical density, lexical sophistication, lexical variation, TTR, verb diversity and

lexical word diversity. The table of detailed measures with coding and calculating method is presented in appendix.

**Table 7: Results of lexical sophistication**

	CE headlines	AE headlines
<b>LD (lexical Density)</b>	0.77	0.68
Lexical Sophistication		
LS1 (Lexical sophistication-I)	0.51	0.53
LS2 (Lexical sophistication-II)	0.47	0.52
VS1 (Verb sophistication-I)	0.25	0.25
VS2 (Verb sophistication-II)	7.50	7.25
CVS1 (Corrected VS1)	1.94	1.90

Lexical density (Lu 2012:191) refers to the ratio of the number of lexical words to the total number of words in a text. Lexical sophistication measures “the proportion of relatively unusual or advanced words in the learner’s text” (Read 2000: 203, cited in Lu 2012: 192). In Table 7, the ratio of lexical density of CE headlines is higher than that of the AE headlines. In terms of lexical sophistication, two indicators (VS2, CVS1) of CE headlines are also higher.

**Table 8: Results of lexical variation**

	CE headlines	AE headlines
<b>NDW</b>		
NDW (number of different words)	552	629
NDWZ (first 50 words)	44	45
NDWERZ (expected random 50)	46.20	45.40
NDWESZ (expected sequence 50)	46.50	45.40

Lexical variation, or lexical diversity, or lexical range, indicates “the range of a learner’s vocabulary as displayed in his or her language use” (Lu 2012:192). There are more numbers of different words in AE headlines, but NDWERZ and NWDESZ of CE headlines slightly higher.

**Table 9: Results of TTR**

	CE headlines	AE headlines
<b>TTR</b>		
TTR (Type/Token Ratio)	0.64	0.57
MSTTR (Mean Segmental TTR)	0.91	0.91
CTTR (Corrected TTR)	13.31	13.42
RTTR (Root TTR)	18.82	18.98
LOGTTR (Bilogarithmic TTR)	0.93	0.92
UBER (Uber Index)	44.72	38.21

Although TTR is widely used in first language and second language studies, Lu (2012) summarized from previous studies that TTR has been proved problematic because the ratio is likely to drop with the increase of sample size. Lu (2012: 193) also pointed out that to improve TTR, methods including mean segmental TTR (MSTTR), which is computed by dividing a sample into successive segments of a given length and calculating the average TTR of all segment

and other transformations of TTR such as Corrected TTR (CTTR), Root TTR (RTTR), Bilogarithmic TTR (LOGTTR) and the Uber Index are also frequently adopted.

The TTR and UBER of CE headlines are higher than AE headlines. However, CTTR and RTTR of CE headlines are slightly lower.

**Table 10:** Results of verbal diversity

	CE headlines	AE headlines
<b>Verb Diversity</b>		
VV1 (Verb variation-I)	88.41	47.21
SVV1 (Squared VV1)	6.65	4.86
CVV1 (Corrected VV1)	0.75	0.74

In terms of verb diversity, all three indicators of CE headlines are higher than that of AE headlines. Hence, with regard to verbs, CE headlines are more diversified.

**Table 11:** Results of lexical word diversity

	CE headlines	AE headlines
<b>Lexical word diversity</b>		
LV (Lexical word variation)	0.86	0.64
VV2 (Verb variation-II)	0.16	0.10
NV (Noun variation)	0.73	0.77
ADJV (Adjective variation)	0.12	0.05
ADV (Adverb variation)	0.01	0.02
MODV (Modifier variation)	0.13	0.07

As apparently shown in Table 11, NV and ADV are two of the 6 indicators in terms of lexical word diversity that CE headlines are lower than that of AE headlines.

The result of the Lexical Complexity Analyzer needs to be discussed along with the lexical findings in section 5.2. Although there are more numbers of words in AE headlines, CE headlines outnumber AE headlines in 15 out of 25 indicators in terms of lexical complexity. Also, while a larger number of verbs are used in AE headlines, the verbs in CE headlines are more diversified. AE adopted more nouns and the ratio of noun variation is higher than that of the CE counterparts.

#### 5.4.2 Syntactic Complexity Analyzer

Lu (2010) selected 14 indicators in the Syntactic Complexity Analyzer by classifying them into groups. Concerning the length of production units, the first group includes three indicators: mean length of sentence (MLS), mean length of T unit (MLT), and mean length of clause (MLC). The second group is about the internal structures with three subcategories: subordinating structures, coordinating structures, and coordinate phrases per clause. The third group concerns particular structures including verb phrase and complex nominal structures. A table with detailed information of the 14 indicators is presented in the appendix.

The result of the current study would also be presented here in forms of tables.

