ISSN 2007-9737

Rhetorical Relations in the Speech of Alzheimer's Patients and Healthy Elderly Subjects: An Approach from the RST

Anayeli Paulino[1], Gerardo Sierra[1], Laura Hernández-Domínguez[2], Iria da Cunha[3], Gemma Bel-Enguix[1]

[1] Universidad Nacional Autónoma de México, Instituto de Ingeniería, Mexico

[2] ETS, Quebec University, Montréal, Canada

[3] Universidad Nacional de Educación a Distancia (UNED), Spain

{apaudje, laudobla}@gmail.com, {GSierraM, GBelE@iingen.unam.mx}, iriad@flog.uned.es

Abstract. The study is aimed to extract discourse relations patterns in conversational speech of subjects with Alzheimer's Disease (AD) and adults with healthy aging processes using the Rhetorical Structure Theory (RST). By means of the RST, we analyzed semi-structured interviews of native Spanish speakers. Seven subjects were in the mild, moderate or advanced stages of AD, and 6 were cognitively intact individuals. The procedure involved the segmentation of each conversational discourse into Semantic Dialog Units (SDUs), the labeling of their rhetorical relations and the construction of tree diagrams. We perform a correlation analysis to determine the significance of the use of rhetorical relations for each group. We found a significantly (p-value $<$ .05) lower rhetorical relations production density in subjects with AD. We also observed that most rhetorical relations used by healthy older subjects were *Elaboration*, *Concession*, *Interpretation*, *Non-Volitional Cause*, *Solutionhood* and *Volitional Result*.

Keywords. Rhetorical relations, conversation analysis, Alzheimer's disease.

1 Introduction

In the catalogue of mental disorders, dementia is defined as a neurodegenerative syndrome that implies a gradual decrease in the ability to think, what finally leads to a degradation of the functional capabilities of the person. Alzheimer's disease (AD) is the prevalent dementia, and also the one most of research is focused on. In order to look for new indices for prognosis and treatment, the cognitive impairments of Alzheimer's have been studied.

In addition to memory, gnosias and visuospatial sketchpad, the linguistic function is one of the most degraded abilities. Since decades ago, some works like Appell [2], Bayles [4] and Nicholas et al. [28] described some areas and linguistic functions that are altered throughout the disease. For example, most patients show, in the first stage, semantic paraphasias, anomias [2, 4, 17, 5], problems keeping a conversation [2] and emotional prosodic impairments [40, 26].

In a moderate stage, some phonetic problems for articulating sounds start, like agrammatism, low levels of complex sentences and several discourse problems, for example, limited ability to make inferences [10], use of indefinite terms and discourse vagueness due to semantic slang, limitations to maintain topics of conversation, presence of perseverations [31, 27], prosodic incomprehension and intonation [40, 38], as well as a constant use of pauses or hesitations in discourse [30].

In recent years, the field of Natural Language Processing and linguistics have formed a very productive interdisciplinary area that has focused

on the early detection of this dementia. Thus, with the help of the new methods of computational linguistics, new data about performance in the language of patients have been found or refuted. The methodology and the linguistic phenomena used have been diverse, whether they are lexical features [39, 6, 3, 1, 22] or discourse and pragmatic phenomena [18].

However, there are still questions to be resolved. The first one lies in the fact that most of the corpus used is only for English-speaking patients and, therefore, the studies are entirely in English. The second question focuses on that, although studies involving PLN have obtained good results in the identification of dementia, most of them focus on the recount of textual markers to measure phenomena of different nature such as pragmatic discourse (for example, continuators or repetitions) and thus try to distinguish the speech of a healthy patient from that of another who suffers AD.

The automatic analysis of lexical clues in a representative linguistic corpus have yielded quantitatively interesting data about communicative idiosyncrasy of subjects with dementia; however, its results are often general. A linguistic approach would provide an insight about the progress of the disease and its effects if we take in account that, in language production, multiple parts and functions are involved.

Therefore, one of the main objectives of our work is to propose a new approach to analysis through the application of a discourse theory whose bases are found both in the generation of automatic language and in the discourse tradition, and which does not necessarily rest on lexical markers. We also use a corpus currently under construction, the Carolinas Conversations corpus in Spanish, which consists of semi-structured interviews conducted to Spanish-speaking adults with AD.

