Diachronic perspectives of language economy in headlinese

Peter Bojo – Daniel Lancaric
DOI: 10.18355/XL.2021.14.02.22

Abstract
The language of newspaper headlines is traditionally marked by a number of lexical and grammatical devices which are used to reach the overall economy of language. The present article is a review of the occurrence and overall behaviour of these lexical and grammatical devices in the course of time. The research is conducted as a diachronic quantitative analysis. The aim of the analysis is to discover whether the distribution of lexical and grammatical devices of language economy is equal or if one prevails over the other in the course of time. Further, the analysis is designed to determine whether the occurrence of the lexical and grammatical devices is random, or if there is a direct correlation between the time and the number of devices applied. In attempt to map the overall behaviour of lexical and grammatical economy devices, two hypotheses are tested: 1: There is a significant difference in the distribution of lexical and grammatical devices in the course of time. 2: There is a positive correlation (a growing tendency) between the time axis and the distribution of lexical and grammatical devices. This article is written as a part of a VEGA project named VEGA project no. 1/0769/20. “The Language of Electronic Media in the English Linguocultural Context”.

Key words: headline, language economy, lexis, syntax, ellipsis, condenser

1 Introduction
Journalism is closely related to language and using language suitably in order to deliver the message. The message needs to be delivered in a way that appeals to the reader and encourages them to read through the rest of the article. A number of linguistic devices are applied and arranged in a desirable way to serve this purpose. Space-saving is quite significant in news language since a maximum amount of information needs to be squeezed into a minimum amount of space. Quite mistakenly, the language of newspapers is considered time-saving. It may save time on the side of the writer. Nonetheless, it requires a lot more time on the side of the reader to decode the message. The question of linguistic opacity in headlines is also determined by the readership. Defining the readership profile may be decisive when it comes to the choice of linguistic means (Shaeda, 2011:10). According to (Shaeda, 2011:8), speech accommodation theories have demonstrated that addressers accommodate speech to that of the audience targeted. Headlines are not an exception, as they are designed to address a broadly identified community of readers, and so, they use appropriate language means.

Due to space-saving measures, newspaper language, and headlines, in particular, is governed by a number of shortened lexical and grammatical structures. These shortened utterances may be treated under the principle of least effort, the concept which was first introduced by the French philosopher G. Ferrero in 1894. Ferrero discussed this principle in his article “L’inertie mentale et la loi du moindre effort.” Later, the principle was adopted by G. K. Zipf, an American philologist at Harvard University in his book “Human behaviour and the principle of least effort.” (1966). Zipf theorised that people attempted to communicate the maximum of their ideas effectively with the least effort. This principle is also known as Zipf’s Law. Humans naturally want the greatest outcome in the least amount of work (Zhu et al., 2018:2-3). Currently, this principle is one of the fundamental linguistic principles which operates on all levels of language (Šípošová, Bojo, 2016:2). The principle of
least effort was subsequently applied to various branches of science, e.g. the field of information retrieval, the field of human behavior, the field of animal behavior, etc. There is probably no better theory which would reflect natural human behavior, including language behavior. The principle of least effort may be clearly traced in the language of newspapers and newspaper headlines as well.

Shortened lexical forms used in headlines may be categorized within the domain of minor word-formation. Not all minor word-formation processes are shortenings (e.g. reduplication), but a great majority of them are dominated by abbreviating of originally longer expressions. Minor word-formative processes will be used in this research as the markers of language economy at the lexical level. The second category to be examined are grammatically (syntactically) shortened forms which, in this research, will be represented by ellipses and non-finite sentence condensers. The question remains whether the language of newspaper headlines has always been marked by these lexical and grammatical devices, i.e. the tendency towards language economy has always been present or it has developed gradually in the course of time. Another question is whether the overall economy in newspaper headlines is met equally at the lexical and grammatical level or whether one of them significantly prevailed at a certain stage in the history of news writing. To answer these questions, the present research will be conducted as a diachronic analysis of the language of newspaper headlines. Prior to the research itself, the term “diachrony” needs to be explained. The distinction between the synchronic and the diachronic approach was first presented by Ferdinand de Saussure in his Course in General Linguistics (1916). Saussure is considered the father of structural linguistics. The notion “structural” means that he understands language as consisting of strings of linguistic objects which are compared and contrasted with other objects within the system. Thus, these objects cannot be viewed in isolation. A synchronic investigation of language deals with a language or its parts (the objects) studied as a fixed phenomenon existing at a given moment of time, whereas a diachronic investigation of language concerns the development of language or its parts over time (Raclavský, 2014:2). Any linguistic problem, if studied at a particular period of time, regardless of its past or future development, is synchronic. If studied across stages of time, it is diachronic. Since this research outlines the situation in headline writing from 1900 to the present, it is diachronic.

Research aim

The aim of the research is to find out whether the two types of the devices of language economy (lexical and grammatical) are distributed in the course of time equally or whether one of the types significantly prevails. Moreover, the research is aimed at identifying a possible direct correlation between time and the amount of lexical and grammatical devices applied (attempts to determine whether the devices of language economy are random or if they have been increasingly applied across the time axis as time has approached the present stage.

Summary of research questions and hypotheses

1. Is there a significant difference in the distribution of grammatical and lexical devices of language economy over time?