**Table 12:** Length comparison

	CE headlines	AE headlines
MLS (mean length of sentence)	8.5700	10.3654
MLT (mean length of T unit)	9.7386	12.8333
MLC (mean length of clause)	9.9651	12.3908

Wolfe-Quintero et al (1998) indicated that MLT and MLC can determine the syntactic development in L2 writing and as stated by Lu (2010), MLC is the best measure to assess L2 writing proficiency. As shown in Table 12, the MLS, MLT and MLC of CE headlines are all lower than the AE headlines. Therefore, it could be referred that CE headline writers, as L2 users, tend to produce shorter sentences and shorter clauses than the AE headline writers. This is also in consistent with the result using my own method which shows that the average length of CE headlines is shorter than the AE headlines.

Eight measures are categorized into three groups by Lu (2010) in measuring internal structure. The number of clauses per sentence (C/S) belongs to the first group sentence complexity. The second group is about subordinating structures, including clauses per T-unit (C/T), dependent clauses per clause (DC/C), dependent clauses per T-unit (DC/T) and complex T-unit ratio (CT/T). The third group concerns coordinating structures, including T-units per sentence (T/S), coordinate phrases per T-unit (CP/T), and coordinate phrases per clause (CP/C).

**Table 13:** Subordinate or coordinate syntactic complexity comparison

	CE headlines	AE headlines
C/S (the number of clauses per sentence)	0.8600	0.8365
C/T (clauses per T-unit)	0.9773	1.0357
DC/C (dependent clauses per clause)	0.0689	0.0920
DC/T (dependent clauses per T-unit)	0.0682	0.0920
CT/T (complex T-unit ratio)	0.0682	0.0833
T/S (T-units per sentence)	0.8800	0.8077
CP/T (coordinate phrases per T-unit)	0.0114	0.1429
CP/C (coordinate phrases per clause)	0.0116	0.1379

Concerning the comparison of subordinate or coordinate syntactic complexity, there are more C/S in the CE headlines. However, AE headlines use more C/T, DC/C, DC/T and CT/T, indicating that CE headline writers use fewer subordinating structure. Regarding to coordinating structures, although CE headlines use more in the number of T/S than AE headlines, there are more CP/T and CP/C in the AE headlines. The results in Table 13 are, to some degree, in line with previous research on EFL Chinese students. Xu et al (2013) stated that as language proficiency increases, the L2 users are likely to apply more clauses per T unit in their writings. Therefore, if the CE headline writers are evaluated as intermediate in terms of language proficiency and their AE counterparts as proficient in language use, then this study follows that the proficient language users would adopt more C/T and DC/C.

**Table 14:** The particular structure comparison

	CE headlines	AE headlines
VP/T (Verb phrases per T-unit)	1.4432	1.2619
CN/T (complex nominals per T unit)	1.4545	1.2857
CN/C (complex nominals per clause)	1.4884	1.2414

Verb phrases per T-unit (VP/T), complex nominals per T unit (CN/T) and complex nominal per clause (CN/C) are classified as particular structures. Lu (2011) defined that complex nominals include nouns plus adjective, possessive, prepositional phrase, adjective clause, participle, or appositive; nominal clauses; gerunds and infinitives in subject but not object position.

As shown in Table 14, the figures of the three particular structures in CE headlines all outnumber that of the AE counterparts. As noted by McNamara et al (2010) that the complexity of verb phrases could indicate writing quality and Lu (2010) also pointed out that complex nominal (CN) structures demonstrate writing performance, a brief conclusion here is that CE headlines tend to be more complex in terms of verb phrases and complex nominals.

## 6. Results and Discussion

With regard to the lexical level, more nouns and verbs are used in the AE sample headlines and auxiliary words are more preferred in the AE data. The AE sample headlines used slightly more prepositions than the CE headlines. Contradictory to Hu & Guo's (2013) research that there are more statistics in Chinese sports news headlines, there is not much distinction in terms of using statistics in two sets of data in the current study. Possible reason can be that readers expect the result of a competition when reading a sports news article but when it comes to political news, statistic plays a less important role. Conjunction presents an apparent distinction because only 4 out of 100 CE headlines employed conjunctions while in the AE headlines, the occurrence of conjunctions is 5 times that of its CE counterpart. Among the 25 measurements in the Lexical Complexity Analyzer, 15 indices of CE headlines are higher than the AE headlines. Therefore, it can be concluded that regardless of being inferior in terms of the number of words, CE headlines tend to be more complex concerning lexical density, type/token ratio, verb sophistication and verb diversity.

In terms of syntax, data analysis from my own method shows that AE headlines tend to be more complex by adopting more complex sentences, interrogatives, combination sentences and omitting predicates. AE headlines use more quotations to "attain readers' attention" (Lee 2012: 337) and there are more auxiliary words entertained in AE headlines to refer to different tenses. Among the 14 measurements in the Syntactic Complexity Analyzer, CE headlines outnumber AE headlines in 5 indices, indicating that AE headlines tend to be more complex in terms of length, subordinating and coordinating structures. However, in terms of verb phrases per T-unit and complex nominal, CE headlines show more complexity.