2 Related Work

2.1 Discourse Abilities in AD

The measurement of discourse coherence together with the analysis of lexical phenomena is one of the most studied approaches within the linguistic state-of-the-art analysis in AD. As of the 90s, we can see a growing number of studies that focused on sketching deficits in the spontaneous discourse of patients, sometimes with the help of short interviews or from referential tasks such as storytelling. Chapman [10], by means of a pictorial image, asked patients with Alzheimer's and healthy adults to retell the story they were shown to them. Through the structure of the clauses that shaped the discourse of each patient, the author discovered that those who suffered AD tend to produce fewer normal or complementary sentences and more interrupted structures compared to older adults in a healthy aging process.

Previously, Marjorie Nicholas et al. [28] undertook a comparative analysis between patients suffering from Wernicke's aphasia and individuals with Alzheimer's dementia. Based on the story count and the utterance of reflections on the part of the patients, the author intended to evaluate the abilities of communicative intention, inferences and distribution of information. The results showed a greater inability to make inferences in the population with AD, while grammatical failures were greater in aphasic patients.

Donald Ellis [14] approaches more closely to the cohesion markers that occur in the middle and late stages of the disease. He discovers that in the middle stage the subject starts having a pragmatic consideration more focused on his or her person than toward the interlocutors. In late stages, the subject began to present, what the author calls, a pre-grammatical ability about the use of the language.

Daniel Ripich [29] performs a longitudinal study to identify differences in cohesive patterns of the speech of men and women with AD. In his study, he evaluates the number and length of statements and cohesion phenomena as referents, ellipsis and conjunctions.

Among his findings, he argues that, in general, all cohesive aspects are used appropriately in the first stage while only the reference errors are visible in the early and middle stages.

Another methodological approach followed in the identification of dementias has considered the cohesive-discourse failures. Obler, Albert

and Helm-Estabrooks [28] analyze empty sentences, indefinite terms, deictics, paraphrases and conjunctions in descriptions of images made by populations with aphasias and with dementia. The most characteristic errors in the population with dementia were the indefinite terms and pronouns without antecedents.

The realization of reference tasks has been useful to extract skills related to a better performance based on the dynamics involved. Feyereisen [16] performed repetitions of description tasks in healthy people and patients with Alzheimer's disease. The experiments indicate that patients suffering from dementia did not maintain an apprenticeship or use of previous exercises to improve their discourse performance in the task. This caused a worse performance in their communicative efficiency. They also used more indefinite terms and less information units than healthy people.

2.2 RST and Conversation Analysis

Although RST already has an extensive range of applications in the area of NLP, such as automatic text generation [25], linguistic comparative study of languages [8, 12], information extraction [33], subjective content analysis [35, 20], textual similarity [13] and creation of Treebanks, this theory has not had the same importance in spoken language or conversational analysis. However, the interest to apply it to more diverse areas is increasing.

Some works have explored the application of RST in oral discourses, while others have devised new proposals in order to cover the phenomena that appear in discourse. Fawcett and Davies [15] applied the RST to monologues of speakers who held a conversation. Based on the concept of Elementary Discourse Units and the RST rhetorical relations catalog, Amanda Stent [34] developed a labeling manual for conversations, which also included additional features such as prosodic phenomena and turn-based separation.

Taboada [37] analyzes a bilingual corpus of telephone conversations in English and Spanish using the standard RST theory. In her proposal, she identifies conversational phenomena and includes them within the analysis, which is established through two levels. With this, the author verifies that the RST is applicable to spoken discourse and that it is valid for languages other than English.

Desiderato [23] extracts features of oral discourse in Brazilian readings. After an analysis based on the RST, the author concludes that the theory can include a variety of speech phenomena, such as repetitions and paraphrasing. From this, Taboada agrees that the RST is suitable for the analysis of conversational discourses.

The presence of new conversational forms due to the massive use of social networks has also been covered by the RST. Sidarenka, Bispin and Stede [32] analyze issues to be considered in a corpus of dialogues on Twitter and highlight phenomena such as the non-adjacent particular order of some dialogues on this social networking service, the use of the tree diagram and the adequacy of the taxonomy relationship.

Other projects have also sought to add more phenomena to speech in order to create an automatic labeling tool such as the DAMSL project [11], which includes the consideration of speech acts, or the ISO project [7], that seeks to create an optimal and self-sufficient labeling system for any conversation.

