

Jordan's Economy: A Corpus-Based Study

Atef Aljbour

Language Center, Hashemite University, Zarqa, Jordan

Tel: 962-772-400-199 E-mail: aljbour2005@yahoo.com

Abstract

The present study attempts to investigate Jordan's economy throughout two periods (2001-2003; 2010-2012). For this purpose a journalistic corpus containing 48,291 news items obtained from the archive of Petra News agency, totaling 13,045,108 words was compiled. Out of this main corpus, an "economy subcorpus" was extracted to carry out a search of the phrase "Jordan's economy" by using 'AntConc' software. The analysis of the contexts reveals, first, that the vast majority of the contexts in the first period (85%) carries positive connotations about the economy outlook in Jordan as opposed to only 25% in the second period, and second, that Jordan economic crisis which was expressed in the second period was largely attributed to the international financial crisis, domestic policies, which has been aggravated by the consequences of the "Arab Spring", and particularly the Syrian crisis. The revival of Jordan economy in the first period is attributed to Jordan's international economic agreements, and the government's privatization programs.

1. Introduction

With the invention of computers and modern technology, new opportunities for researchers have been opened to study and analyze language. This availability of modern computer technology, and consequently, the ability to computerize large amounts of language data and to access and retrieve this data led to the emergence of corpus linguistics. This research approach that has evolved recently has facilitated the analysis of a large body of information to aid exploring language use and revealing results of greater validity than would otherwise be possible.

Leech (1992: 106) emphasized the role of computer technology in linguistic research as follows:

The computer's ability to search, retrieve, sort, and calculate the contents of vast corpora of text, and to do all these things at an immense speed, gives us the ability to comprehend, and to account for, the contents of such corpora in a way which was not dreamed of in the pre-computational era of corpus linguistics.

The word ‘corpus’ has a Latin origin meaning ‘body’. In linguistics, a corpus is "a large collection of linguistic data, either written texts or a transcription of recorded speech, which can be used as a starting point of linguistic description or as a means of verifying hypotheses about a language (Crystal, 1995). According to (Dash 2005: 12), a corpus should represent both the common and special linguistic features of the language from which it is designed and developed; It should be large and wide enough to contain materials from a variety of fields, and it should be accessible in a computerized format.

Corpus linguistics, on the other hand, is the study of large, computerized collections of real life language use. This method “encompasses the compilation and analysis of collections of spoken and written texts as the source of evidence for describing the nature, structure, and use of languages” (Kennedy, 2001). Some corpus linguists refer to this field as computer corpus linguistics (e.g., Leech, 1992: 106). Others considered it as a methodology (e.g., Giries, 2009: 1).

In his article ‘What is Corpus Linguistics?’ Giries (2009) described corpus linguistics as one of the growing approaches in contemporary linguistics. The author also discussed some of the concepts and methods that are central to this field (collocations, concordances, frequency lists, balancedness, representatively, etc.) In his answers to questions raised by linguists who have not used corpus-based approaches so far, the author insisted that corpora are “text files collected and stored in the form of text Resume ASCII (American Standard Code for Information Interchange) or Unicode text files that can be loaded, manipulated, and processed platform-independently and all you can get out of such files is distributional (or quantitative/statistical) information”.

Therefore, corpus linguistics is based on the analysis of descriptive data rather than on linguistic theories. As for the texts that make up the corpus, the author maintained that they must have been produced in a natural communicative setting, i.e., they were written or spoken for an authentic purpose, but not for the purpose of putting them into a corpus such as the newspaper articles that are produced in a natural setting to convey something to the readers, but not to just fill a linguist’s corpus.

2. Literature Review

Plenty of research in a wide array of languages have made use of corpus technology to handle a wide range of questions pertaining to issues in linguistic and cultural studies, sign language, information technology, language teaching and learning. These multifarious uses of corpus technology evidently demonstrate the great prospects and opportunities that corpus linguistics can provide for researchers, language teachers and students to inspect and collaborate. For example, in her investigation of recent development on corpus linguistics and corpus-based research in Hong Kong, Xie (2013) identified 29 on-going research projects from the eight research-active universities in Hong Kong. These studies were classified according to their primary

focus as follows: Teaching and learning (12), Linguistic research (11), Cultural studies (3), Information technology (2) and Sign language (1).