H0: Newspaper headlines are supplied equally by the grammatical and lexical devices, irrespective of the time (represented by the decades from 1900 – until the present).
H1: There is a significant difference in the distribution of lexical and grammatical devices in some periods across the time axis.

This hypothesis tests the level of productivity of the two types (lexical and grammatical devices) in the course of time in an effort to determine whether the lexical and grammatical devices are equally productive (equally important for reaching a level of economy in headline language) or whether one of the two types significantly prevails.

2. Is there a positive correlation (growing tendency) between the time and the distribution of devices of language economy?

H0: The distribution of grammatical and lexical devices is random, independent of the time periods (the decades).

H1: There is a positive correlation (growing tendency) between the time axis represented by the decades and the distribution of grammatical and lexical devices. This hypothesis attempts to determine whether the devices of language economy were abundantly applied to newspaper headlines even in earlier stages and if their occurrence is random or if there is a gradual growing tendency of applying these devices across the time axis, possibly reflecting the fact that the earliest headlines did not use these devices at all, whereas the recent perspectives of the economy in language behavior may have caused an increase in their usage.

In the following chapters, the theoretical framework concerning the typology and the features of lexical and grammatical devices of language economy will be introduced. Based on the framework, a quantitative diachronic analysis will be conducted. Two hypotheses will be formulated and tested statistically.

2 Review of literature

Newspaper language has a rich history. A detailed account of the history of newspaper and headline writing was provided by Morison (1932) and Straumann (1935). As (Morison, 1932:5) puts it, the first printed newspapers are believed to have been published in China. The history of news writing and publishing newspapers in Europe goes back to the early Roman times. In the period of the Roman Empire, political events and happenings were published which gave birth to the first newspaper (ibid.). The so-called modern history of newspapers goes back to Guttenberg’s time. With the invention of printing, the first pieces ever published were pamphlets. In the 16th century, the first newspapers called “avisî” were published in Venice, Italy. Avisî portrayed the current social and political affairs, they were not published on a regular basis. Regular printing in Europe did not start before 17th century.

Since the very early beginning, newspaper articles were introduced by forerunners of what is now called headlines to capture attention and readership. According to (Morison, 1932: 4-7), the first newspapers to have been known in Britain were weekly newspapers published in the 1920s. (Straumann, 1935:87) introduced the idea that the forerunner of present newspapers was the 16th-century Ballad News. In this newspaper, the first headlines appeared which were referred to as headings - the introductory sentences which provided the reader with the summary of the article. The first real headline appeared in a newspaper in 1622 (Straumann, 1935:87).

To gain readership, the language of headlines obtained a degree of obscurity. It was this obscurity that quickly caused the language of headlines to be in the focus of linguists all over the world. Consequently, the newspaper language
developed into an independent genre commonly referred to as journalese, the jargon which abundantly uses block language and is quite frequently governed by lexical and grammatical irregularities. Webster’s dictionary (https://www.merriam-webster.com) names these irregularities colloquialisms. It refers to journalese as English of a style featuring the use of colloquialisms, superficiality of thought or reasoning, clever or sensational presentation of material, and evidence of haste in composition, considered characteristic of newspaper writing. Journalese is further defined as a hyperbolic language, sometimes saturated with a great number of abbreviations. In correspondence with journalese, the sub-genre of newspaper headlines is referred to as headlinese. This term first appeared in a handbook called Headlines Deadlines were written in 1933 by two journalists, R. E. Garst and T. M. Berstein. Headlinese is often perceived as an inferior genre because of elliptical headlines which are often typical of British tabloid press (Shaeda, 2011:7).

Straumann was one of the pioneering linguists who elaborated on the structure of newspaper headlines (Straumann, H.: (1935) Newspaper Headlines: A study in Linguistic Method), and described the patterns and irregularities of their writing. Straumann introduced the idea of “block” language, i.e. noun strings without / or with reduced verb phrases and with no functional words (conjunctions, articles, etc.), which he refers to as expanding elements. Straumann identified the language of headlines as a separate variety. Due to the absence of the predicate part, the “block language” structures are traditionally treated as non-sentential headlines. Unlike non-sentential headlines, some headlines may be treated within the regular SVO structure, which are referred to as sentential.

