

Using Brain Imaging to Gauge Difficulties in Processing Ambiguous Text by Non-native Speakers

Imtiaz Hussain Khan

Department of Computer Science
King Abdulaziz University Jeddah Saudi Arabia

Abstract—Processing ambiguous text is an ever challenging problem for humans. In this study, we investigate how native-Arabic speakers face problems in processing their non-native English language text which involves ambiguity. As a case study, we focus on prepositional-phrase (PP) attachment ambiguity whereby a PP can be attached to the preceding noun (aka low attachment) or the preceding verb (aka high attachment). We setup an experiment in which human participants read text on a computer screen and their brain activity is monitored using near infrared spectroscopy. Participants read two types of text: one involving PP-attachment ambiguity and the other unambiguous text which is used as a control for comparison purposes. The brain activity data for ambiguous and control text are clustered using hierarchical-clustering technique available in Weka. The data reveal that Arabic speakers face more difficulty in processing ambiguous text as compared to unambiguous text.

Keywords—Prepositional-phrase attachment ambiguity; near-infrared spectroscopy; Arabic speakers; hierarchical clustering

I. INTRODUCTION

Ambiguity is a challenging problem for human when they process natural language for different tasks, including text comprehension, text summarization or machine translation. State-of-the-art research in natural language processing suggests that people use different strategies in comprehending potentially ambiguous text [1, 2, 3, 4, 5, 6]. Research also suggests that generally human are good at comprehending potentially ambiguous text because of the availability of the context in which the text is processed or because they use their common sense knowledge in an effective manner to disambiguate text. At present, English is considered as a universal language for communication worldwide. A large body of the existing work focusses on how human process their native language, but there is very little work reported in literature on how non-native speakers process their second language, which involves ambiguities. This is an important problem to address in a systematic manner because there is more than one-third population whose native language is not English.

In the Arabic world, English language is commonly used as the second language, mostly as a medium of instruction in educational institutions or among the business community. Therefore, it is imperative to study systematically how native Arabic speakers process English language, especially when the language text involves ambiguities. This is important because once we are able to understand how Arabic speakers process English language, it would be easier to develop sophisticated language processing tools to aid Arabic

speakers. For instance, a personalized language model could be built to inform the design of, for example, a paraphrasing system to produce text which is easy to comprehend as well as fluent for a special group of readers.

As a starting point, in this study, we undertake the task of how Arabic speakers process English language text, which involves one of the most notorious ambiguities in almost every human language called prepositional-phrase (PP) attachment ambiguity. This is the first study of its kind which thoroughly investigates the problem using brain imaging by near infrared spectroscopy (NIRS) [7, 8]. Briefly, we conducted an empirical study in which human participants read text and their brain activity was monitored using NIRS to gauge the difficulty they faced during the text reading task. Participants read two types of texts, one involving PP-attachment ambiguity and the other unambiguous text, which is used as a control for comparison purposes. The brain activity data for ambiguous and control text were recorded. Subsequently, the NIRS data are clustered using hierarchical-clustering technique to discriminate between the processing of ambiguous and unambiguous texts. We found that Arabic speakers face significant difficulty in processing ambiguous text as compared to unambiguous text; our NIRS results corroborate with the previous findings, which are mainly based on self-paced reading paradigm.

The rest of this article is organized as follows. In Section II, state-of-the-art research work is presented. Section III describes the empirical study. Results are discussed in Section IV. Finally, the article concludes in Section V with pointers and open questions for future work outlined in Section VI.

II. RELATED WORK

Ambiguity is an interesting phenomenon which has received considerable attention by researchers in language studies [1, 3, 2, 9, 10]. Among these researchers, in an earlier interesting work [2], the authors proposed minimal attachment and right association principles. These principles have been dominating the field for many years and the authors argued that they are effective in resolving ambiguities of different kinds, especially PP-attachment ambiguity. Other notable contribution which received much attention is by [10] who proposed probability based constraint satisfaction theories for structural ambiguity resolution. Their work paved the way for many modern

corpus-based approaches to resolve ambiguities of different kinds.