To study the frequent use of noun phrases in academic writing which make it difficult for less competent readers to figure out a text, Ahmadi, et al. (2020), analyzed noun phrase modifiers in applied linguistics research article abstracts between expert non-native English Persian writers and international writers. They constructed 38,762-word corpus consisting of 109 international academic research articles and 100 Persian English-medium research articles. Data analysis revealed that international writers differed significantly from Persian writers in the use of total noun phrase modifiers, relative clauses, and post-modifying prepositional phrases; Persian writers used lexical bundles to modify noun phrases more frequently than international writers. The findings of this study offered a deep understanding of the way international and non-native academic writers in applied linguistics make use of phrasal features for complexifying research article abstracts.

Much literature has also been devoted to discussing topics associated with the design of certain corpora. In designing a Corpus of Contemporary Arabic (CCA) which represents the state of the Arabic language at the present time, Al-Sulaiti and Atwell (2003), for example, conducted a survey of the needs of teachers of Arabic as a foreign language (TAFL) and language engineers. The results of the survey revealed that existing Arabic corpora were restricted in use; they were too narrowly limited in genre; they were generally compiled for a specific research project rather than as a resource for teachers of Arabic as a foreign language; and were not representative of Contemporary Arabic. Therefore, the researcher took into consideration some factors when designing the corpus. These include the need for a corpus of contemporary Arabic covering a broad range of text-types; giving priority to a number of text types to be included in the corpus; the inclusion of Arabic texts which are available and current; and the integration of parallel English-Arabic texts.

Another corpus commonly used for diachronic change analysis is the Corpus of Historical American English (COHA). The corpus contains more than 400 million words in more than 100,000 texts covering the period from the 1810s to the 2000s. It includes texts extracted from various genres, including fictions, non-fictions, newspapers, magazines and so on (Davies, 2012). According to the author, COHA allows users to easily quest the 400 million words to see how words, phrases and grammatical constructions have increased or decreased in frequency, how words have changed meaning over time, and how stylistic changes have taken place in the language. It was obvious that when designing the Corpus of Historical American English, some of the key consideration for building and designing a corpus were taken into consideration. First, the corpus is designed to include large data to accurately track and study recent changes in the language. It is taken into consideration that the corpus should be tagged for part-of- speech. It is

also designed to allow for a wide range of research on lexical changes, morphological shifts, syntactic shifts (due to very accurate lemmatization and part of speech tagging), and semantic change (by comparing collocates over time, as well as searches that use data from the integrated thesaurus and customized word lists).

COHA is also related to other large corpora including the Corpus of Contemporary American English (COCA) which is by far the most widely-used of these corpora. The corpus contains more than one billion words of data, including 20 million words each year from 1990. This makes COCA the only corpus of English that has a wide range of genres (Davies, 2020).

The corpus, as stated by the author, was designed to allow users to potential uses, including researching common American idioms; double-checking prepositions and verbs; comparing styles between spoken and academic American English; finding the right word for each medium. Other features have also been considered in designing the corpus. These include its size (1.0 billion words), how up-to-date it is (texts 1990- 2020), the wide range of genres (e.g. spoken, TV/movies, newspapers, magazines, Academic, web, and other web pages). This provides useful information about the frequency of words, phrases, and grammatical constructions across the genres. Searching the corpus, the author added, was also designed to allow users to carry out queries in a variety of ways that make searches useful for all types of users including, among others, language learners and teachers, and those who are interested mainly in academic English.