A deeper linguistic interest in headlines goes back to the 1960s, the time when headlines had already been assigned their obscurity. Moreover, the idea to study the language of journalism was introduced as a result of this language not being simplistic and transparent (Zelizer, 2012:111). The increasing popularity of headline language and its diversion from the regular gave rise to numerous linguistic theories, classifications and functions of headlines… etc. established by linguists worldwide, such as Bremner (1972); Mardh (1980); Howard (2000); Dor (2003); Hargreaves (2005); Aryani (2009); Bonyadi and Moses (2012) and others. Czech and Slovak linguists have discussed various issues of newspaper language as well Křístek (1973); Droppa (1998); Burcl (2011), Findra (2013)… etc. Among other functions, (Dor, 2003:28) emphasized functions and features which appear to be essential to most of the headlines. One of the functions introduced by Dor was providing summaries of news items in a telegraph-like manner. Headlines may provide an abstract of the article. Some headlines, however, may focus on single details instead. They may present a quotation or any other important details that the author or the editor decided to push to the foreground. (Bell, 1991:188 in Dor, 2003:697) makes the distinction between the headlines which provide a brief summary (abstract) of the story and those which focus on what he calls “secondary event” or “secondary detail.” Dor puts into contrast quality press headlines and tabloid headlines, stating that tabloid headlines are not telegram-like, nor are they even informative. Instead, they present the reader with a “fairly complex riddle” which first triggers schemata in the reader’s mind and then persuades them to read through the text. As for the types of headlines, linguists present quite varied typologies, each considering a different aspect of headline writing. Contentwise, (Křístek, 1973:2) distinguishes 1. Appellative headlines – their function is to denote and to specify the main topic of the article; 2. Evaluative headlines – the headlines in which the author evaluates the events he is writing about and provides his own opinion. 3. Headlines which are appeals towards doing / not doing something. While the typology by Křístek laid the focus on the content and the pragmatic value, (Mardh, 1980:25) brought a rather structural typology of headlines...
introducing the following types of sentential headlines: 1. simple headlines – the headlines introduced by a single independent clause, sometimes extended by optional clause elements; 2. compound headlines – the headlines which may be formed of two or more independent clauses linked by a coordinator; 3. complex headlines - those which may be analysed within the pattern of independent + dependent clauses, the dependent ones being finite, non-finite or verbless.

3 Grammatical and lexical tools of language economy

In this research, two major types of tools of language economy are applied, the grammatical devices represented by ellipses and sentence condensers, and lexical devices represented by abbreviated forms treated under the concept of minor word-formation.

3.1 Grammatical tools

3.1.1 Ellipses

Ellipses are traditionally characterized as omissions of certain items. Depending on the type of the omitted item, ellipses may be lexical, grammatical or structural. In headlines, though, a distinction needs to be made between an ellipsis and elision. (McCarthy, 2012:193) focused on the act of omission and introduced the definition of ellipses being such instances in which a word or a structure is left out which would normally be expected in a sentence if grammatical rules were followed. On the contrary, (Halliday, 1976:362) emphasizes the concept of naturality in that he defines ellipses as “substitution” by zero where “zero” stands for the items which receivers do not find incomplete because they may be sufficiently filled by the preceding or the following context; e.g. Have you spoken to him? Yes, I have (spoken to him). The sentence feels natural without the “spoken to him” part. (Abbas, 2012:1) develops this approach to ellipses and defines them as a sentence where, for reasons of economy, emphasis or style a part of it is omitted; e.g. She might sing but I don’t think she will (sing). The word “sing” is left out in order to avoid repetition. Nevertheless, the utterance is acceptable and does not seem defective without the word “sing”. (Biber, 2011:751) introduced various grammar contexts in which ellipses may occur: (1) Ellipses in coordinated clauses - these ellipses share elements with a preceding clause: I thought they were on the seat, but they were not (on the seat); (2) Ellipses in comparative clauses - comparative clauses largely mirror the structure of a preceding clause: She looks older than my brother (does) (3) Ellipses in question (answer sequences) - these are restricted to the conversation: “Have you got an exam on Monday?” (I’ve got) “two exams” (on Monday) 4) an elided full interrogative sentence: “What was the mileage when we got there?” (Was it) “a hundred and eleven?” (5) Omission of function words and situational ellipses: unstressed function words are left out in conversation: Do you know the woman at the end of our road? (She) takes me to school every day. In addition to these types, (Coppock, 2014:2) introduced the types of ellipses viewed predominantly from semantic and pragmatic perspectives. He introduced the so-called gapping referring to instances consisting of two clauses where the second clause does not contain a verb. The absence of the verb is then referred to as a “gap”, as in John likes caviar and Mary, beans (Mary likes beans). Coppock (ibid) calls the first clause antecedent (as the utterance which existed prior to the second, the referent sentence). The role of the antecedent clause is to provide ground for identifying the missing verb.

All elliptical instances introduced in this paragraph share one common feature – the absence of words or structures (the ellipsis) does not feel unnatural to the receiver. The structures in newspaper headlines differ in that omissions of words are considered defective. E.g. “US General Shot” (FT, Wed. 6th August 2014); an auxiliary verb is omitted. Unlike in the case of ellipses, the omission is neither fully
recoverable nor does it feel natural to the receiver. For this reason, such instances in newspaper headlines are referred to as elisions rather than an ellipsis. These elisions occur in the predicate part of a clause, and their omission will be considered the marker of grammatical defectiveness. Nominal ellipses or ellipses of the articles will not be justified as a marker of grammatical defectiveness because their usage is often arbitrary, based on the speaker’s understanding of the message. Elisions in the predicate part of the headline (if there is a predicate) will be set as one of the criteria of language economy.

3.1.2 Non-finite condensers

Other devices of language economy at the syntactic level include non-finite sentence condensers (infinitive, gerund, and participle). All these non-finite forms may enter various syntactic functions as commented by (Biber, 2011:198). Performing these syntactic functions, all non-finite structures are in the role of dependent clauses. They are more compact and less explicit than finite clauses, they are not marked for tense and modality, and they frequently lack an explicit subject and subordinator. Biber listed the following examples: I don’t know what to write about – non-finite, I don’t know what I should write about – represents the finite structure which is more explicit with a subject in the subordinate clause. These examples demonstrate that to fully understand the non-finite clauses, the finite clause or even a broader context is needed.