Compared with previous research, both similarities and differences are attested. Current results in the present study are in parallel with Wang's (2016) research who found that EFL Chinese learners tend to produce shorter sentences and clauses. In Lee's (2012) research, KE headlines employ NP + to V and verb-less constructions more than the AE headlines. In the current study, although the verb-less construction is adapted into "None predicate", there are



indeed more NP + to V and None predicate constructions in CE sample headlines. Another similarity that accords with Lee's (2012) research is the use of tenses. Lee's (2012) study shows that KE headlines employ fewer varieties of tenses. After analyzing CE and AE headlines, I also find out that the latter presents more varieties of tenses with the use of auxiliary words. New finding is that in the present study, there appears to be more sentence structures in AE headlines compared to Lee's (2012). In the current data, there are several cases that one headline contains two or even more sentences connected by a full stop or a question mark. This makes the AE headlines more complex than the CE headlines in this respect. Another new finding is that in the current study, headlines that are purely noun phrases scarcely appear either in AE data or in CE data. However, in Lee's (2012) research, noun phrases as headline appear 4 times in the AE data and 6 times in KE data.

The reason to the difference of linguistic complexity between AE and CE headlines lies in not only the language proficiency of the headline writers but also the target readers. Although more complex syntax, greater lexical diversity and less frequent words indicate more sophisticated language, it might be difficult for readers to process because "less complex sentences, less diversity of words across the text and more familiar words will generally facilitate reading" (McNamara et al 2010: 63). The majority of readers of CE headlines are ESL/EFL Chinese learners who are interested in reading news both home and abroad in English language or may use the news articles as learning sources. English language teachers may also use these articles as teaching materials. Thus, CE headlines, to some degree, should be easier for the readers to understand.

## 7. Conclusion

The paper contrasted 200 headlines collected from American English and Chinese English news websites, aiming to explore the linguistic complexity between the two sets of data at lexical and structural levels. The results shed more light on the linguistic features of headlines because the 200 headlines are analyzed first using my own method, then by adopting computational systems as a supplement.

Despite the findings summarized in the above section, the present study is still limited and requires improvement for further investigation.

First, more data ought to be collected. The current study contains only 200 political news headlines within five days' period. Thus, there could be distinctions in the results if headlines covering other topics such as business and entertainment are taken into consideration. Therefore, the dataset of the study should be enlarged and a longitudinal method should be adopted in order to ensure the accuracy and sophistication of the results.

Second, the computational systems, namely the Lexical Complexity Analyzer and Syntactic Complexity Analyzer are designed originally for assessing L2 learners' writing proficiency. In our study, however, half of the 200 headlines are written by native speakers of English, which, to some extent, contribute to the lack of persuasiveness of the results generated automatically from the systems.

Thirdly, although the Lexical Complexity Analyzer and Syntactic Complexity Analyzer have been proved to be adoptable and useful in examining writing quality, this does not mean

that these two systems are perfect. Noun phrases, adjective phrases, noun clauses and attributive clauses are included in the complex nominal structures by Lu (2011), however, as Wang (2016: 85) stated “it is still yet impossible to conclude whether the phrase structure in the complex nominal structure would be an effective measure over the traditional measure using clause or T unit”.

It is necessary for future studies to expand the analysis to a larger comparable dataset within a broader scope of time using multiple automated systems such as Coh-Metrix and the Lexical and Syntactic Complexity Analyzer. Such expansion is likely to enhance our understanding of linguistic complexity in news headlines as well as to contribute to the better use of English language for ESL/EFL teachers and students.

### **Conflict of interest statement**

I declare that I have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled, “The New York Times and China Daily Online - Linguistic Complexity in News Headlines”.

### **About the Author**

The author finished her master programme in Germany, focusing on pragmatics and applied linguistics. The author is now working as a teaching assistant in a university in China and is responsible for giving English lectures to non-English major students.

### **References**

- (2020). Retrieved 18 October 2020, from <https://www.degruyter.com/viewbooktoc/product/44514>. Date of access 01.03.2019
- 2016. Reuters Institute Digital News Report 2016. <https://reutersinstitute.politics.ox.ac.uk/our-research/digital-news-report-2016>
- Bao. 2009. “Syntactic complexity in EFL learners’ essays: A multidimensional perspective.” *Foreign Language Teaching and Research* 41 (4): 291-297
- Brendel, Elke & Meibauer, Jörg & Steinbach, Markus. 2011. “Exploring the Meaning of Quotation.” In: Brendel, Elke et al, eds. *Understanding Quotation*. Berlin/New York: de Gruyter. 161-193.
- Cambridge Dictionary*. <https://dictionary.cambridge.org/> Date of access 02.03.2019.
- Casanave, Christine. 1994. “Language Development in Students’ Journals.” *Journal of second Language Writing* 3 (3): 179-201.
- China Daily*. <http://www.chinadaily.com.cn/> Date of access 27.02.2019.
- Dijk, Teun A. Van. 1988. *News as Discourse*. New Jersey: Lawrence Erlbaum Associates.
- Ferris, Dana R. 1994. “Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency.” *TESOL Quarterly* 28 (2): 414-420.
- Gao, Lifen. 2016. “A Comparative Study of Conceptual Metaphors in English and Chinese Economic News Headlines.” *Creative Education* 7: 2629-2639.
- Halliday, M.A.K. and Colin Yallop. 2007. *Lexicology. A short Introduction*.

