September 2018

Discourse-Pragmatic and Processing-Related Motivators of the ordering of Reason Clauses in an Academic Corpus

Abbas A. Rezaee
University of Tehran, Iran, aarezaee@ut.ac.ir

Majid Nemati
University of Tehran, Iran, nematim@ut.ac.ir

Seyyed Ehsan Golparvar
University of Tehran, Iran, segolparvar@ub.ac.ir

DISCOURSE-PRAGMATIC AND PROCESSING-RELATED MOTIVATORS OF THE ORDERING OF REASON CLAUSES IN AN ACADEMIC CORPUS

ABBAS A. REZAEE
University of Tehran, Iran
aarezaee@ut.ac.ir

MAJID NEMATI
University of Tehran, Iran
nematim@ut.ac.ir

SEYYED EHSAN GOLPARVAR
University of Tehran, Iran
segolparvar@ub.ac.ir

Abstract
The present research is aimed at examining the relative importance of the competing motivators of the sequencing of reason clauses in a corpus of research articles of applied linguistics. All the finite reason clauses accompanied by their main clauses in this corpus were collected. Random forest of conditional inference trees is the statistical modelling in this study. The findings showed that sentence-final reason clauses outnumber sentence-initial ones. Moreover, subordinator choice and bridging, which are discourse-pragmatic constraints on clause positioning, emerged as the two more powerful predictors of the ordering of reason clauses in this corpus. Furthermore, the complexity of the clause turned out to be a stronger processing-related predictor than the length of the clause.

Keywords: positioning, reason clauses, subordinator, bridging, complexity

1. Introduction

Adverbial clauses are primarily positioned in initial and final slots (Aarts, 1988; Kirk, 1997; Diessel, 2001; Givón, 2011), each of which serve distinctively different functions in discourse, including academic discourse. Adverbial clauses that are sentence-final regularly have a local function: elucidating the situation of their matrix clause by specifying reasons, temporal circumstances, results, etc. Further, post-posed adverbial clauses are mainly unidirectional, i.e., they are associated with their main clauses that have been already mentioned. In addition, final adverbial clauses offer information that is more integrated with the main clause at the local level (Thompson, Longacre, & Hwang, 2007). Moreover, these adverbial clauses are often in the middle position of a paragraph, that is, final

adverbial clauses are consistently in the middle of a thematic chain which is tightly coherent (Givón, 2001). With respect to semantics, the information provided by sentence-final clauses is often in line with the information offered in clauses in coordination (Thompson, 1985; Ford, 1993; Givón, 2001).

Sentence-initial adverbial clauses, however, do not often have such a limited local function, but play a wider function in the organization of discourse, by introducing a new frame for the discourse that follows or connecting it back to the discourse that has come. Furthermore, the cohesive function of pre-posed adverbial clauses may realize at different levels, from the whole discourse to inter-paragraph and inter-sentential levels. The inter-sentential function can be considered as a local back-referencing function that establishes a close connection between two sentences, while the higher-level function marks the episode boundary or thematic discontinuity. It should be noticed that either local or global, the function of pre-posed adverbial clauses tends to be bidirectional. To put it differently, these clauses link what has been stated before to what is to be expressed. In addition, the semantic information that is offered by pre-posed clauses carries less significance since they regularly repeat or provide predictable information from what has already been mentioned (Thompson et al., 2007).

Therefore, the two sequencing patterns are not interchangeable in academic discourse and writers of research articles should know when to use each in their texts. Hence, exploring the factors that condition the ordering of subordinate adverbial clauses and the relative importance of these factors will provide us with fresh insight into the use of reason clauses in academic discourse.

The present study aims at investigating the constraints on the positioning of finite reason clauses in a corpus of research articles of applied linguistics. Further, this research purports to measure the weight of processing-related and discourse-pragmatic constraints on the ordering of finite clauses of reason by means of random forest modelling, which has been shown to be more efficient than ordinary regression models (Rezaee & Golparvar, 2017; Tagliamonte & Baayen, 2012; Wiechmann & Kerz, 2013).

2. Background

The sequencing tendency of adverbial clauses in English has been investigated by two approaches. The first approach postulates that the ordering of linguistic items, including finite adverbial clauses, is mainly influenced by information structure. Proponents of this line of research (Chafe, 1984; Birner & Ward, 1998; Greenbaum & Nelson, 1996) have put forward the argument that speakers and writers tend to produce new, inaccessible information reflected in the main clause after given, accessible information which is reflected in the dependent clause.