Some of the non-finite structures may have a condensing function. The issue of sentence condensation by means of these condensational devices was introduced by the Czech linguists Mathesius, Vachek, and Hladký (Malá:38). As Hladký puts it, sentence condensers are predominantly nominal forms derived from verbs such as infinitives, participles, gerunds, and verbal nouns. Vachek conducted contrastive research on the productivity of non-finite sentence condensers in English and Czech. The findings of this research determined participles as the most productive type of condenser, followed by infinitives and gerunds (Malá, 2013:38).

The infinitive condenser: Infinitives, unlike other sentence condensers, are able to express the full range of combinations of verbal categories (e.g. to do, to have done, to be doing, to have been doing). In the same time, they allow passive alternatives (e.g. to be done, to have been done, to be being done, to have been being done) (Quirk at al, 1991:153). Some of these forms refer to the present and some refer to the past. Grammarians use the terms present and past infinitive. As (Nováková, 2008:17) states, the present infinitives, i.e. the infinitives used in non-perfective structures, have their tense in accordance with the subordinate clause (e.g. He is believed to be a good student). On the other hand, no such correspondence may be seen in the perfective infinitives (e.g. He is believed to have been stolen), where the first part of the sentence refers to the present, and the perfective infinitive refers to the past. The full-time / tense reference may be seen when the non-finite infinitive structure is replaced by a finite subordinate clause (e.g. It is believed that he is a good student / It was believed that he was a good student). With respect to time, (Nováková, 2008:17) also states that non-finite infinitive structures may refer to futurity (e.g. I expect him to come) which could otherwise be replaced by the finite structure (e.g. I expect that he will come). The positioning of infinitive structures is bound to the regular word order. Infinitive structures most frequently take the initial or a final position in a sentence. It needs to be noted that not all infinitive structures have the condensing function. Condensing is considered those infinitives which may be fully replaced by a non-condensed dependent clause. For example, in the headline (Judge Archbald Likely to Know his Fate Today) the infinitival structure is not condensing a possibly longer sentence with a finite subordinate clause. Thus, the infinitive does not have the condensing function. On the contrary, in the headline
(Franco-Soviet pact to boost trade) the infinitive structure “to boost trade” functions as a condenser because it is possible to replace it by a finite clause. (Franco-Soviet pact which will boost trade).

The participle condenser: English participles are traditionally divided into the present and the past participles, e.g. Going home I saw him might be replaced by a past alternative (having gone). Participles are of two types – the “ing” participle (the example introduced earlier) and the “ed” participle (e.g. The man, covered with paint, is decorating living-room). While in the former category of English infinitive, the condensing infinitives were distinguished from those infinitives which do not have a condensing function, a great majority of participles may be regarded as condensers because they can be replaced by a finite clause. For example, when told by the police how badly injured his victims were, he said: “Good, I hope they die” (Biber, 2012:199) The participle structure (when told by the police) may be replaced by a finite clause (when the police told me...)

The gerund condenser: Gerund clauses are clauses in which the first verb in the verb phrase is a gerund, an “ing form”. Gerunds may perform as the so-called full gerunds, the ones whose subject is in the genitive or possessive case (e.g. It all depends on their helping us) and the half gerunds whose subject is in the accusative case (e.g. It all depends on them helping us). Like infinitives and participles, gerunds may as well perform in all aspects and voices and their combinations: I was surprised at having lost the race – perfect, They asked me about him/his meeting with known felons – perfect progressive, John is pleased at being given the “student of the year” award by his classmates – passive, Having been attacked by bears at the ZOO convinced me not to visit any more – perfect passive. These gerund structures may, again, be replaced by a finite clause. Thus, they are regarded as condensers.

Following the literature review, the language economy at the syntactic level is jointly represented by the following structures:

- Grammatical ellipses in the predicate part of sentence headlines
- Non-finite sentence condensers (infinitive, gerund, participle)

3.2 Lexical devices of language economy

Economical behavior at the lexical level of language may be observed in various forms. Most of these forms are commonly classified as abbreviations which also have their full counterparts. Abbreviated lexical forms are a part of the lexicon which, rather than consisting of a closed set of items, contains various productive processes (Panasenko, 2018:165). It is just abbreviated lexical processes which have recently been discussed by linguists who examine the overall economy of language. Crystal (2014) introduced a detailed work on the language of texting in which he discussed efforts to abbreviate language in electronic communication. Slovak linguists have introduced in-depth analyses of this type of language as well. To mention Lančarič (2008, 2009, 2011) who introduced a systemic-categorical typology of lexical abbreviations in English, French, and Slovak. Typologically, he divides abbreviations into non-lexicalized abbreviations (e.g. graphical abbreviations, numeronyms, graphic-phonetic abbreviations, etc.), lexicalized abbreviations (e.g. truncations), and related word-formative processes (conversion and affixation of abbreviations). In fact, a great number of abbreviations may simultaneously be categorized within the framework of minor word-formation processes introduced by lexicologists and lexical morphologists. The questions of word-formation have been elaborated by Slovak linguists Kvetko (2009), Pavlík (2017, 2018), Lančarič (2016). The overlap may also be caused by the fact that word-formation processes have not yet been clearly defined. The word word-formation does not have clear-cut, universally accepted usage. Sometimes, the notion “word-formation” is used to refer
to all processes connected with changing the form of the word or shifting the word from one category to another.