PP-attachment ambiguity has thoroughly been investigated in literature and different approaches, including corpus based [4, 11], statistical [12, 13], and machine learning [14, 15] have been proposed to resolve it. Some interesting work is reported in literature where researchers have tackled such ambiguities in Arabic language. Notable among them is work done by [16, 17, 18, 19]. In these studies, the authors have primarily tried to resolve different ambiguities, including structural ambiguity in Arabic language. These are important contributions, for example [19] used a corpus-based approach to resolve PP-attachment ambiguity in Arabic. However, little attention has been paid to investigate how people, especially Arabic speakers process text involving PP-attachment ambiguity in English as their second language. One notable exception [20], where authors used a sample size of 60 Arabic native speakers to investigate how they interpret English language text involving different types of ambiguities. Unlike our methodology, they asked the participants to translate the given text, involving ambiguity, in to Arabic language. The number of sentences in their experiment was only 18, which arguably a small sample size keeping in view the format of their study. They found that participants faced significant difficulties in translating all type of ambiguous sentences.

Some other interesting work is reported in literature where researchers tried to investigate how people generally process their second language. In one such work [21], the authors reported on an empirical study on Greek students to examine the difficulties they face while learning English as their second language. They found that Greek students who were facing difficulties in their native language also face difficulties in their second language, thus concluding that there is a strong linkage between processing native and non-native language. In another related work [22], the authors investigated in a cognitive-based experimental setting the effects of their native language on the non-native language. They showed that those constructions which are governed by the same linguistic rules, they exhibit similar neural processing. However, the lexically-driven constructions, which are different, similar neural activity does not occur for them. In [23], the authors report on a detailed literature review showing the effects of first language on the second language. They have observed that by and large researchers are agreed that the role of the native language in learning the second language depends on some similarities and dissimilarities between the two languages. In [24], the author investigated how the Spanish students learn and process the new content word in English as their second language. The limitation of their work is that they do not report convincingly on their findings, so it is difficult to assess their contribution. In [25], the researchers explored three different hypotheses for relative clause processing in the learners who learnt English as their second language. They observed a complementary relationship among the hypotheses in processing the relative clauses. In an interesting work [26], the authors studied the cognitive aspect of learning English as their second language. They found that those students who

take English as their second language as medium of instruction, their reasoning capabilities in science-based subjects dropped significantly. In [27], the researchers investigated that second language learners perform better in judging long-distance object extraction than subject extraction. They conducted their study with a medium sample of 25 Chinese students whose second language was English. They found that it is parsing difficulty and lack of grammatical understanding which makes the sentence processing task challenging to the second language learners.

NIRS is a popular noninvasive brain-imaging technique [8], which has successfully been used in different studies, including language processing task [28, 29, 30, 31]. It uses light in a very narrow band (700-1000 nm), which propagates through different tissues in the human body and ultimately either it is absorbed or scattered. In the brain, changes in the optical absorption are related to the changes in both oxygen hemoglobin (oxy-Hb) and deoxygenated hemoglobin (deoxy-Hb). The multiple wavelengths' changes in the absorption of light can be used to estimate the concentration of oxy-Hb and deoxy-Hb, thereby localizing the brain areas, which were active or otherwise against a given task like language comprehension. Such changes in the brain concentration are assumed to be due to human's response to a given stimulation task, for example finger tapping, called functionally-evoked responses. In literature, it is very little reported how such changes in brain's concentration are happening when people process ambiguous text. Therefore, our study is timely and it is expected that our findings will pave way for further investigation in this interesting area.

III. EMPIRICAL STUDY

We conducted an experiment with human participants whose task was to read text on a computer screen. While the participants were reading the text, their brain activity was monitored using NIRS. Each participant completed 24 trials, where a trial consisted of a target sentence followed by a comprehension question. The trials were presented to the participants one trial at a time fashion, where all participants received all trials in exactly the same order.