- <http://www.personal.psu.edu/xx113/downloads/l2sca.html> Date of access 17.03.2019.
- Hu & Guo. 2013. "A Corpus-Based Study on the Differences of Chinese and American Online English Sports News Headlines from the Perspective of Cross-Cultural Communication." *International Journal of English Linguistics* 3 (5): 87-94.
- Kleinke, Sonja. 2000. "Women and Headline-Policy in German and English Local Daily Newspapers." In: Ungerer, Friedrich, ed. *English Media Texts Past and Present*. Amsterdam/Philadelphia: John Benjamins. 67-84.
- Lee, Hye-Kyung. 2012. "Linguistic Variations between English News Headlines in the U.S. and Those in Korea." *Linguistic Research*. 29(2): 329-350
- Lewis, Diana M. 2003. "Online news, a new genre?" In: Aitchison, Jean and Lewis, Diana M, eds. *New Media Language*. London and New York: Routledge. 95-104.
- Linguistic Complexity Analyzer and Syntactic Complexity Analyzer*.  
<http://www.personal.psu.edu/xx113/downloads/l2sca.html> Date of access: 17.03.2019.
- Lu, Xiaofei. 2010. "Automatic analysis of syntactic complexity in second language writing." *International Journal of Corpus Linguistics* 15 (4): 474-496.
- Lu, Xiaofei. 2012. "The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives." *The Modern Language Journal* 96 (ii): 190-208.
- McNamara, Danielle S. et al. 2010. "Linguistic Features of Writing Quality." *Written Communication* 27 (1) 57–86.
- Napu, Novriyanto. 2018. "English and Indonesian Newspaper Headlines: A Comparative Study of Lexical Features." *European Journal of Literature, Language and Linguistics Studies* 2 (4): 103-116.
- Newman, Nic et al. 2018. Reuters Institute Digital News Report 2018. <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/digital-news-report-2018.pdf>
- Ortega, Lourdes. 2003. "Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing." *Applied Linguistics* 30 (4): 555-578.
- Quirk, Randolph et al. 1972. *A Grammar of Contemporary English*. London: Longman.
- Reah, Danuta. 1998. *The Language of Newspapers*. London and New York: Routledge.
- Schneider, Kristina. 2000. "The Emergence and Development of Headlines in British Newspaper." In: Ungerer, Friedrich, ed. *English Media Texts Past and Present*. John Benjamins: Amsterdam/Philadelphia. 45-66.
- Skehan, Peter. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- The New York Times*. <https://www.nytimes.com/> Data of access 27.02.2019.
- U.S. and Those in Korea." *Linguistic Research*. 29 (2): 329-350.
- Verspoor, Marjolijn and Kim Sauter. 2000. *English Sentence Analysis-An Introductory Course*. Amsterdam/Philadelphia: John Benjamins.
- Wang, Sue & Slater, Tammy. 2016. "Syntactic Complexity of EFL Chinese Students' Writing." *English Language and Literature Studies* 6 (1): 81-86. <http://dx.doi.org/10.5539/ells.v6n1p81>
- Wolfe-Quintero et al. 1998. *Second language development in writing: Measures of fluency, accuracy and complexity*. Honolulu, HI: University of Hawaii Press.
- Xu et al. 2013. "A study on the syntactic complexity of English essays by in Chinese students of English." *Foreign Language Teaching and Research* 45 (2): 264-320.