(Pavlík, 2017:28) divides word-formative processes into primary, secondary, and marginal. The categories of secondary word-formation correspond to what (Kvetko, 2009:63) refers to as minor word-formation. These word-formative processes are interpreted as minor or secondary because, in contrast with the major / primary word-formation process, they are typically assigned a lower level of productivity. In the category of secondary word-formation, (Pavlík, 2017:29) introduced shortening, blending, back-formation, reduplication, and lexical ellipses. The category of shortening falls into several sub-categories, including clipping, acronyms, initialisms, and graphical abbreviations. Clipping, as a sub-category of shortening, is removal of usually several letters of a word. Different parts of words may be clipped. Based on the position of the clipped part, the sub-categories are initial clipping – initial part is clipped (e.g. cello - violoncello, phone – telephone, gator – alligator), medial – the middle part is clipped (e.g. fridge - refrigerator, flu – influenza, shrink – head-shrinker), final – the final part of the word is clipped (e.g. doc - doctor, ad – advertisement, memo - memorandum) or combinations of these, also called complex clipplings (e.g. cablegram – cable telegram, opart – optical art). In some cases, both halves of a compound are clipped like in navicert – navigation certificate. Here, it may be difficult to classify the resulting form which might be classified as clipping or blending. Another sub-category of shortening are the so-called acronyms. Acronyms are abbreviations pronounced as regular words (e.g. OPEC /ˈoʊ.pek/ – Organization for Petroleum Exporting, FIST /fiːst/ – Federation of Inter-State Truckers)…; as opposed to initialisms in which initial letters are pronounced separately (e.g. TUC /tiːjuːsiː/ - Trades Union Congress, EU /iːjuː/). Both acronyms and initialisms may shorten word-groups or even sentences; the latter happens predominantly in informal and electronic communication (e.g. ICQ – I seek you). Sentential acronyms are not discussed in this article because they scarcely occur in headlines. Graphical abbreviations are relevant for newspaper writing because, in fact, they are used only in written language. When read, these abbreviations are pronounced as full words (e.g. &; $, £). Blending is commonly characterized as combining two words into one. There is a close link between blending and clipping because it is not rare that words are clipped before they are blended, e.g. workaholic – work (alco)holic. Another device of language economy are lexical ellipses. These are commonly defined as omissions of parts of multiword expressions (e.g. disabled - disabled people).

Previous research into the lexical features of newspaper headlines detected acronyms and initialisms as the most productive forms in this type of writing, which may be caused by their informative function and their ability to save space.

The criteria for identifying language economy at the lexical level are the following:

- shortening (clipping, acronyms, initialisms, graphical abbreviations),
  blending, back-formation, lexical ellipses.

4 The description of the dataset and methodology

The research is conducted as a diachronic quantitative analysis of lexical and grammatical devices of language economy. Data have been collected manually from digitalized online libraries and databases. It has been decided for manual extraction because no software was available which would fully meet the requirements of this research. The final dataset is represented by a collection of 500 randomly selected newspaper headlines from broadsheet and tabloid newspapers from

XLinguae, Volume 14 Issue 2, April 2021, ISSN 1337-8384, eISSN 2453-711X

309
3 English-speaking countries (the UK, the USA, and Canada). The diachronic time axis was divided into decades, and each decade was supplied by more or less the same number of newspaper headlines from that period. The heterogeneity of the dataset is secured by the type of newspaper (broadsheet / tabloid), by the represented countries (3 countries present), and by the period of extraction (1900 – present). See sample below.

| No. | Sample Headline | Year | Source |
|-----|-----------------|------|--------|
| 25 | The king of Italy’s visit to Paris | 1903 | The Observer |
| 26 | The riots in Madrid. Many persons killed and injured. | 1903 | The Observer |
| 27 | The anti-christians measures in France | 1903 | The Observer |
| 28 | The home rule story from Ireland | 1903 | The Observer |
| 29 | The doctress of the American civil state | 1904 | The Guardian |
| 30 | Disappointing yield of the death duties. | 1909 | The Guardian |
| 31 | Flying Machine Takes To The Air | 1903 | London Herald |
| 32 | Wright Brothers Make First Powered Flight | 1903 | London Herald |
| 33 | Suicide in Tombs Prison | 1903 | The Sun |
| 34 | Disease in Mrs. Fofly’s house | 1903 | The Sun |
| 35 | Again Thrown Rock Island | 1903 | The Sun |
| 36 | Funeral Carriages Empty | 1903 | The Sun |
| 37 | Ramseyer Decision Opens 500 New Homes To Race | 1905 | / |

Table 1: dataset sample

Prior to collecting all of the dataset, a pre-sample of 50 headlines was extracted from various periods randomly, and pilot research was conducted in order to test whether the dataset would be searchable. At the next stage, the criteria for identification of lexical and grammatical economy were set (introduced in the review of literature), and the dataset was checked for the desired cases. The occurrence of cases varied greatly. Some headlines contained only lexical and some only grammatical items; in some headlines, both types were detected, and, on the contrary, some of the headlines were not represented by any cases at all. Further, the binomial test and chi-square statistics are employed in order to test the significance of differences in the distribution of lexical and grammatical devices across the time stages. The second part of the quantitative analysis tests the possible positive correlation between the time and the productivity of lexical and grammatical devices of language economy, as introduced in the hypotheses.
5 Hypotheses – testing

H: There is a significant difference in the distribution of lexical and grammatical devices in some periods across the time axis.