A. Materials and Methods

We constructed a balanced dataset of 24 trials where 12 sentences involved potentially ambiguous PP-attachment noun phrases and the remaining 12 sentences were unambiguous; the latter are used as control for comparative analysis. A comprehension question was also associated with each sentence to ensure that participants carefully read the text to make sense of it. For instance, in one trial, the target sentence was 'Hassan saw Bilal with the telescope and then went away', and the comprehension question was 'Did Bilal held the telescope?'. For the purpose of brain imaging during the reading task, we opted for a block-design paradigm, in which a Rest-Task-Rest pattern is used. Since our task consisted of reading a piece of text followed by a comprehension question, we adapted the design as Rest-Task-Task, where first task was associated with reading the text and the second task was associated to answer the comprehension question. The task protocol setting is shown

in Fig. 1, which indicates that 5 seconds each are assigned to rest, the first task and the second task, respectively. The 5 seconds for reading the target sentence and comprehension question each were informed by a self-paced reading pilot study before running this experiment proper. This setting requires 15 seconds to complete a trial, and a total 360 seconds (24 trials x 15 seconds, i.e. 6 minutes) to complete the experiment for one participant.

For the brain-imaging purpose, we used Shimadzu's 44-channel LABNIRS system, with its default parameter settings, for example, the sampling time was 26 milliseconds for data acquisition. We used the well-known international 10-20 reference system [32] (as shown in Fig. 2) for positioning the channels (i.e., emitter-detector pair). Even though the channels were covering the area between occipital and temporal lobes, we are primarily interested in those channels that were positioned on Wernicke's area (as shown in Fig. 3, encircled area # 11) as it is associated with language comprehension.

Fig. 4 depicts a sample of the experiment setting during NIRS data acquisition while the participant is in the rest mode, i.e., with the blank computer screen. The NIRS computer screen shows the time-course for concentration in oxy-Hb, deoxy-Hb, and average of oxy-Hb and deoxy-Xb, which is extracted for a selected channel and shown in Fig. 2 for further analysis.

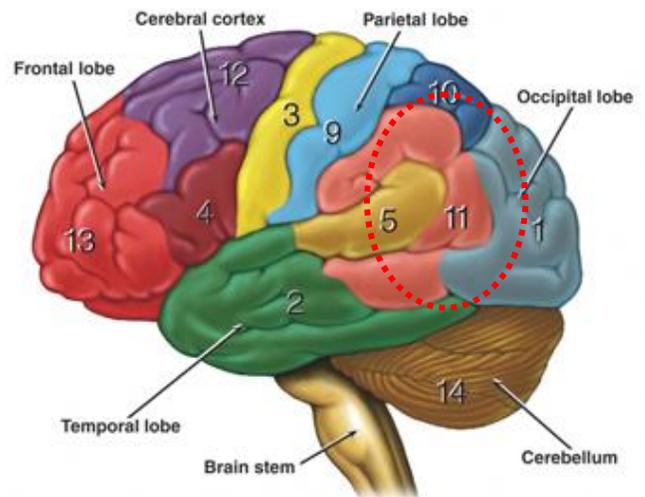


Fig. 3. Functional Areas of Human Brain.

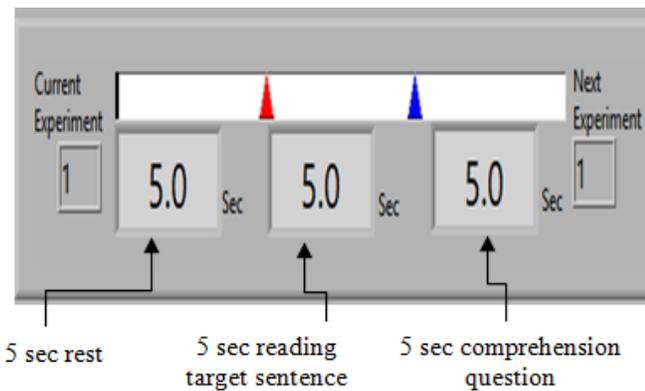


Fig. 1. Task Protocol.

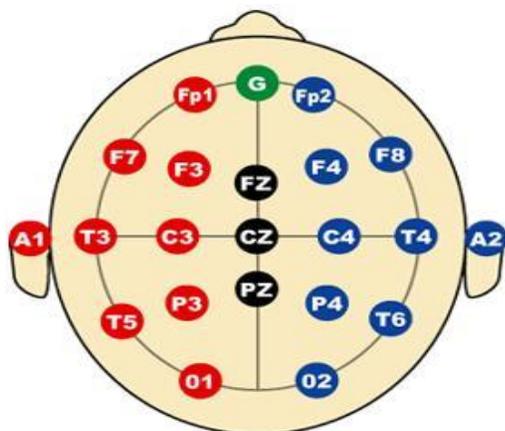


Fig. 2. International 10-20 System.