## Appendix

### A. List of 100 headlines from China Daily

- CD.01: Jilin to become the third province to deregulate industrial hemp cultivation
- CD.02: New media facilitates crackdown on corruption in China
- CD.03: East China province eyes large 5G coverage
- CD.04: Bureau modifies tax certificate rules
- CD.05: Supermoon to light up China's Lantern Festival
- CD.06: Over 90 pct young people in big cities find commuting inconvenient: survey
- CD.07: As Tianjin develops, it preserves old styles
- CD.08: Interest in air travel perks up in response to super moon
- CD.09: Snow-making catalyst is safe, experts say
- CD.10: Domestic workers look for better status and conditions
- CD.11: Bookstores bouncing back in internet era
- CD.12: Experts wary of shrinking population
- CD.13: Sanquan Foods moves to contain fallout
- CD.14: Tests of processed pork sold in Gansu find African swine fever
- CD.15: PLA welcomes launches of private rockets
- CD.16: Saudi crown prince visits Pakistan on ties, cooperation
- CD.17: India hosts first 'yoga on a pole' world competition
- CD.18: Thai polls cast political uncertainty
- CD.19: Hanoi prepares for Trump-Kim summit
- CD.20: Venezuelan FM meets US representative to address bilateral relationship
- CD.21: Greater Bay Area aims to become world leading center
- CD.22: China launches campaign against drunk driving
- CD.23: China's top disciplinary authorities to tighten political inspection
- CD.24: China unveils guidelines on prioritizing agricultural, rural development
- CD.25: Shenzhen to pilot 5G commercial use in 2019
- CD.26: Lantern show held in Beijing's Yanqing district
- CD.27: Hebei demolishes illegal buildings on mountain
- CD.28: Infant, children service centers set up in cities
- CD.29: Film academy revokes actor Zhai's doctorate
- CD.30: Hong Kong population grows by 0.9 pct to over 7.48 mln
- CD.31: Hong Kong society sees Greater Bay Area development plan an opportunity for complementary growth
- CD.32: Insurance reduces all risks
- CD.33: China greenlights more cancer medicines in 2018
- CD.34: Grading will aid choices of nursing homes
- CD.35: Research on treating cancer with malaria microbes ongoing
- CD.36: PLA Rocket Force video fires up Weibo
- CD.37: Myanmar forms joint committee for amending constitution
- CD.38: Pilot dead as 2 aircraft crash in India during aerobatics practice
- CD.39: Indonesia submits bid to host 2032 Olympics
- CD.40: Over 10,000 Thais verified as electoral candidates in individual constituency mode
- CD.41: China, Russia, India FMs to hold meeting in Zhejiang
- CD.42: Japan books largest trade deficit in almost 5 years in January
- CD.43: Former Malaysian deputy PM slapped with additional charge over criminal breach of trust
- CD.44: US, DPRK consider exchanging envoys
- CD.45: Japan set to host Trump a second time
- CD.46: Online food-delivery scrambling more than the restaurant industry
- CD.47: Bernie Sanders kicks off second bid for White House

- CD.48: US confirms its envoy has met with Venezuelan officials  
CD.49: Yale president seeks to expand China ties  
CD.50: Trump wants California to pay back billions for bullet train  
CD.51: Trump plans to nominate Rosen as Justice Dept No 2: Official  
CD.52: Trump signs directive in step to create US Space Force  
CD.53: Trump aides pushed to share nuclear tech with Saudis, report says  
CD.54: Sailor in iconic V-J Day kiss photo dies at 95  
CD.55: Former Trump adviser Stone ordered to appear in court over Instagram posts  
CD.56: Putin says preliminary results of national projects to be summed up in 2020  
CD.57: Britain announces drone exclusion zone over airport runways  
CD.58: Three lawmakers quit UK PM May's Conservatives over 'disastrous' Brexit  
CD.59: Former Spanish PM Rajoy to testify in Catalan separatists' trial  
CD.60: 15 people injured after ship collision in Baltic sea  
CD.61: Cuba, China have much to do together for a better world: Cuban diplomat  
CD.62: Polluters of waterway to face harsher punishments  
CD.63: Sewage exits along Yangtze River to be mapped  
CD.64: Beijing to provide delivery men with rental housing  
CD.65: Nation's logging capital goes green  
CD.66: Chinese scientists develop nanogenerator for cancer drug delivery system  
CD.67: Experts discuss ethical issues of China's neuroscience development  
CD.68: Robots employed in China's railway system during Spring Festival travel rush  
CD.69: Xi urges boost in high-tech development  
CD.70: Ningxia to ensure safe drinking water for all impoverished population  
CD.71: Beijing promises to improve air quality  
CD.72: China to severely punish Yangtze River Economic Belt polluters  
CD.73: Trump optimistic about upcoming summit with Kim  
CD.74: At least 81 dead as blaze rips through buildings in Bangladeshi capital  
CD.75: China, US kick off 7th round of high-level trade talks in Washington  
CD.76: Trump threatens tariffs on EU cars if no trade deal reached  
CD.77: Peru hotel robbed, no Chinese injury  
CD.78: Mandarin duck survey begins in urban Beijing  
CD.79: Canada PM Trudeau calms restless MPs over shock resignations  
CD.80: Death toll from Brazilian dam collapse climbs to 171  
CD.81: State Council appoints, removes officials  
CD.82: Chinese president meets Saudi crown prince  
CD.83: State Council addresses overexploitation of underground water in North China  
CD.84: CPC meeting discusses government work report  
CD.85: Eight Shanghai bakery employees found guilty of selling substandard produce  
CD.86: Beijing IPR court handles more cases in 2018  
CD.87: Beijing to build two new metro lines in 2019  
CD.88: 2,600 foreign reporters register to cover US-DPRK summit in Vietnam  
CD.89: Govt notice aids women's rights to work  
CD.90: Former official with China's top court under investigation  
CD.91: Ex-official sentenced to jail term of 20 years  
CD.92: Beijing to target hospital scalpers with facial recognition  
CD.93: Rail travel to tourist attractions becomes more popular  
CD.94: Maduro says he is considering closing Venezuela border with Colombia  
CD.95: Canadian provinces still eager for China's business, tourism  
CD.96: EU to end trade talks if US imposes car tariffs: EU Trade Commissioner  
CD.97: Ethiopian Airlines to operate all-woman flight for International Women's Day  
CD.98: ASUS strengthens presence in East Africa  
CD.99: US, China open talks on trade dispute