The overall distribution of lexical and grammatical devices is outlined in the frequency tables below.

**Table 2:** The descriptive statistics for the grammatical and the lexical devices

In 506 observed cases (some headlines were marked by more than one case), lexical devices were tagged in 13.4% of cases, and grammatical devices were tagged in 22.5%. Overall, grammatical devices prevail over lexical ones. The crosstabs below demonstrate the distribution of lexical and grammatical devices in the individual decades in an attempt to find out whether the distribution of lexical and grammatical devices (separately) differs significantly in individual periods.

**Table 3:** Chi-Square test of independence, period – lexical device
The crosstabs for the distribution of lexical devices show almost identical values (the ratio between the “device” and “no device”) in all decades except for the last decade (2011–2020) where the distribution of lexical devices was significantly higher. $x^2 (11, n=506) = 29.728, p = 0.002$

The $p$ value of 0.02 proves an outstanding value in distribution only in one of the decades. The higher occurrence was caused by the type of situations and events and the way they were reported. It was observed in the corpus that the latest period was saturated by the vocabulary closely related to the European Union (e.g. the acronym EU, or the blends eurosceptics, Brexit, etc.). The higher occurrence could also be caused by the enormous efforts mentioned earlier to put a minimum effort into delivering a message. Equally, the expressions such as eurosceptics and Brexit are neologisms which have been created to refer to some latest events. This may reflect the way neologisms are coined. In fact, many of them are not created as brand-new words but appear to use the already existing expressions (two words are blended into one or some neologisms were coined as just polysemous words where the existing expression has acquired a new secondary meaning).

The crosstabs for the distribution of grammatical devices show the values which are close to identical, with little aberrances in individual decades across the explored time axis (1900–present). The statistic chi-square test was applied, giving the following results: $x^2(11, n=506) = 15.240, p = 0.172$. The $p$ value shows no significant aberrance in the distributions in any of the explored decades, i.e the distribution of grammatical devices of language economy is completely arbitrary, irrespective of time.

To test the overall equality in the distribution of grammatical and lexical devices, the ratio was set to 50:50 (lexical – grammatical). The following table shows the observed values:

Table 5: Chi-square goodness of fit test
The observed values (lexical: 68, grammatical: 114) demonstrate a significant prevalence of grammatical devices in the dataset $x^2(1, n=182) = 11.626, p < 0.001$). Whereas the previous calculations were done for each category (lexical and grammatical) separately, the following data demonstrate the possible influence of the decade on the mutual ratio between the lexical and the grammatical devices applied.

| decade | Count | Lexical | Grammatical | Total |
|--------|-------|---------|-------------|-------|
| 1      | 4     | 7       | 11          |       |
|        |       | 4.1     | 6.9         | 11.0  |
|        |       | 0.0     | 0.1         |       |
| 2      | 6     | 13      | 19          |       |
|        |       | 7.1     | 11.9        | 19.0  |
|        |       | 0.0     | 0.0         |       |
| 3      | 6     | 14      | 20          |       |
|        |       | 7.5     | 12.5        | 20.0  |
|        |       | -0.7    | -0.7        |       |
| 4      | 5     | 7       | 12          |       |
|        |       | 4.5     | 7.5         | 12.0  |
|        |       | -0.3    | -0.3        |       |
| 5      | 4     | 8       | 12          |       |
|        |       | 4.5     | 7.5         | 12.0  |
|        |       | -0.3    | -0.3        |       |
| 6      | 5     | 13      | 18          |       |
|        |       | 6.7     | 11.3        | 18.0  |
|        |       | -0.9    | -0.9        |       |
| 7      | 2     | 5       | 7           |       |
|        |       | 2.6     | 4.4         | 7.0   |
|        |       | -1.5    | -1.5        |       |
| 8      | 2     | 8       | 10          |       |
|        |       | 3.7     | 6.3         | 10.0  |
|        |       | -1.2    | -1.2        |       |
| 9      | 6     | 8       | 14          |       |
|        |       | 5.2     | 8.8         | 14.0  |
|        |       | 0.4     | 0.4         |       |

Table 6: Chi-Square test of independence, period – type of device

It may be observed from the table that the real counts are proportional to the expected counts in decades 1 – 11. However, decade 12 is marked by a disproportional occurrence of the real count of lexical devices against the expected count $x^2(11, n=506) = 7.904, p=0.722$. The p value indicates the level of significance which subsequently permits the acceptance of H0. It may be stated that the decades have not had an impact on the distribution of lexical and grammatical devices.