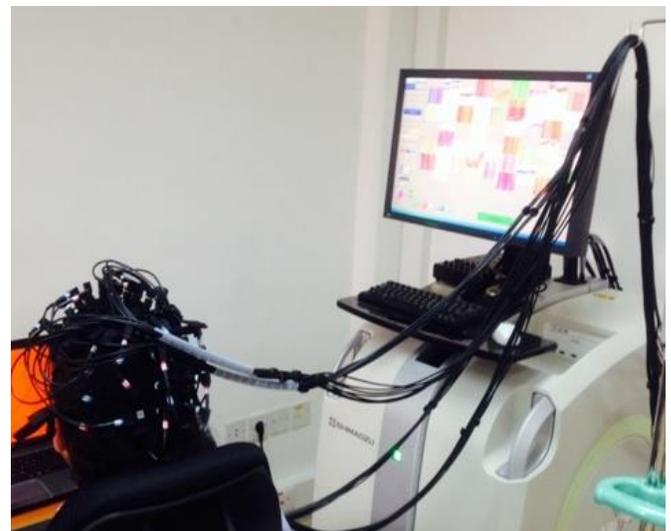


Fig. 4. Data acquisition using LABNIRS

B. Participants and Procedure

Twelve native-Arabic speakers who were self-reportedly fluent in English took part in this study. The participants were male students having normal vision with average age of 20 years. The study was conducted in accordance with the university code of ethics and all participants signed consent before the experiment. Before conducting the experiment, they were briefed about the purpose of the study and any effects of NIRS brain imaging. Before running the experiment proper, participants were presented with a sequence of 5 practice trials to familiarize themselves with the experiment procedure. The data from the practice trials were not recorded for analysis purposes. Participants were also briefed that they can willingly withdraw from the experiment at anytime if they want to; however, all the participants complete all the experiment trials.

Participants were presented with a sequence of trials where each trial began with a 5-second rest period during which participants were just relaxing. After the rest period, the target sentence appeared in a customized moving window fashion, in three parts. That is, first, a sequence of n words appeared on the screen for 1 second, then another sequence of m words appeared on the screen for 1 second and finally the last sequence of k words appeared on the screen for 1 second. For example, in the above example target sentence 'Hassan saw Bilal with the telescope and then went away', 'Hassan saw Bilal' appeared first for 1 second, then 'with the telescope' appeared and finally 'and then went away' appeared. This moving window fashion was driven by necessity in this study as we were interested more in the brain activity just after processing the ambiguity-inducing part (i.e., the second part 'with the telescope', in this example). The whole target sentence stayed on the screen for 5 seconds, which was sufficient time to comprehend the text. After the target sentence, the comprehension question appeared for 5 seconds. The participant had to answer the comprehension question by pressing left (Yes) or right (No) arrow key, which they were briefed as part of the instructions.

IV. RESULTS AND DISCUSSION

We recorded the NIRS data for 10 channels at a sampling rate of 26 milliseconds. Each channel recorded both oxy-Hb and deoxy-Hb, but we used only oxy-Hb for analysis purposes as studies suggest that oxy-Hb and deoxy-Hb generally complement each other. Each channel recorded the time-series data for 24 trials (12 ambiguous sentences and 12 control sentences). Since NIRS provides markers to mark the end of a particular trial, we extracted the time series data for each trial from each channel for data-mining purposes. Additionally, NIRS also provides 3 markers to separate the Rest-Task-Task periods, we also extracted another set of data for 5 seconds, which covers the brain activity during which the target sentence was onset; these data are used for statistical analysis by using a paired t-test.

We used the hierarchical clustering in Weka for clustering the trials. This clustering technique is widely used in literature for data mining of time-series data, especially when the data size is reasonably large. There were a total of 2880 rows (= 10 channels x 24 trials x 12 participants) of time-series data. Initially, hierarchical clustering considers all 2888 data points as separate clusters and then it keeps merging them on a similarity basis until some stopping criterion is met. Since in our case, there are two types of sentences, ambiguous and unambiguous, the clustering process stopped when there were only two clusters left in the hierarchy. The results elucidate that the hierarchical clustering was able to achieve 87% accuracy in discriminating between the brain activity required for processing ambiguous and unambiguous texts.