Harb (2016), in an experimental study, emphasized the role of COCA in helping ESL undergraduate's students improve their writing quality. Results revealed the experimental group with a better writing proficiency than the control group in terms of showing more proper use of vocabulary and synonyms in their writing and more lexical sophistication. The findings of the study were also influential in showing teachers the lexical features that they should consider for developing their students' writings and the effective role of COCA in helping educators develop students' vocabulary competence.

Many other corpus-based research studies were carried out employing applicable corpus linguistics techniques with the aim of collecting data for a specific-purpose corpus. Beloso (2015), for example, presented the criteria needed in building the CADCE (Corpus of Architecture Discourse in Contemporary English). This corpus includes around 500.000 words of written language representing the language used in Architecture, a field that involves many other relevant sub-fields such as landscape architecture construction, interior design, etc. It includes texts published from 2007-2008 in North-American, British, Irish, Canadian, and Australian publications. According to the study, the corpus to be built had to meet three main standards. First, it had to be "representative" (size, topic, sources, and level of technicality) to represent what it was intended for. The second criteria that should be accomplished is "contemporariness" (online, free-

accessed, computerized texts). Finally, besides being representative, the corpus had to be accessible to the professionals and people who are learning and teaching Architecture.

3. Current Study

The current research is based on an Arabic corpus compiled by the researcher in March 2022 and made up of the Arabic versions of news items obtained from the archive of Petra News Agency. The news items covered the period between 1992 up to 2012. The coverage of the news included one month from each year. It consisted of 48,291 news items distributed across eight genres (economy, investigation, education, international, sports, Arabic, local, and miscellaneous), totaling 13,045,108 words. The following table shows detailed information about the news items and the number of words in each year from 1992 to 2012, and the total number of the items and the words, which constituted the entire corpus.

Table 1:

Year	News Items	Words
1992	1114	710020
1993	668	526728
1994	407	233830
1995	5304	1106298
1996	3151	979902
1997	3761	809821
1998	4613	1255313
1999	3464	718683
2000	2041	426728
2001	929	202160
2002	3314	957692
2003	4845	667886
2004	1905	650995
2005	2517	721127
2006	1976	485963
2007	1680	484753
2008	1820	386952
2009	1756	324837
2010	2053	372154
2011	1623	298754
2012	2317	704512
Total	48291	13045108

The aim of this study is to investigate the economic situation in Jordan during two time periods starting from 2001 up to 2012. The time periods are divided as follows (2001-2003; 2010-2012) and to make a comparison between the economy outlooks across the two time periods.

4. Methodology

The researcher extracted an economic subcorpus from the main corpus (the journalistic corpus). The subcorpus was composed mainly of the news items which originated under the category 'economic'. Thus, this subcorpus contained only economic texts that were distributed across the period from 1992 to 2012. In order to investigate Jordan's economy during the abovementioned periods (2001-2003; 2010- 2012), the subcorpus was uploaded onto "AntConc" program which is the concordance software chosen by the researcher to carry out the needed search and investigation of the data. The phrase 'Jordan's economy' was used as a search phrase by entering it in the search box to generate all occurrences of that phrase. In the KWIC results, the key phrase appeared with its context. Clicking on the instance of the key term takes us to the 'File View' for the text file in which it occurred, showing the term in its wider context. Another way to reveal the words that collocate with this phrase was through generating a list of the common collocates of the search phrase by using the 'Collocates' Tool Tab at the top of AntConc window. This allows us to sort out the words (adjectives, verbs) that collocate with the phrase "Jordan's economy" (e.g., *booming, buoyant, dynamic, healthy, sound, stable, strong, ailing, depressed, nagging, fragile, Stagnant, weak, develop, expand, flourish, collapse, contract, fail, slow, stagnate, recover, stabilize, etc.*). This obviously will give us an idea about the status of the economy in a particular period and enable us to make a comparison with that in the other one.