The users of a language usually put adverbial clauses in the initial slot in light of two factors, namely the 'bridging' function and the 'setting the stage' function. Adverbial clauses in the final slot play local functions, whereas sentence-initial

adverbial clauses have discourse-organizing functions. Establishing the link with the previous discourse or suggesting new frames for upcoming discourse are instances of discourse-organizing functions (Ford, 1993; Verstraete, 2004; Thompson et al., 2007; Givón, 2011).

In line with Wiechmann & Kerz (2013), in the present study, we solely examine one discourse-organizing function, which is bridging, referring to a context in which a sentence-initial adverbial clause serves a bridge-like function connecting the preceding and the upcoming discourse. The presence of an anaphoric item in a sentence-initial adverbial clause indicates the bridging function in that clause. In example (1), the underlined part is a sentence-initial reason clause and the anaphoric item **she** plays a bridging function, linking the sentence with the preceding discourse.

(1).
   "To find out why the teacher did the activities or made the choices recorded during the observations, a follow-up interview was held with the teacher. Since **she** was not aware of the specific research questions, in an unstructured interview she was asked to explain 'how' she taught the course and comment on course objectives, materials, in-class teaching and testing activities." (Saif, 2006, p. 20)

The other constraint explored in this line of research is the semantic nature of the subordinate clause. The semantic disparity detected among different types of adverbial clauses (i.e., adverbial clauses of time, condition, concession, and reason) leads them to assume different positions in a complex sentence. (Quirk et al., 1985; Biber et al., 1999; Diessel, 2005). Diessel (2001, 2005) found that conditional clauses are regularly pre-posed, clauses of cause are usually post-posed, and temporal clauses to be roughly equally divided between the two ordering patterns. In a similar vein, Diessel (2001) showed that adverbial clauses of reason and purpose largely follow their matrix clauses. Adverbial clauses of concession show a modest preference for the final slot (Biber et al., 1999; Diessel, 2001; Wiechmann & Kerz, 2013). Subtle meaning differences exist between clauses that are introduced by different subordinators. Thus, any subordinator selected for dependent clauses is viewed as a motivator of the sequencing of adverbial clause (Wiechmann & Kerz, 2013). For instance, concessive clauses introduced by ALTHOUGH are usually sentence-initial, while clauses headed by WHEREAS are mainly sentence-final (Wiechmann & Kerz, 2013).

The second approach trying to explain the sequencing of dependent clauses considers processing-related factors. These accounts investigate the ordering of an adverbial clause on the grounds of variables such as the relative length of the clausal string, its complexity, and its deranking status. The most famous supporter of this account is John Hawkins (Hawkins, 1994; Hawkins, 2004), who pointed out that the constituent order is mainly determined by processing difficulty. He has asserted that information structure matters only when two alternative orders are equally demanding with respect to processing.

The first factor conditioning the ordering of reason clauses is the length of the clause. Previous empirical research has vividly demonstrated that in languages like English longer constructions regularly come after shorter ones (Quirk et al., 1985). This tendency can be explained in light of the assumption that the online processing of the whole structure appears to be more efficient with this order (Hawkins, 1994; Hawkins, 2004; Gibson, 1998; Gibson, 2000). Based on Hawkins' performance-based theory of constituent ordering (Hawkins, 2004), constituents that are perceived as heavy tend to be placed in the final slot, because this pattern of sequencing is cognitively more efficient in languages that are head-initial, causing both production and comprehension to be easier.

In a similar vein, the dependency locality theory proposed by Gibson (1998, 2000) postulates that the processing complexity of a linguistic string is contingent upon the length of its syntactic dependencies. The ordering complexity effects are associated with the integration cost component which proposes that longer distance attachments are more expensive to make than shorter distance ones (Bever, 1970). Adverbial clauses of reason that are pre-posed introduce longer dependencies and are thus more demanding to process.

A pragmatic, information-structural account can also provide an explanation for the trend of 'lighter' constituents to precede 'heavier' ones based on the 'given-new' principle (Arnold et al., 2000), assuming that new information, in comparison with given information, requires more linguistic materials to be encoded. Discourse-pragmatic explanations have also demonstrated that for clauses and multi-clause constructions, the informativeness increases towards the end of each construction. Thus, length is a salient predictor of positioning of adverbial clauses of reason.

The other constraint on the ordering of reason clauses that is associated with processing difficulty is complexity. Several definitions have been proposed for complexity such as relative complexity (see Dahl, 2004; Vulanovic, 2007), absolute complexity (see (Miestamo, 2004), language complexity (Hawkins, 1994; Hawkins, 2004), and complexity with respect to informativeness (Li & Vitányi, 1997). Adverbial clauses can be complex in different degrees. It may be thought that sentence-initial adverbial clauses of reason are structurally less complex. Following Diessel (2008) and Wiechmann and Kerz (2013), in this study we regard as complex only those reason clauses that involve another subordinate clause of any kind. It should be noticed that there exists a close connection between linguistic complexity and the length of adverbial clause. Reason clauses that have another subordinate clause – complex reason clauses - tend to be longer and therefore are more burdensome to process. Consequently, we can make the assumption that complex adverbial clauses of reason are generally post-posed.