Fig. 5 shows the brain-activity data, where high peaks of oxy-Hb are associated to the time span when an ambiguous sentence was onset. Fig. 6 shows the brain-activity data for a channel that was positioned in the brain location, which does not involve in language related tasks. We used such channels

as controls to ensure that NIRS data can be faithfully used to observe brain activity in certain regions of the brain and contrast them from other regions that remain passive during a certain task. It is clear from Fig. 5 that as expected, for the control sentences, oxy-Hb values are relatively higher than the rest period; however, these oxy-Hb values are significantly lower than the ones corresponding to ambiguous sentences.

Descriptive statistics of the oxy-Hb concentration during the task is shown in Table I. It is clear from Table I that as shown in Fig. 5, participants used more effort (mean oxy-Hb ≈ 0.027) in processing ambiguous sentences than the unambiguous sentences (mean oxy-Hb ≈ 0.018). The mean oxy-Hb values for the rest period (oxy-Hb ≈ 0.011) is less than reading both ambiguous and unambiguous sentences. A paired t-test was administered over the oxy-Hb peak values. We found that the difference between oxy-Hb for ambiguous and control text is significant ($p < 0.05$).

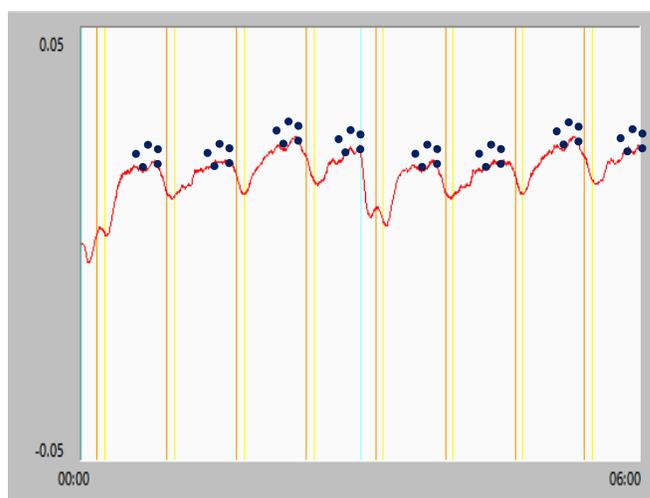


Fig. 5. Time-Course for Concentration in oxy-Hb for the Select Channel Indicating Brain Activity (Encircled Peaks Correspond to Brain Activity when a Trial Related to an Ambiguous Sentence was Onset).

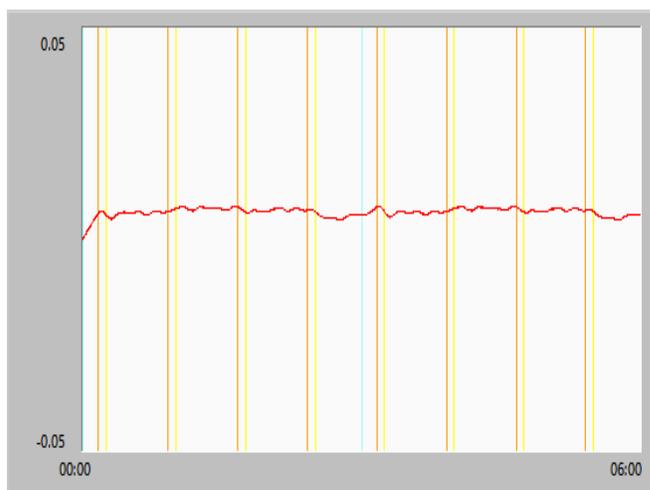


Fig. 6. Time-Course for Concentration in oxy-Hb for the Select Channel Indicating no Brain Activity in the Brain Region where the Channel was Positioned.