5. Findings and Discussion

By searching the phrase "Jordan's economy" in the first (2001-2003), one hundred and fourteen (140) occurrences of the phrase in question appear as illustrated in the image blow followed by an image of an example of a text file in which the phrase occurred:

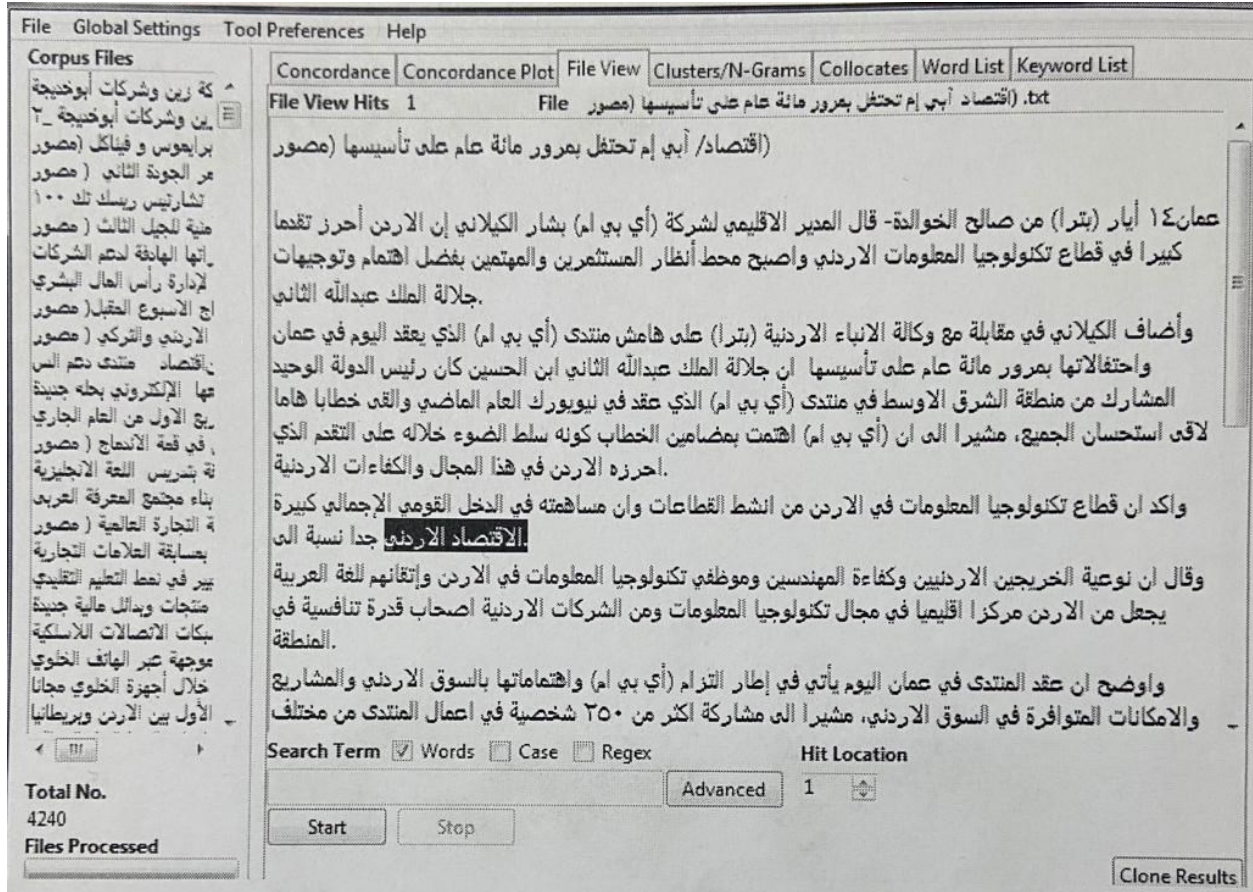
Image 1: Occurrences of the phrase 'Jordan's economy'

The screenshot shows a concordance tool interface with the following components:

- Menu Bar:** File, Global Settings, Tool Preferences, Help
- Corpus Files:** A list of files on the left side, including 'كافة زين وشركات أبوخديجة', 'زين وشركات أبوخديجة_٢', 'برايوس و فيناكل (مصور)', etc.
- Concordance Hits:** A table with columns 'Hit', 'KWIC', and 'File'. It displays 14 search results for the phrase 'الاقتصاد الاردني'.
- Search Term:** A text box containing the search term, with options for 'Words', 'Case', and 'Regex'.
- Search Window Size:** A control set with an 'Advanced' button and a value of '50'.
- Buttons:** 'Start', 'Stop', 'Sort', and 'Clone Results'.
- Kwic Sort:** A section with checkboxes for 'Level 1 1R', 'Level 2 2R', and 'Level 3 3R'.

Hit	KWIC	File
1	ساهمته ب ١٤ بالمئة من الاقتصاد الاردني يؤثر فيما يزيد على ٨٠ الف وظيفة في	روني في المنطقة
2	ان الحكومة تبذل جهودا الاقتصاد الاردني وزارة الصناعة والتجارة مها علي ابرز مؤ	اقتصادي مع كندا
3	وعية الخريجين الاردنيين. الاقتصاد الاردني الدخل القومي الإجمالي كبيرة جدا نسبة الى	تأسيسها (مصور
4	الى صعوبات ستواجهها الاقتصاد الاردني املة. وتناول الدكتور نور التحديات التي ي	حالي ... اضافة ٢
5	حالي يتمثل بعجز الموازنة الاقتصاد الاردني الدكتور نور ان التحدي الثاني الذي يواجه	حالي ... اضافة ٢
6	ار (بتر)- عبرت السفيرة الاقتصادية الاردني اقتصاد/ السفارة الفرنسية واثقة بقدره	الاقتصاد الاردني
7	الازمة الاقتصادية الحال الاقتصاد الاردني في عمان كورين بروزيه عن ثقتها بقدره	الاقتصاد الاردني
8	بحالة الربيع العربي وأز الاقتصاد الاردني المالية اخيرا حول اثر المتغيرات الخارجية	مالي واقتصادي
9	بحالة الربيع العربي وأز الاقتصاد الاردني المالية اخيرا حول اثر المتغيرات الخارجية	لي واقتصادي_٢
10	بوعات رئيسية هي العمالة الاقتصاد الاردني وصنفت الدراسة العمالة فيذي القطاع الح	مالة غير الرسمية
11	ب ٨ بالمئة منها ما نسب الاقتصاد الاردني ان نسبة العمالة في القطاع الخاص في	مالة غير الرسمية
12	الاقتصاد الاردني الخصوص في حل المشاكل والتحديات التي	ر الريادي للاردن
13	وبين الطابع ان التحديات التي ي	مافة ثانية وأخيرة
14	الاقتصاد الاردني على سوريا نظرا لتاثير هذه العقوبات على	في منتدى دافوس

Image 2: Example of a text file in which the phrase ' Jordan's economy ' occurs



By studying the contexts in which the phrase occurred it was found that only 15% of them refer to negative connotations of the economy status, particularly in 2003, and most of the negative contexts attacked the international economic cooperation and agreements such as Jordan-European Association Agreement that was signed on November 24, 1997. For example, we found expressions like:

- *Jordan Economy faces difficult challenges.*
- *Jordan economy suffers several troubles due to the siege against the occupied territories.*
- *The agreement (Jordan –European Association) needs revision.*

On the other side, the vast majority of the contexts (85%) described Jordan economy in a positive way. Examples of such expressions included:

- *promising economy outlook*
- *economy revival*
- *Jordan witnesses an open-door economy.*

- *Aqaba zone attracts a broad investments economy.*

It is also found that there was a link between the positive contexts and the certain factors such as the international cooperation agreements (Jordan –European Association Agreement, the United States-Jordan Free Trade Agreement which was signed on October 24, 2000) and the government's Privatization Program which were the same elements that receive attacks from some Jordanian policy makers at the end of this period.