CD.100: Chinese language being introduced into Botswana school curriculum

### **B. Table of 100 headlines from The New York Times**

- NYT.01: Saudi Crown Prince Promises \$20 Billion in Investments for Pakistan  
NYT.02: Bonaire: Where Coral Cactus Thrive, and the Sea Soothes the Soul  
NYT.03: Kim Jong-un's 'Butler' Visits Vietnam to Prepare for Trump Meeting  
NYT.04: The Disappearing Chimney Sweeps of Paris  
NYT.05: Poland and Israel Try to Improve Ties, but History Intrudes  
NYT.06: Trump Delivers Blunt Warning to Venezuela Military Over Aid Impasse  
NYT.07: Shinzo Abe Won't Say if He Nominated Trump for a Nobel Prize  
NYT.08: Honda Confirms Plan to Leave Britain as Brexit Looms  
NYT.09: Vatican's Secret Rules for Catholic Priests Who Have Children  
NYT.10: Anti-Semitic Taunts by Yellow Vests Prompt French Soul-Searching  
NYT.11: Vatican Hopes Meeting on Child Sex Abuse Will Be a Turning Point  
NYT.12: 7 Labour Lawmakers Resign in U.K., in Rebuke of Jeremy Corbyn  
NYT.13: Kamala Harris Faces Questions About 'Democratic Socialist' Label and Jussie Smollett in N.H. Debut  
NYT.14: Elizabeth Warren's Native American Ancestry Draws a Shrug From These Voters  
NYT.15: Becoming Greta: 'Invisible Girl' to Global Climate Activist, With Bumps Along the Way  
NYT.16: When Does Kicking Black People Off Juries Cross a Constitutional Line?  
NYT.17: Obama Quietly Gives Advice to 2020 Democrats, but No Endorsement  
NYT.18: Anthony Weiner Is Out of Prison and in a Halfway House in New York  
NYT.19: Elizabeth Warren Wants a Wealth Tax. How Would That Even Work?  
NYT.20: Cancellation of Pakistan Trip Revives Debate Over Taliban's Ability to Travel  
NYT.21: Kaiser Permanente's New Medical School Will Waive Tuition for Its First 5 Classes  
NYT.22: Another Sick Migrant Dies in Border Patrol Custody in Texas  
NYT.23: Southern Baptists Announce Plans to Address Sexual Abuse  
NYT.24: 16 States Sue to Stop Trump's Use of Emergency Powers to Build Border Wall  
NYT.25: George Mendonsa, 95, Most Likely the Sailor in a Famous Photo, Dies  
NYT.26: Bearing Witness to Jim Crow in Mississippi With Uncompromising Candor  
NYT.27: Marijuana Tax in New Jersey? It Could be \$42 an Ounce  
NYT.28: Tanzania Court Convicts 'Ivory Queen' for Trafficking Elephant Tusks  
NYT.29: The Life-or-Death Struggle of Refugees Braving the Mediterranean Crossing  
NYT.30: After New York City's War with Amazon, Uber Could Be Next  
NYT.31: An Alabama Editor Urges the K.K.K. to 'Ride Again,' and His Small Town Winces  
NYT.32: West Virginia Teachers Walk Out (Again) and Score a Win in Hours  
NYT.33: Egypt Turns Back Veteran New York Times Reporter  
NYT.34: Huawei Founder Slams U.S. Charges as 'Politically Motivated'  
NYT.35: Family of Shamima Begum, ISIS Teenager, Says Britain Plans to Revoke Her Citizenship  
NYT.36: 'Insult to China': A Model's Freckles Spark an Online Storm  
NYT.37: Is 2019 the Year When New York Passes Congestion Pricing?  
NYT.38: Trump Has Publicly Attacked the Russia Investigation More Than 1,100 Times  
NYT.39: Florida Student, 11, Arrested After Dispute Over His Refusal to Say Pledge of Allegiance  
NYT.40: Intimidation, Pressure and Humiliation: Inside Trump's Two-Year War on the Investigations Encircling Him  
NYT.41: Survivors of Sexual Abuse Want Church Reform. Here's Why It Might Not Happen  
NYT.42: A U.S. Ambassador Promises a Global Push to Decriminalize Homosexuality  
NYT.43: A Staggering Exodus: Millions of Venezuelans Are Leaving the Country, on Foot  
NYT.44: Venezuela Closes Border to 3 Caribbean Islands Ahead of Aid Showdown  
NYT.45: Alabama Woman Who Joined ISIS Can't Return Home, U.S. Says  
NYT.46: Huawei Risks to Britain Can Be Blunted, U.K. Official Says, in a Rebuff to U.S.