TABLE I. DESCRIPTIVE STATISTICS OF THE OXY-Hb CONCENTRATION DURING THE TASK

Task	Mean oxy-Hb	Standard deviation
Reading ambiguous sentence	0.0269	0.0107
Reading unambiguous sentence	0.0183	0.0096
Rest	0.0107	0.0018

These results suggest that Arabic speakers face significant difficulty while processing ambiguous text as compared to unambiguous text. Therefore, it is important to use sophisticated techniques to avoid (or manage) ambiguity while authoring English language documents for Arabic speakers who learn or use English as their second language. Our findings also corroborate with previous findings that non-native speakers generally face difficulty in processing their second language.

One limitation of our work is that we explored a small dataset and the number of participant was also relatively small. However, depending on the nature of the study this limitation is not very much daunting as in such studies it is generally difficult to recruit large number of participants and also it is difficult to manage such experiment when the number of trials increases. The latter is true because with six minutes of NIRS data recoding for a participant, the size of one NIRS generated data goes beyond 35 MB; so with large number of participants and the large number of trials it would be very hard to analyze the data in a meaningful way.

V. CONCLUSION

In this study, we investigated empirically how native-Arabic speakers process English as their second language, especially when English language text involves ambiguity. As a case study, we focused on PP-attachment ambiguity, which is a challenging task for human in processing language. We conducted an empirical study with 12 native Arabic speakers who were fluent in English language. Their task was to read two types of texts, one involving PP-attachment ambiguity and the other unambiguous text (control), during which their brain activity was also monitored using NIRS to gauge the difficulty they faced during the reading task. The NIRS data for ambiguous and control text were recorded and subsequently analyzed using hierarchical-clustering technique to discriminate between the processing of ambiguous and unambiguous text. We found that Arabic speakers generally face difficulty in processing ambiguous text. Our findings are consistent with previous findings in this area, especially using self-paced reading paradigm with native speakers.

VI. FUTURE WORK

In future, we intend to extend this study by using other types of ambiguities as well on relatively a large pool of dataset. We also intend to pursue a self-study paradigm, with non-native English speakers, to examine whether self-paced reading data can be trustfully used to predict human brain activity, which otherwise needs more expensive and difficult techniques to be administered like brain imaging using NIRS or functional magnetic resonance imaging (aka fMRI).

ACKNOWLEDGMENT

This work was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under grant No. 611-010-D1433. The author, therefore, acknowledge with thanks DSR technical and financial support.

REFERENCES

- [1] G. T. Altmann, "Ambiguity in sentence processing.," Trends in cognitive sciences., vol. 2, no. 4, pp. 146-152, 1998.
- [2] L. Frazier and K. Rayner, "Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences.," Journal of memory and language., vol. 26, pp. 505-526, 1987.
- [3] F. Ferreira and C. Clifton, "The independence of syntactic processing.," Journal of memory and language., vol. 25, no. 3, pp. 348-368, 1986.
- [4] D. Hindle and M. Rooth, "Structural ambiguity and lexical relations.," Computational Linguistics, pp. 103-120, 1993.
- [5] J. Kimball, "Seven principles of surface structure parsing in natural language.," Cognition., vol. 2, pp. 15-47, 1973.
- [6] C. Phillips, "Right association in parsing and grammar.," in Proceedings in language processing and acquisition., 1995.
- [7] H. Heekeren, H. Obrig, R. Wenzel, K. Eberle, J. Ruben, K. Villringer, R. Kurth and A. Villringer, "Cerebral haemoglobin oxygenation during sustained visual stimulation: a near infrared spectroscopy study.," Philosophical Transactions of the Royal Society of London B: Biological Sciences, vol. 352, no. 1354, pp. 743-750, 1997.
- [8] A. Villringer and B. Chance, "Non-invasive optical spectroscopy and imaging of human brain function.," Trends Neurosci, vol. 20, no. 10, pp. 435-442, 1997.
- [9] S. Stevenson, "A Competition-based explanation of syntactic attachment preferences and garden path phenomena.," in Proceedings of the 31st annual association for computational linguistics., 1993.
- [10] J. C. Trueswell and M. K. Tanenhaus, "Toward a lexicalist framework of constraint-based syntactic ambiguity resolution.," in Proceedings in perspectives on sentence processing., Hillsdale, NJ, US, 1994.
- [11] P. Nakov and M. Hearst, "Using the web as an implicit training set: application to structural ambiguity resolution.," in Proceedings of the conference on human language technology and empirical methods in natural language processing., 2005.
- [12] A. Ratnaparkhi, J. Reynar and S. Roukos, "A maximum entropy model for prepositional phrase attachment.," in Proceedings of the ARPA human language technology workshop., 1994.
- [13] M. Collins and J. Brooks, "Prepositional phrase attachment through a backed-off model.," in Proceedings of the third workshop on very large corpora., 1995.
- [14] S. Zhao and D. Lin, "A nearest-neighbor method for resolving PP-attachment ambiguity.," in Proceedings of the first international joint conference on natural language processing (IJCNLP-04)., 2004.
- [15] M. Olteanu and D. Moldovan, "PP-attachment disambiguation using large context.," in Proceedings of human language technology conference and conference on empirical methods in natural language processing (HLT/EMNLP)., 2005.
- [16] K. Daimi, "Identifying syntactic ambiguities in single-parse Arabic sentence.," Department of mathematics and computer science, University of Detroit Mercy, 2001.
- [17] E. Othman, K. Shaalan and A. Rafea, "A chart parser for analyzing modern standard Arabic sentence.," in MT summit IX workshop on machine translation for semitic languages: issues and approaches, New Orleans, Louisiana, USA, 2003.
- [18] E. Othman, K. Shaalan and A. Rafea, "Towards resolving ambiguity in understanding Arabic sentence.," in Proceedings of international conference on Arabic language resources and tools, 2004.