2.3 RST and Clinical Linguistics

So far, the RST has yet to be used in the field of linguistic disorders, but it is a potential option that is beginning to be discovered. One of the most complete studies carried out in this field is the work presented by Kong et al. [24] who analyze the coherence of patients with aphasia and healthy ones. From two different genres, narration and description, Kong performs the segmentation following oral, semantic and phonetic characteristics.

Then, with the catalog of rhetorical labels, he performs the construction of tree diagrams. Finally, through various criteria such as communicative effectiveness, formulation of complete ideas, connectivity between units, and complexity and interrupted fluency, the author concludes that healthy patients performed better in all criteria and

recorded greater use of relationships background, explanation and elaboration.

3 Rhetorical Structure Theory

RST has been defined as a functional theory of text structure whose basis consist on the functional relations description between parts of a text. Mann and Thompson argue that its function is centering in the fact that "It describes the relations between text parts in functional terms, identifying both the transition point of a relation and the extent of the items related. It provides comprehensive analyses rather than selective commentary" [36, p. 243]. RST respond to a hierarchy principle and organization that is established between parts of a textual or oral discourse. The union of these parts maintain a semantic or cohesive association.

3.1 Fundamental Units

The fundamental structures or units are called Elementary Discourse Units (EDUs) which, in general terms, correspond to a sentence. The EDUs are organized by pairs or pairs of relationships that are called *spans*. This association between two units is fundamental and is shaped by the following principle: in a discourse, there are more important units that would be the backbone of the text, without them the logical sense of any discourse would be lost. These units will be called nuclear units.

The nuclear units are accompanied by complementary units, which add more information; these, in turn, have to join other higher structures to form a tree diagram. The units that complement the nuclear unit and are subjugated to it are called satellite relations.

Spans can be integrated by a nuclear relation and a satellite relation, or two nuclear relations.

In this last case, a *span* can consist of more than one relation.

3.2 Rhetorical Relations

Rhetorical relations give an account of the semantic concept of connections between units. Such relationships are establishing the discourse coherence throughout the text. In the state of art, these relations have also been called relations of coherence, discourse relations and relations of conjunction. In an analysis, these relationships are usually indicated through labels and are organized according to their intention, the semantic concept they contribute and the class of units they point to. Currently, there are several manuals that offer [9, 34] an extensive catalog of these connections.

The representation of the *spans* is usually done through schemes that indicate the order of each element. The result of the analysis is expected to be represented by a tree diagram under which the most general and important relationships remain at the most superficial level and the information descends as it becomes more detailed.

4 Method

This paper follows the proposal of Taboada [37] due to its applicability to Spanish and its methodological adequacy to oral analysis. We include certain adjustments due to the phenomena that we find throughout the analysis.

4.1 Sample and Transcription

For this analysis, we decided to use the Spanish corpus Carolinas Conversations [21] that has been compiled since 2015 as a joint project between the Language Engineering Group (LEG) of the UNAM, the École de Technologie Supérieure, of Montréal, Canada, the Medical University of South Carolina (MUSC) and the University of North Carolina at Charlotte (UNCC). The corpus consists of semi-structured interviews with healthy elderly patients, with dementia or with some other neurodegenerative disease, such as Parkinson's disease or other mental disorders such as Bipolar disorder.

The interviews are collected periodically, and on each occasion they are appended to new patients.

ISSN 2007-9737

The transcription and alignment of the interviews were done with the help of the Transcriber 1.5.1 program[1]. The transcription conventions used were given by the compilers of the CCC corpus in Spanish. These conventions allow to mark punctuation marks, prosodic phenomena (changes in tone, intonations), kinesthetic elements (facial gestures, gestures) and contextual elements (external interruptions, noises, etc.). When the interview was completed it was imported into plain text format for better handling.

4.2 Segmentation Process

The initial parameter that Taboada proposes for speech analysis is the nature of the conversation: If the conversational discourse is considered to be a whole in which each exchange is a structural part, then the analysis should be done through the whole text. If, on the other hand, each turn is conceived as an autonomous and coherent structure in itself, then the analysis will be intraturn. For this work, we will adopt the last premise. In this way, our initial unit was the turn of each patient and each caregiver.