- 
- NYT.47: Oil Companies May Be Complicit in Atrocities in South Sudan, U.N. Panel Says
- NYT.48: Putin Doubles Down on Threats Against the U.S.
- NYT.49: A European Army? The Germans and Dutch Take a Small Step
- NYT.50: 7 Children in Syrian Refugee Family Die in Fire in Canada
- NYT.51: America's U.N. Ambassador Post Is Empty. Is That a Problem?
- NYT.52: Thousands Rally Against Anti-Semitism in France
- NYT.53: Is Britain's Political System at the Breaking Point?
- NYT.54: As a 'No Deal' Brexit Looms, the Art World Prepares for the Fallout
- NYT.55: Just How Endangered Is Watchmaking?
- NYT.56: In Beijing, a Communist Funeral for an Inconvenient Critic
- NYT.57: Egypt Executes 9 Convicted of Assassinating Top Prosecutor
- NYT.58: Amid Trouble with the West, Saudi Arabia Looks East
- NYT.59: Overlooked No More: Dorothy Bolden, Who Started a Movement for Domestic Workers
- NYT.60: CNN Is Criticized for Hiring Trump Administration Aide as a Political Editor
- NYT.61: Stop Asking Women About Childbearing Status, China Tells Employers
- NYT.62: China Uses DNA to Track Its People, With the Help of American Expertise
- NYT.63: Made on the Inside, Worn on the Outside
- NYT.64: Pope Francis Opens Summit on Sexual Abuse
- NYT.65: Julie Bishop Quits Australian Politics, Adding to Exodus of Conservative Women
- NYT.66: Prosecutors Broke Law in Agreement Not to Prosecute Jeffrey Epstein, Judge Rules
- NYT.67: Chinese Girl Finds a Way Out of Tedious Homework: Make a Robot Do It
- NYT.68: Japan's Hayabusa2 Spacecraft Lands on Ryugu Asteroid
- NYT.69: Pope Opens Meeting on Clerical Sex Abuse Under Great Pressure from Victims for Change
- NYT.70: New Zealand's Former Leader Is Depicted by China as Praising Its Policies
- NYT.71: 2 Israel Centrists, Benny Gantz and Yair Lapid, Join Forces Against Netanyahu
- NYT.72: A Victory for the Civil War 'Cyclorama'
- NYT.73: Poroshenko Builds Support Among Ukrainians Abroad in Visit to U.S.
- NYT.74: Bangladesh Fire Tears Through Crowded Neighborhood, Killing 70
- NYT.75: Black History Trail Makes 200 Stops Across Massachusetts
- NYT.76: India Threatens a New Weapon Against Pakistan: Water
- NYT.77: Scores Dead in Bangladesh Fire: 'This Isn't About Poverty, It's About Greed'
- NYT.78: Seweryn Bialer, 92, Defector and Soviet Union Scholar, Dies
- NYT.79: Social Media Posts Reveal NATO Soldiers' Activities, Report Says
- NYT.80: For U.K.'s Labour, Anti-Semitism and Corbyn Are as Divisive as Brexit
- NYT.81: Terrorism Watchdog Castigates Pakistan Over Aid to Militants
- NYT.82: Facing Protests, Sudan's Leader Declares Yearlong State of Emergency
- NYT.83: They Left Canada for ISIS. Should They Be Allowed Home?
- NYT.84: Pablo Escobar's Home Is Demolished in Colombia, Along With a Painful Legacy
- NYT.85: 5 Children Die of Food Poisoning in Pakistan, Spurring Food Safety Concerns
- NYT.86: Saudi Sisters' Plan to Escape to Australia Is Halted in Hong Kong
- NYT.87: Billionaire Wins Defamation Case Against Australian Media Group
- NYT.88: Indonesian Art Collective to Curate Next Documenta Exhibition
- NYT.89: Japan Battles Worst Measles Outbreak in Years
- NYT.90: South Korea Awaits 2nd Kim-Trump Summit with Both Hope and Fear
- NYT.91: U.S. Catholics Wanted a Vatican Response on Sex Abuse. Is a New Proposal Enough?
- NYT.92: The Vatican Is Talking About Clerical Abuse, but Italy Isn't. Here's Why.
- NYT.93: ISIS Cases Raise a Question: What Does It Mean to Be Stateless?
- NYT.94: Why a Tax Cut Might Not Mean a Bigger Refund
- NYT.95: Trump Administration Blocks Funds for Planned Parenthood and Others Over Abortion Referrals
- NYT.96: Hate Crime Hoaxes Are Rare, but Can Be 'Devastating'
- NYT.97: Jeffrey Skilling, Former Enron Chief, Released After 12 Years in Prison
- NYT.98: Kamala Harris Says Jussie Smollett Developments Are Disappointing