- [19] R. Al-sabbagh and K. Elghamry, "A Web-based approach for Arabic PP-attachment.," in Proceedings of the 6th international conference on informatics and systems., Cairo, Egypt, 2008.
- [20] M. I. Khawalda and E. M. Al-Saidat, "Structural ambiguity interpretation: a case study of Arab learners of English.," Global journal of human social science., vol. 12, no. 6, 2012.
- [21] G. Andreou and M. Segklia, "Learning difficulties in first and second language: preliminary results from a cross-linguistic skills transfer.," English linguistics research., vol. 6, 2017.
- [22] L. Sabourin and L. Stowe, "Second language processing: When are first and second languages processed similarly?," Second language research., vol. 24, pp. 397-430, 2008.
- [23] P. Durán Escribano, "Exploring cognition processes in second language acquisition: the case of cognates and false-friends in EST.," Ibérica revista de la asociación Europea de lenguas para fines específicos., pp. 87-106, 2004.
- [24] A. Derakhshan, "The interference of first language and second language acquisition.," Theory and practice in language studies., vol. 5, no. 10, pp. 2112-2117, 2015.
- [25] S. Izumi, "Processing difficulty in comprehension and production of relative clauses by learners of English as a second language.," Language learning., vol. 53, no. 2, pp. 285-323, 2003.
- [26] A. H. Johnstone and D. Selepeng, "A language problem revisited.," Chemistry education: Research and practice in Europe., vol. 2, no. 1, pp. 19-29, 2001.
- [27] A. Juffs and M. Harrington, "Parsing effects in second language sentence processing: subject and object asymmetries in wh-extraction.," Studies in second language acquisition., vol. 17, no. 4, pp. 483-516, 1995.
- [28] A. Ehlis, M. Herrmann, M. Plichta and A. Fallgatter, "Cortical activation during two verbal fluency tasks in schizophrenic patients and healthy controls as assessed by multi-channel near-infrared spectroscopy.," Psychiatry Research., vol. 156, no. 1, pp. 1-13, 2007.
- [29] B. Ambridge and C. F. Rowland, "Experimental methods in studying child language acquisition.," Cognitive science., 2015.
- [30] O. Takeuchi, M. Ikeda and A. Mizumoto, "The cerebral basis for language learner strategies: A near-infrared spectroscopy study.," Reading in a foreign language. , vol. 24, no. 2, pp. 136-157, 2012.
- [31] M. A. Hall, Temporal mapping and connectivity using NIRS for language related tasks., Florida international university., 2012.
- [32] V. Diekmann, W. Becker, R. Jürgens, B. Grözinger, B. Kleiser, H. Richter and K. Wollinsky, "Localisation of epileptic foci with electric, magnetic and combined electronic models.," Electroencephalogr Clin Neurophysiol, vol. 106, pp. 297-313, 1998.