It may be observed from the table that the real counts are proportional against the expected counts, $x^2(11, n=506) = 7.904, p=0.722$. Even though, decade 12 is marked by a higher occurrence of real count lexical devices against the expected count, it is not proven significant in relation to all periods of time. The p value indicates the level of significance which subsequently does not allow the falsification of H0. It may be stated that the decades do not impact the distribution of lexical and grammatical devices.

H: There is a positive correlation (growing tendency) between the time axis represented by the decades and the distribution of grammatical and lexical devices.

The hypothesis is testing a possible gradual increase in the occurrence of lexical and grammatical devices in the time. The hypothesized positive correlation is based on the variable time (the decades from 1900 to 2020) and the device variable (the counts of identified grammatical and lexical devices) A two-tailed chi-square test
will be used to test the possible positive correlation between the time and the increasing occurrence of devices.

### Table 7: Pearson’s correlation coefficient

Pearson’s correlation coefficient did not prove significant for both lexical devices ($r = 0.394; p = 0.205$) and grammatical devices ($0.021; p = 0.949$). For lexical devices, the mediocre association between the decade and increase in the amount of observed lexical devices may be the result of the significant change in lexical devices (counted in the test above) observed only in the final decade. As the variable decade explains only 15.5 % of the variability in the variable lexical devices, it is safe to suppose that there is no effective positive correlation between the time (increasing decade) and the increasing occurrence of devices. For grammatical devices, the variable decade explains less than 0.1 % of the variability in the variable grammatical devices. The results are illustrated in the graph below:
The two graphs illustrate the behavior of lexical and grammatical devices of language economy in time. As mentioned earlier, the decade 2010 – present was more abundant in lexical devices than any other decade. Graph 1 includes the 2010 – 2020 decade. The black line demonstrates a slight increase in the distribution of lexical devices. The red line marks the behavior of grammatical devices – this shows no correlation with time. To compare, graph 2 does not include the 2010 – 2020 decade proving the fact the distribution of lexical and grammatical devices of language economy is random, irrespective of time.

Based on the lack of evidence for falsifying the H0, an alternative hypothesis (H1) cannot be accepted. Thus, in accordance with H0, it may be stated that there is no positive correlation between the time axis from the 1900s to the present (2020) and the distribution of lexical and grammatical devices of language economy.

6 Conclusion

The language of newspaper headlines, headlinese, is traditionally considered a sub-genre of journalese. Both journalese and headlinese are commonly viewed as an obscure type of language due to a number of irregular and grammatically ill structures. There may be several reasons for using these irregularities; probably, two major ones are concerned with semantics and pragmatics of headlines – headlines are created to capture the reader’s attention and to put a maximum amount of information in a minimum of space. Decoding of irregularities may be conditioned by an overall metalinguistic awareness of the reader (Hlava, Šipošová, Hankerová, 2020:18). This space-saving factor directly calls for applying various lexical and grammatical abbreviated forms. This research paper was concerned with an exploration of the distribution and productivity of abbreviated lexical and grammatical forms in time, i.e. it is a diachronic study of how these abbreviated lexical and grammatical forms have behaved in different time stages. Two major questions were raised: 1. Is there a significant difference in the distribution of lexical and grammatical devices of language economy across time? 2. Is there a positive correlation between the time and the distribution of devices of language economy? Two hypotheses were formulated in accordance with these questions.
The first hypothesis tested a possible significant difference in the distribution between the lexical and the grammatical devices of language economy in the decades from 1900 up to the present. Testing the first hypothesis showed a significant prevalence of lexical devices only in the last decade (2010 – present). A more detailed analysis of corpus findings showed that the prevalence of lexical devices was caused by using a lot more EU vocabulary, which is quite abundant in abbreviated forms e.g. EU (the European Union), Eurosceptic (a person who is opposed to increasing the powers of the European Union), Brexit (a blend of British and exit) and others. In the other decades, there were no significant differences in the distribution between lexical and grammatical devices of language economy. The higher occurrence of the real count of lexical devices in decade 12 is not proven significant in relation to all other decades. The p value indicates a level of significance which does not allow the falsification of a null hypothesis. It is safe to state that newspaper headlines are more or less equally supplied by lexical and grammatical devices of language economy and that their distribution is irrespective of time. A partial result drawn from testing this hypothesis is that, regardless of time, the total number of grammatical devices of language economy prevails over the lexical ones.

Among other lexical devices of language economy, the following were detected: acronyms and initialisms (e.g. NATO in NATO worries about détente and deterrents), lexical ellipses (e.g. House Agrees to Conference Report On District Bill) where the house refers to the House of Commons or Gertrude Ederle Swims Channel where the Channel refers to the English Channel. The instances with a missing lexical word in the headline, which would certainly be present in a regular context were treated as a lexical ellipsis as well (e.g. Fire in Factory Kills 148 (people); Girls leap to their death). Some shortened expressions were included even though they are not typical only of the language of headlines, e.g. TV or A.M. like in War is Over! Fighting At 9. A.M. Our Time Stopped. Some headlines were marked by clipped expressions, e.g. Even easier than taking snaps - making movies with a Ciné-Kodak. Where the word “ciné” refers to the cinema. Some cases were marked by a synonymous use of words, each synonym having its own clipped form (nuclear bomb - nuke / atomic bomb – A-bomb) like in the headlines: NATO worries about détente and deterrents, Britain's forgotten A-bomb victims.