3. Method

3.1. Corpus

In order to conduct this study, a corpus of 100 research articles of applied linguistics were utilized. All the articles selected were written by native speakers of English, determined by the authors' affiliation. There are 801 tokens of reason clauses in this corpus. All the articles that are incorporated in this corpus deal with applied linguistics and language teaching and learning. The article length has not been considered as a variable. It should be noticed that the corpus of this research will include articles which are published from 2001 to 2014. All the journals used to collect the corpus are peer-reviewed both in terms of content and language. Ten articles were randomly selected from each journal. The title of these ten journals are as follows: Annual Review of Applied Linguistics, Applied Linguistics, ESP Journal, EAP Journal, Language Learning, Language Teaching Research, System, Second Language Research, language Testing, TESOL Quarterly.

3.2. Variables

The dependent variable in this study is the ordering of adverbial clauses of reason that is measured as a binary factor having two levels that are final (POS 1) and initial (POS 0). In addition, the predictive variables are subordinator, bridging, length, and complexity. Subordinator is a nominal variable with two levels, namely BECAUSE (SUB 0) and SINCE (SUB 1). According to Quirk et al. (1985), these two subordinators are the most frequent reasons subordinators in academic register.

Bridging is a categorical variable with two levels of having an anaphoric item suggesting a bridging context (BRG 0) and absence of such an item (BRG 1). Complexity is also a binary variable with two categories that are simple (COM 0) and complex (COM 1). Finally, length (LNG) is measured on a continuous scale which is defined as the proportion of the length of the reason clause to that of the whole complex sentence involving that clause. It should be noted that there were no instances of deranked reason clauses in this corpus; therefore, deranking, which is one of the processing-related constraints on clause positioning in Wiechmann and Kerz (2013), was excluded from this analysis.

3.3. Data Analysis

Conditional inference trees and the random forest developed from these trees is the modelling approach utilized in this research. Forests are a collection of multiple decision trees used for the purpose of variable selection. One single decision tree is simple and capable of coping with missing values; nevertheless, it might be unstable because minor changes in the input variables may cause huge

changes in the output. Consequently, a random forest of such trees is a more robust tool for selecting variables (Breiman, 2001).

The acceptance of random forest modelling lies in the fact that it is an unbiased method for selecting variables in the individual classification trees allowing us to reliably assess the relative weight of variables which are measured on different scales or that differ as regards to the number of their factor levels. This is the scenario where traditional tree-based models have trouble and the coefficients of logistic regression models are far more complex to interpret (Wiechmann & Kerz, 2013).

The purpose of classification trees in general is to predict a typically binary outcome on the basis of a number of predictors. The algorithms related to classification trees typically work through the data and determine a set of if-then logical (split) conditions producing accurate classification of cases. In other words, in the first step, the algorithm will split the data in accordance with the most salient predictor and will continue to split each resulting subset of the data until it can no longer find statistically meaningful associations between any of the predictors and the dependent variable (Breiman, 2001; Hothorn, Hornik, & Zeileis, 2006).

The random forest, nevertheless, is not prone to these kinds of problems, although the *cost of the computational complexity increases due to bootstrap resampling and permutation-based evaluation of variable importance. A researcher adopting random forest modelling will consider all variables in their own place, and determine which of these variables turn out to be more robust predictors. In a bid to specify how the variables operate together in the random forest, a conditional inference tree can be grown which will illustrate the way different predictors interact (Hothorn et al., 2006; Wiechmann & Kerz, 2013).