Investigating the contexts in which the phrase "Jordan Economy" occurred in the second period (2010-2012) revealed that, unlike the first period ((2001-2003), the majority of the contexts in this period (75%) carried negative expressions. Examples of such expressions included:

- *Pressing troubles facing Jordan economy*
- *Economy is influenced by the financial international crisis.*
- *Deficit in the public budget*
- *Where has the privatization revenues gone?*
- *Rise in the unemployment rates*
- *Fighting corruption is a priority.*
- *There is a need to reconsider privatization programs.*
- *'Arabic Spring' affects Jordan economy.*
- *New taxes impose on a number of commodities.*
- *A need to exclude Jordan from the sanctions on Syria*

What was remarkably noticed is that only 25% of the contexts included positive or neutral expressions. Examples of such expressions included:

- *The monetary policy is in an appropriate stance".*
- *Government starts enacting legislations to attract investments and stimulate economy.*
- *'Information technology' witnesses steady development.*

It was also noticed that the regional and international conflicts, in addition to the international finance crisis were present when talking about the negative situation of Jordan economy in this period.

6. Conclusion

As shown above, it was found that 85% of the contexts in the first period (2001-2003) reflected positive expressions towards Jordan economy. Most of the negative contexts which constituted only 15% as mentioned above were noticed at the end of 2003, and they expressed pessimism towards Jordan economic agreements with Western World. Unlike the first period, the majority of the contexts in the second period (75%) carried negative expressions.

Jordan economic crisis which was expressed in the second period was largely attributed to the international financial crisis, domestic policies, which have been exacerbated by the consequences of the ‘Arab Spring’, and particularly the Syrian crisis. The superficial revival of Jordan economy in the first period was attributed to Jordan's international economic agreements, and the government’s privatization programs.

References

- Ahmadi, M., Rajab E. & Abbas A. (2020). Corpus-based Study of Noun Phrase Complexity in Applied Linguistics Research Article Abstracts in Two Contexts of Publication. *Iranian Journal of English for Academic Purposes*, 9(1): 76-94.
- Al-Sulaiti, L. and Eric, A. (2003). The Design of a Corpus of Contemporary Arabic. *International Journal of Corpus Linguistics*, 11(1): 1-36
- Beloso, B. S. (2015). Designing, Describing and Compiling a Corpus of English for Architecture. *Procedia – Social and Behavioral Sciences*, 198, 459-464.
<https://doi.org/10.1016/j.sbspro.2015.07.466>
- Crystal, D. (1995). *The Cambridge Encyclopaedia of the English Language*. Cambridge: Cambridge University Press.
- Dash, N.S. (2005). *Corpus Linguistics and Language Technology*. New Delhi: Mittal Publications.
- Davies, Mark. (2012). Expanding horizons in historical linguistics. *Corpora*, 7 (2): 121-157.
- Davies, Mark. (2020). the Corpus of Contemporary American English. https://www.english-corpora.org/coca/help/coca2020_overview.pdf
- Harb, G. (2016). The Application of COCA Corpus in Vocabulary Instruction to Assist Lebanese Students Improve their Writing: A Case Study in Les Doctoriales 2016, pp. 19-39, Kaslik: PUSEK
- Gries, S.T. (2009). What is Corpus Linguistics? *Language and Linguistics Compass*, 3(5): 1-17, <https://doi.org/10.1111/j.1749-818X.2009.00149.x>
- Kennedy, Graeme. (2001). Review of “An Introduction to Corpus Linguistics. *Studies in Language*, 25(1):163-167. <https://doi.org/10.1075/sl.25.1.10leu>
- Leech, G. (1992): "Corpora and theories of linguistic performance", *Directions in Corpus Linguistics: Proceedings of Nobel Symposium 82*, ed. J. Svartvik. Berlin; New York, NY: Mouton de Gruyter. 105-122.
- Xie, Qin. (2013). Corpus Linguistics and Corpus-Based Research in Hong Kong: A State-of-Art Review. *English Language and Literature Studies*, 3(3): 48-55.
-