Grammatical devices of the economy in language were supplied by the elisions of the verb “to be” in the predicate part of the headlines, e.g. Washington and Oregon in Line for Prosperity (The Observer 1901), The riots in Madrid. Many persons killed and injured (The Observer 1903) and by non-finite condensers like in Griffith Charges Attempt Was Made to Switch Balls (The Guardian 1915).

The second hypothesis tested a possible positive correlation between the time and the productivity of lexical and grammatical devices. This hypothesis tested whether the devices of language economy are applied in headlines randomly or whether there is a gradual increasing tendency towards applying these devices as the time approaches the present. At the very beginning, newspapers or their forerunners either did not contain proper headlines, or the headlines were a sort of summary of the newspaper content. The assumption behind this hypothesis was the fact that first headlines were regular statements which used no abbreviated forms. Later, hand in hand with an increasing need for an economy of language, shortened headlines were created as a reaction to increasing economization tendencies in language. Again, not enough evidence was provided for the falsification of the null hypothesis, and it is safe to state that the distribution of grammatical and lexical devices is random, independent of the time. No direct correlations between the time axis and the distribution of lexical and grammatical devices of language economy were detected.

In conclusion, the findings of this research only proved the obscure and devious nature of the headline. The generalizations to be made here are that the devices of language economy (the abbreviated lexical and grammatical forms) are by
no means regular. They are used by authors spontaneously and randomly. Their usage is, to a great extent, influenced by their readership.

This article is written as a part of VEGA project named VEGA project no. 1/0769/20 “The Language of Electronic Media in the English Linguocultural Context”.

Bibliographic references

ABBAS, F. N. – UGAIR, K. S. 2012. Elliptical Analysis Between Linguistic Economy and Attention Focus. In: Language in India, vol. 12, n.2, pp. 2-15. ISSN 1930-2940

ARYANI, D. 2009. Ellipsis in English Printed Media Headlines. Skripsi. English Department. Cimanggis: Faculty of Letters: Indonesia. Gunadarma University, Available online: https://www.semanticscholar.org/paper/Ellipsis-in-English-Printed-Media-Headlines-Aryani/7866d4d1ef41ebafdad695fba9593f49b8f6ee

BIBER, D. et al. 2011. Longman Grammar of Spoken and Written English. Harlow: Pearson Education Limited. ISBN 0582237262

BONYADI, A. – MOSES, S. 2012. Headlines in Newspaper Editorials: Contrastive Study, International Research Journal of Social Sciences, vol. 1, n. 3, Available online: https://journals.sagepub.com/doi/full/10.1177/2158244013494863

BREMNER, J. B. 1972. A Study in News headlines. School of Journalism. Kansas: University of Kansas.

BURCL, P. 2011. Jazyk novinových titulkov a ich preklad do slovenskeho jazyka, Dissertation. Nitra: FF UKF.

CARTER, R. – McCARTHY, M. 2006. Cambridge Grammar of English. Cambridge: Cambridge University Press. ISBN 9780521674393

COPPOCK, E. 2011. Gapping: In Defence of Deletion. Northwestern University, Available online: http://eeccoppock.info/Gapping.pdf, quoted in April 2014

CRYSTAL, D. 2014. Language, Communication and Pragmatics, Available online: http://cw.routledge.com/textbooks/9780415602679/pdfs/synopsis1.pdf

DOR, D. 2003. On Newspaper Headlines as Relevance Optimizers. Journal of Pragmatics 35, pp 695-721. Department of Communications Tel Aviv University, Tel Aviv 2003, Available online: https://www.researchgate.net/publication/229005694_On_Newspaper_Headlines_as_Relevance_Optimizers

FINDRA, J. 2013. Stylistika Slovenciny. Bratislava: Oveta. ISBN 9788080634049

HALLIDAY, M. A. K. - HASAN, R. 1976. Cohesion in English. London: Longman Group. ISBN 13 978 0 582 55041 4

HARGREAVES, I. 2005. Journalism, A Very Short Introduction. Oxford: Oxford University Press. ISBN 0199686874

HLAVA, T. – SIPOSOVA, M. – HANKEROVA, K. 2020. Vztahy a interakcie deklaratívnej a procedurálnej roviny vo vyučove anglickej gramatiky. Bratislava: Univerzita Komenského v Bratislave. ISBN 978 80 223 4879 9

HOWARD, P. 2000. The Press Gang, The World of Journalese, Monograph series, no 40. London: The Institute for Cultural Research pp: 1622-1932. ISBN 0904674320

KRISTEK, V. 1973. Současné novinové titulky. Nase rec. vol. 56, n. 5, Available online: http://nase-rec.ujic.cas.cz/archiv.php?art=5724

KVETKO, P. 2009. English Lexicology in Theory and Practice, 2nd Revised Edition. Trnava: Univerzita sv. Cyrila a Metoda. ISBN 97888081051104

LANCARIC, D. 2008. Jazykove skratky – Systemovo-kategorialny opis a ekvivalentacna semantizacia anglickych a francuzskych abreviatur. Bratislava: Lingos Bratislava. ISBN 978 80 89328 08 6