Random forests build a huge number of conditional inference trees (the random forest). Each tree in the forest is developed for a subset of the data that is produced by random sampling without drawing a replacement (subsampling) from observations and predictors. The statistical metaphor is to place part of the observed data into a bag. The data that is put in the bag is called the 'in-bag' observations, while the data points that are not included in the sample are referred to as the 'out-of-bag' observations. The result of this process is that for each tree a training set (the in-bag observations) is coupled with a test set (the out-of-bag observations). The accuracy of a tree's predictions tends to be measured by drawing a comparison between its predictions for the out-of-bag observations and the actual values that are obtained for the out-of-bag observations (Hothorn et al., 2006). Figure 1 illustrates an instance of conditional inference tree modelling taken from Rezaee and Golparvar (2017).
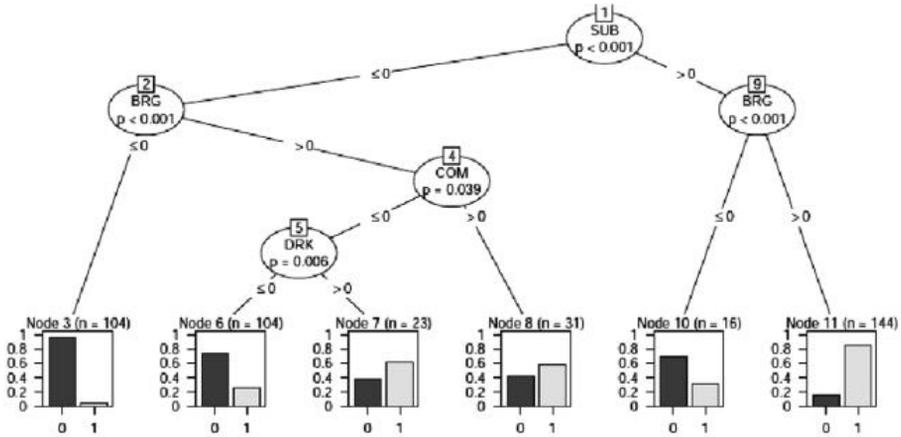
Figure 1. An Instance of Conditional Inference Tree

The analysis of this tree indicates that among the five predictors of the ordering of concessive clauses, four of them, i.e. subordinator, bridging, complexity, and deranking are significant predictors. The boxes at the bottom demonstrate the proportions of initial and final adverbial clauses in each subset, which are labeled as '0' and '1' respectively (0 represents initial clauses and 1 represents final clauses). In the white oval shapes in which the name of the variables is stated, the split variable and the p value indicating the significance level are observed. The numbers on the lines connecting the nodes of the tree suggest the particular categories of the nominal predictors or range of values of the numerical predictors (the only numerical predictor in this study is length).

4. Results

The results of this study demonstrated that a considerable proportion of adverbial clauses of reason (67.7%) are in final position and 32.3% of these clauses are sentence-initial. Moreover, the majority of them are simple (80%), have no anaphoric item suggesting a bridging context (88.2%), and are headed by *Since* (56.9%). Moreover, their average length relative to the size of the whole complex sentence is 0.45. Table 1 reports some descriptive statistics with regard to the sample.

Table 1. Descriptive Statistics for Reason Clauses

Dependent Variable	POS	Initial 32.3%	Final 67.7%
Predictors	BRG	Bridging 11.8%	Non-bridging 88.2%

|	COM	Simple	Complex
		80%	20%
	SUB	Because	Since
		43.1%	56.9%
	LNG	Mean	Standard Deviation
		0.45	0.25

The distribution of these five sequencing motivators across the two clause slots is demonstrated in figure 2. According to Figure 2, there exists a significant distribution difference between initial and final adverbial clauses with respect to subordinator and bridging. In addition, according to Figure 2, clauses having a bridging function are mostly in initial position, whereas those without a bridging context are mainly sentence-final. With regard to complexity, it is observed that in both simple and complex clauses, sentence-final clauses outnumber sentence-initial ones.
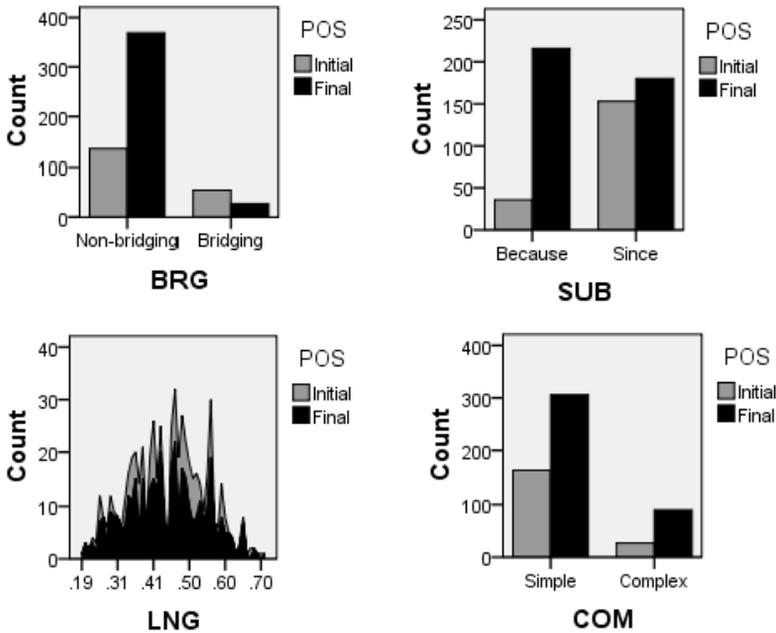
Figure 2. Distributions of the Positioning Motivators across the Two Positions of Reason Clauses