NYT.99: Christopher Hasson, Coast Guard Officer, Plotted Attacks at His Desk, Filings Say

NYT.100: Peace Treaty, and Peace Prize, for North Korea Appear to Tempt Trump

**Table 1:** Measures of Lexical Complexity System

Measures of Lexical Density and Sophistication

Measure	Code	Formula	Examples
Lexical Density	LD	$N_{lex}/N$	Engber (1995)
Lexical Sophistication-I	LS1	$N_{stlex}/N_{lex}$	Linnarud (1986), Hyltenstam (1988)
Lexical Sophistication-II	LS2	$T_s/T$	Laufer (1994)
Verb Sophistication-I	VS1	$T_{sverb}/N_{verb}$	Harley & King (1989)
Corrected VS1	CVS1	$T_{sverb}/\sqrt{2N_{verb}}$	Wolfe-Quintero et al. (1998)
Verb Sophistication-II	VS2	$T_{sverb}^2/N_{verb}$	Chaudron & Parker (1990)

Measures of Lexical Variation

Measure	Code	Formula
Number of Different Words	NDW	$T$
NDW (first 50 words)	NDW-50	$T$ in the first 50 words of sample
NDW (expected random 50)	NDW-ER50	Mean $T$ of 10 random 50-word samples
NDW (expected sequence 50)	NDW-ES50	Mean $T$ of 10 random 50-word sequences
Type-Token Ratio	TTR	$T/N$
Mean Segmental TTR (50)	MSTTR-50	Mean TTR of all 50-word segments
Corrected TTR	CTTR	$T/\sqrt{2N}$
Root TTR	RTTR	$T/\sqrt{N}$
Bilogarithmic TTR	LogTTR	$\text{Log}T/\text{Log}N$
Uber Index	Uber	$\text{Log}^2 N/\text{Log}(N/T)$
D Measure	D	Based on $D$ in Equation (1)
Lexical Word Variation	LV	$T_{lex}/N_{lex}$
Verb Variation-I	VV1	$T_{verb}/N_{verb}$
Squared VV1	SVV1	$T_{verb}^2/N_{verb}$
Corrected VV1	CVV1	$T_{verb}/\sqrt{2N_{verb}}$
Verb Variation-II	VV2	$T_{verb}/N_{lex}$
Noun Variation	NV	$T_{noun}/N_{lex}$
Adjective Variation	AdjV	$T_{adj}/N_{lex}$
Adverb Variation	AdvV	$T_{adv}/N_{lex}$
Modifier Variation	ModV	$(T_{adj} + T_{adv})/N_{lex}$

Table extracted from Lu 2012: 193, 195)

**Table 2:** Measures of Syntactic Complexity Analyzer

Measure	Code	Definition
<b>Type 1: Length of production unit</b>		
Mean length of clause	MLC	# of words / # of clauses
Mean length of sentence	MLS	# of words / # of sentences
Mean length of T-unit	MLT	# of words / # of T-units
<b>Type 2: Sentence complexity</b>		
Sentence complexity ratio	C/S	# of clauses / # of sentences
<b>Type 3: Subordination</b>		
T-unit complexity ratio	C/T	# of clauses / # of T-units
Complex T-unit ratio	CT/T	# of complex T-units / # of T-units
Dependent clause ratio	DC/C	# of dependent clauses / # of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses / # of T-units
<b>Type 4: Coordination</b>		



Coordinate phrases per clause	CP/C	# of coordinate phrases / # of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases / # of T-units
Sentence coordination ratio	T/S	# of T-units / # of sentences
<b>Type 5: Particular structures</b>		
Complex nominals per clause	CN/C	# of complex nominals / # of clauses
Complex nominals per T-unit	CN/T	# of complex nominals / # of T-un
Verb phrases per T-unit	VP/T	# of verb phrases / # of T-units

Table extracted from Lu 2012: 193, 195)

Creative Commons licensing terms

Author(s) will retain the copyright of their published articles agreeing that a Creative Commons Attribution 4.0 International License (CC BY 4.0) terms will be applied to their work. Under the terms of this license, no permission is required from the author(s) or publisher for members of the community to copy, distribute, transmit or adapt the article content, providing a proper, prominent and unambiguous attribution to the authors in a manner that makes clear that the materials are being reused under permission of a Creative Commons License. Views, opinions and conclusions expressed in this research article are views, opinions and conclusions of the author(s). and European Journal of Literature, Language and Linguistics Studies shall not be responsible or answerable for any loss, damage or liability caused in relation to/arising out of conflicts of interest, copyright violations and inappropriate or inaccurate use of any kind content related or integrated into the research work. All the published works are meeting the Open Access Publishing requirements and can be freely accessed, shared, modified, distributed and used in educational, commercial and non-commercial purposes under a [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).