A total set of 500 trees were grown by means of bootstrapping technique, taking 500 different random subsamples from the original data The resulting model is statistically significant, indicating that three of the predictors exert a significant

effect on the positioning of reason clauses. The model shows good performance in predicting adverbial clause ordering. The index of concordance C (area under curve – ROC is 0.80) and the overall error rate of the model is 0.21. Figure 3 depicts the conditional inference tree for the positioning of reason clauses.
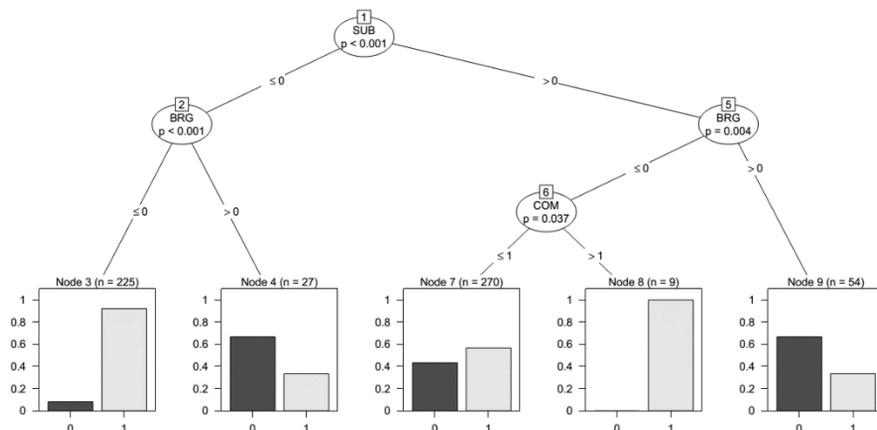
Figure 3. Conditional Inference Tree for the Positioning of Reason Clauses

Figure 3 illustrates the conditional inference tree. The analysis of this tree indicates that among the four predictors of the ordering of reason clauses, three of them, i.e. subordinator, bridging, and complexity are significant predictors. The boxes at the bottom show the proportions of initial and final adverbial clauses of reason in a given subset, which are labeled as '0' and '1' respectively (0 represents initial clauses and 1 represents final clauses). The rest of the symbols are the same as those in Figure 1.

In the first subset of the data, the first split is made based on subordinator (Node 1). The left split represents clauses of reason that are headed by *Because* (SUB ≤0) and the right one represents clauses of reason headed by *Since* (SUB > 0). Figure 3 illustrates that in both subsets of the data, a further split is made based on bridging (Node 2 and Node 5). Clauses that are headed by *Because* and do not have a bridging function (BRG ≤ 0, Node 3) are predominantly in final position. This is true for 225 cases, which is observed in Node 3. In contrast, *Because* clauses having an anaphoric item indicating a bridging context (BRG<0) are mostly sentence-initial (Node 4, 27 cases).
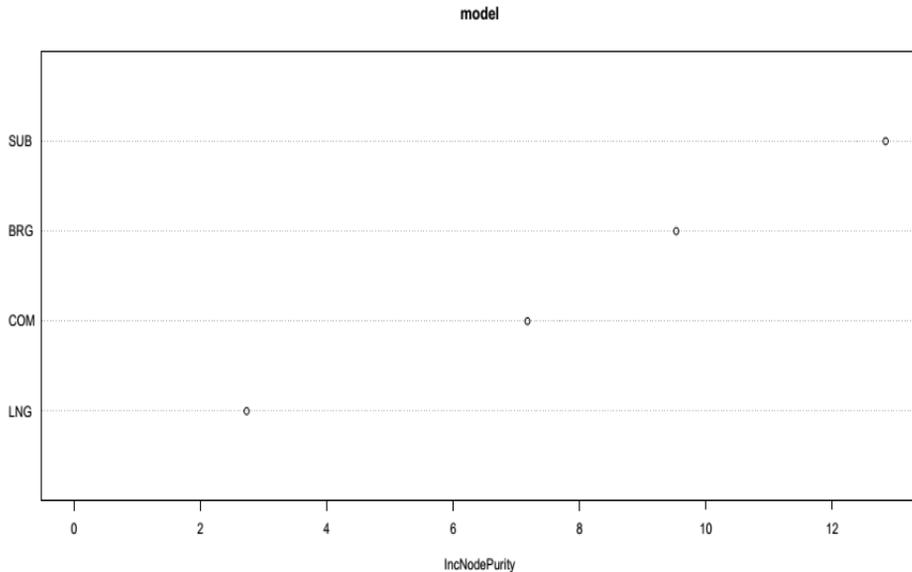
Figure 4. Variable Importance Plot for the Predictors of the Ordering of Reason Clauses

The right side of the tree depicts adverbial clauses of reason that are headed by *Since* (SUB>0). This subset of the data is further split based on bridging (Node 5). Adverbial clauses of reason beginning by *Since* with a bridging context (BRG>0) mostly precede their associate main clauses (Node 9, 54 cases), whereas those without such a function (BRG≤ 0) are further split based on their complexity (Node 6). Among clauses that are simple or complex (COM ≤ 1), sentence-final positions outnumber sentence-initial ones (Node 7, 270 cases).). Figure 4 depicts the variable importance plot for all predictors measured by the random forest model.

As shown in Figure 4, subordinator turns out to be the strongest predictor of adverbial clauses of reason, followed by bridging. Complexity turns out to be a stronger predictor than length. Finally, length has the lowest contributions to the prediction of clause ordering in this corpus of reason clauses.

5. Discussion

The analysis of a corpus of reason clauses produced by researchers of applied linguistics revealed that they tend to use these clauses in final position. This is in line with Quirk, et al. (1985), Biber et al. (1999), and Diessel (2001). Moreover, the majority of these clauses do not have an anaphoric item indicating a bridging context. In addition, only 20 percent of these reason clauses are complex,

containing another subordinate clause of any type. Moreover, clauses headed by *Because* outnumber those headed by *Since*.

The results of this study also demonstrated that *Because* clause are predominantly in final position. In addition, in reason clauses headed by *Since*, post-posed clause slightly outnumber pre-posed ones. Furthermore, in both *Because* and *Since* clauses, final clauses outnumber initial ones. It was shown that *Because* clauses are predominantly in final position, while sentence-final *Since* clauses outnumber sentence-initial ones. In other words, those reason clauses that are pre-posed are mostly headed by *Since*.

In addition, random forest modelling of conditional inference trees demonstrated that the ordering of reason clauses in a corpus of research articles of applied linguistics is firstly predicted by subordinator. To put it differently, whether the adverbial clause of reason is headed by *Because* or *Since* is the most important determinant of the sequencing of these clauses. Based on Wiechmann and Kerz (2013), the semantic disparity between reason subordinators is the most salient motivator of the positioning of reasons clauses. This finding is also in line with Rezaee and Golparvar (2017) who found that subordinator is the most powerful predictor of the sequencing of concessive clauses in a corpus of concessive clauses written by non-native speakers of English. Examples (2) and (3) illustrate this finding.

(2)

"The participants' OPI ratings were rather high even before studying abroad,
most likely because they were highly motivated, enough to opt to study
abroad." (Iwasaki, 2010, p. 50)

(3)

"*Since the test items were not discrete point but were nested within one of four tasks (each*
with their own theme), by endorsing the interactionist view of construct definition, effects
of these four themes (context) on individual items were also investigated." Vafaee, Basheer,
& Heitner (2012: 1)

The random forest of conditional inference tree modelling revealed that having an anaphoric item indicating a bridging context is the second most powerful predictor of the sequencing of these clauses in research articles of applied linguistics. This is in line with Vandepitte (1993) mentioning that the information value of the reason clause impacts both its position in relation to matrix clause and the choice of its subordinator; therefore, reason clauses offering given, recoverable information usually occur in initial position, while reason clauses presenting new, unrecoverable information are placed in final position. This is also supported by the principle of end-focus and the principle of end-weight (Quirk et al., 1985; Mukherjee, 2001) asserting that the information in a message is often processed in a way to achieve a linear presentation from low to high information value.

This finding offers support for the fact that when the function of adverbial clauses of reason is to organize the flow of information in the discourse, and their

use is impacted by factors associated with information structuring and cohesion, they are mostly placed in the initial slot (Givón, 2001; Verstraete, 2004; Diessel, 2005; Diessel, 2008; Rezaee & Golparvar, 2017; Wiechmann & Kerz, 2013). To put it differently, the anaphoric relation in the discourse is the second most significant motivator of the positioning of adverbial clauses of reason in this academic corpus. Examples (4) and (5) illustrate this finding.

(4)

 "This item required a gap in a dialogue to be filled with an utterance containing an expression which, in hindsight, even the highest ability candidates to the university were unlikely to have learned, hence possibly eliciting random guessing behaviors. <u>Since **this** anomaly had to do with the correct option of this specific item, not with the number-of-options factor</u>, it was decided to drop this item in subsequent analyses." (Shizuka et al., 2006, p.43)

(5)

"The English words were selected from the 5,000 most frequent words in Collins COBUILD corpus (Bank of English). <u>Because **these** frequencies might not apply to FL/L2 learners</u>, the selected words were checked against a word list based on EFL textbooks used in the Netherlands." (Schoonen et al., 2011, p. 45)

In (4) and (5), the underlined part is a reason adverbial clause in which 'this anomaly' and 'these frequencies' are anaphoric items indicating a bridging context. These anaphoric items and the reason clauses in which they are embedded create a link between the matrix clauses and the previous discourse. The results of this study showed that the majority of these bridging-functioning clauses are sentence-initial.

The variable that is most closely associated with processing-based explanations is complexity, which only emerged as the third predictor of ordering in reason clauses. To put it differently, adverbial clauses of reason that incorporate another subordinate clause tend to be put in sentence-final position; however, the impact of this constraint, i.e. complexity, is less than that of bridging and subordinator. This finding is in line with Wiechmann and Kerz 2013), demonstrating that processing-related factors are less powerful in predicting the positioning of adverbial clauses. This offers additional support for the assumption that the sequencing of adverbial clauses in general, and reason clauses in particular, is first and foremost determined by discourse-pragmatic motivators rather than processing-based constraints. Example (6) is an illustration of this point, in which the underlined part is a complex clause of reason and the bold part is a relative clause embedded in it.

(6)

"It is reasonable to expect working memory and short-term memory to be correlated <u>because the tasks **that measure the two constructs** are very similar</u>." (Trude & Tokowicz, 2011: 262)

6. Conclusion

This study investigated the positioning of adverbial clauses of reason in a corpus of 100 research articles published by writers of research articles of applied linguistics for whom English is considered as a native language. It was revealed that they tend to use these clauses in final position. Moreover, it was found that the ordering of reason clauses produced in this academic corpus is firstly predicted by subordinator type, and the presence of an anaphoric item indicating a bridging context is the second most powerful predictor of the sequencing of these clauses. In addition, this research lends further support for previous research on clause positioning (Diessel, 2005; Wasow, 2002; Diessel, 2008: Wiechmann & Kerz, 2013), indicating that the sequencing of adverbial clauses of reason is co-determined by principles of cognitive processing and discourse-pragmatics.

Further, motivators related to discourse-pragmatics (subordinator and bridging) are significantly more robust predictors of clause ordering than processing-related motivators (complexity and length). Moreover, the complexity of the dependent clause has a more significant contribution to the positioning of the reason clauses in comparison with other processing-related factors. Finally, random forest analysis proved to be a robust statistical means for predicting the relative weight of these constraints.

References

Aarts, Bass. 1988. Clauses of Concession in Written Present-day British English. *Journal of English Linguistics* 2. 39–85.

Arnold, Jennifer E., Losongco, Anthony, Thomas Wasow and Ryan Ginstrom. 2000. Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering. *Language 76*(1). 28–55.

Bever, Thomas G. 1970. The Cognitive Basis for Linguistic Structures. In: John R. Hayes (ed), *Cognition and the Development of Language,* 279–362. Hoboken: Wiley.

Biber, Douglas, Johansson, Stig, Leech, Gwoffrey, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English.* London: Longman.

Birner, Betty J. and Gregory Ward. 1998. *Information Status and Noncanonical Word Order in English*. Amsterdam: John Benjamins Publishing.

Breiman, Leo. 2001. Random Forests. *Machine Learning* 45 (1). 5–32.

Chafe, Wallace. 1984. How People Use Adverbial Clauses. *Berkeley Linguistics Society* 10. 437–49.

Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins Publishing.

Diessel, Holger. 2001. The Ordering Distribution of Main and adverbial Clauses: A typological Stud. *Language* 77 (3). 433–455.

Diessel, Holger. 2005. Competing Motivations for the Ordering of Main and Adverbial Clauses. *Linguistics* 43 (3). 449–470.

Diessel, Holger. 2008. Iconicity of Sequence: A Corpus-based Analysis of the Positioning of Temporal Adverbial Clauses in English. *Cognitive Linguistics* 19 (3). 465–490.

Ford, Cecilia E. 1993. *Grammar in Interaction*: *Adverbial Clauses in American English Conversations*. Cambridge: Cambridge University Press.

Gibson, Edward. 1998. Linguistic Complexity: Locality of syntactic Dependencies. *Cognition* 68 (1). 1–76.

Gibson, Edward. 2000. The Dependency Locality Theory: A Distance-based Theory of Linguistic Complexity. In: Alec Marantz, Yasushi Miyashita and Wayne O'Neil (eds.), *Image, Language, Brain,* 95–126. Cambridge, MA: MIT Press.

Givón, Talmy. 2001. *Syntax: An Introduction.* vol. 1. Amsterdam & Philadelphia*:* John Benjamins.

Givón, Talmy. 2011. *Ute Reference Grammar*. Amsterdam: John Benjamins Publishing.

Greenbaum, Sidney and Gerald Nelson. 1996. Positions of Adverbial Clauses in British English. *World Englishes* 15 (1). 69–81.

Hawkins, John. A. 1994. *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.

Hawkins, John A. 2004. *Efficiency and Complexity in Grammar*. Oxford: Oxford University Press.

Hothorn, Torsten, Hornik, Kurt and Achim Zeileis. 2006. Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics* 15 (3). 651–674.

Iwasaki, Noriko. 2010. Style Shifts Among Japanese Learners Before and After Study Abroad in Japan: Becoming Active Social Agents in Japanese. *Applied Linguistics* 31 (1). 45–71.

Kirk, John M. 1997. Subordinate Clauses in English. *Journal of English Linguistics* 25 (4). 349–364.

Li, Ming and Vitányi, Paul. 1997. *An Introduction to Kolmogorov Complexity and Its Applications*. Heidelberg: Springer.

Miestamo, Matti. 2006. On the feasibility of complexity metrics. In FinEst linguistics, proceedings of the annual Finnish and Estonian conference of linguistics, Tallinn, 11-26.

Mukherjee, Joybrato. 2001. Principles of pattern selection. *Journal of English linguistics* 29 (4). 295-314.

Quirk, Randolph. et al. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.

Rezaee, Abbas Ali and Seyyed Ehsan Golparvar. 2016. The Sequencing of Adverbial Clauses of Time in Academic English: Random Forest Modelling. *Journal of Language Modelling 4*(2), 225-244.

Rezaee, Abbas Ali and Seyyed Ehsan Golparvar. 2017. Conditional Inference Tree Modelling of Competing Motivators of the Positioning of Concessive Clauses: The Case of a Non-native Corpus. *Journal of Quantitative Linguistics 24*(2-3), 89-106.

Saif, Shahrzad. 2006. Aiming for Positive Washback: A Case Study of International Teaching Assistants. *Language Testing* 23 (1). 1-34.

Schoonen, Rob. et al. 2011. Modelling the Development of L1 and EFL Writing Proficiency of Secondary School Students. *Language learning* 61(1). 31-79.

Shizuka, Tetsuhito, Takeuchi, Osamu, Tomoko Yashima and Kiyomi Yoshizawa. 2006. A Comparison of Three-and Four-Option English Tests for university Entrance Selection Purposes in Japan. *Language Testing* 23 (1). 35-57.

Tagliamonte, Sali A. and Harald R. Baayen. 2012. Models, Forests, and Trees of York English: Was/were Variation as a Case Study for Statistical Practice. *Language Variation and Change* 24 (2). 135–178.

Thompson, Sandra A., Rober A. Longacre, and Shin Ja J. Hwang. 2007. Adverbial Clauses. In: Timothy Shopen (ed.), *Language Typology and Syntactic Description,* 237–300. Cambridge: Cambridge University Press.

Thompson, Sandra A. 1985. Grammar and Written Discourse. Initial and Final Purpose Clauses in English. In: Talmy Givón (ed.), *Quantified Studies in Discourse*. *Special Issue of Text,* 5, 55–84.

Trude, Alison M. and Natasha Tokowicz. 2011. Negative Transfer from Spanish and English to Portuguese Pronunciation: The Roles of Inhibition and Working Memory. *Language Learning* 61(1). 259-280.

Vandepitte, Sonia. 1993. *A Pragmatic Study of the Expression and the Interpretation of Causality: Aonjuncts and Conjunctions in Modern Spoken British English*. Brussel: Paleis der Academiën.

Vafaee, Payman, Basheer, Nesrine and Reese Heitner. 2012. Application of Confirmatory Factor Analysis in Construct Validity Investigation: The Case of the Grammar Sub-Test of the CEP Placement Exam. *Iranian Journal of Language Testing* 2 (1). 1-19.

Verstraete, Jean-Christophe. 2004. Initial and Final Position for Adverbial Clauses in English: The Constructional Basis of the Discursive and Syntactic Differences. *Linguistics* 42 (4). 819–853.

Vulanovic, Relja. 2007. On Measuring Language Complexity as Relative to the Conveyed Linguistic Information. *SKY Journal of Linguistics* 20. 399–427.

Wasow, Thompson. 2002. *Postverbal Behavior*. Stanford: CSLI Publications.

Wiechmann, Daniel and Kerz, Elma. 2013. The Positioning of Concessive Adverbial Clauses in English. *English Language and Linguistics* 17. 1–22.