In standard theory, an Elementary Discourse Unit (EDU) was the minimum unit within the discourse. Broadly speaking, the EDU were clauses that had a nominal nucleus and a verbal nucleus and expressed a complete sense.

In our work, instead, we prefer the Semantic Dialogue Unit (SDU), introduced by Taboada, which is adapted to the features of a conversation, and is based on discourse theories [19]. Apart from considering the usual criteria of an EDU (syntactic criteria), an SDU can be delimited by prosodic (intonation, pauses) and semantic phenomena (semantic completeness) [24].

However, in a population that presents disorders in pragmatic characteristics and discourse elements, the ability to formulate structural and semantically coherent SDUs decreases considerably. Instead, we note that his speech is often full of incomplete utterances, circumlocutions and unusually long pauses.

For this reason, we allow ourselves to make some concessions: first, we wanted the SDUs to be kept as complete as possible, unless there was a discourse marker that obviated the presence of a rhetorical relation:

> *053 Mr. Buendía: Yo soy de la provincia del Carchi Cantón Mira, <ya> — [entonación ascendente] donde se acuesta uno — y se levantan tres [laughs].*

> *053 Mr. Buendía: I am from Carchi Cantón Mira's province, <ok> —[ascending intonation] where one goes to sleep — and wakes up with other three [laughs]*

If the syntactic string was abnormal or incomplete, it was allowed to consider it as an SDU whenever the semantic idea could be understood:

> *064 Mrs. San Juan: Sí. [background conversations] — Me dijeron ahora [pron=ora] que sabía que Mrs. Regina, este, que - que eso es lo que me mandaron decir,— pero que- que me ponga así la ropa, que el de saco [background conversations] [step noise].*

> *064 Mrs. San Juan: Yes. [background conversations] — They told me that now that I knew that Mrs. Regina, amm, that - that is what they were going to tell me, — but that, that I must wear the clothes just like that, and the coat [background conversations] [step noise].*

Likewise, some pre-done sentences where considered to be SDUs if they provided and important meaning to the turn built by the patient:

> *054 Mrs. Cortés: - - - <mmm> Un, una cosa buena se presenta entonces, cambia de- - - — deja de estar tristona y... — Eso es.*

> *054 Mrs. Cortés -<mmm> An, a good thing happens then, changes the - - - — the one is not sad anymore... — That's it.*

In case a segmentation limit was not clear, then prosodic features such as tone or intonation were applied to determine the start or end of the SDUs.

[1] http://trans.sourceforge.net/en/presentation.php

ISSN 2007-9737

4.3 Relations Labeling and Discourse-Trees Elaboration

Once all the patients' turns were segmented, their organization and labeling of rhetorical relations were carried out. This operation consisted in associating pairs or *spans* of SDUs and considering their hierarchical relevance in the turn as well as the type of union that existed between them.

The standard catalog of the RST proposed by Mann and Thompson was chosen. This catalog consists of 32 rhetorical relationships defined and exemplified in the official website of the RST[2]. The RSTTool program allows the integration of an indefinite relationship catalog, which is deployed when establishing the labeling between two SDUs.

It is worth mentioning that not all turns can form relationships. The linguistic features and limitations that constrained the speakers according to the progress of the disease were evident when formulating turns of a single statement and even sentences or monosyllables. In the first approach where the text is conceived as a whole, the shifts of a single SDU can be complemented by the preceding or subsequent interventions of the interlocutor.

After the linguistic discourse analysis of the RST, the statistical analysis was carried out, which involved the counting and type of relationships of each individual in both populations. We call production density the frequency of relations produced by each subject of the sample. The number of relations was provided by the RSTTool program. Although the caregivers' shifts were also analyzed, we only took the turns of patients and healthy subjects into account.

Certainly, a greater production of one or another relation in the general count does not guarantee that it is precisely an exclusive relationship of one or another population. A correlation analysis was carried out to determine which rhetorical relations were connected to the disease and which corresponded to cognitively healthy patients. To corroborate this data, the *p- value* was calculated and data that had an adequate index of co-dependability was recorded.

[2] http://www.sfu.ca/rst/index.html

5 Results

5.1 Rhetorical Relations Production Density

Figure 2 shows the production density of the relations. Of all the rhetorical relationships, *Elaboration* had a higher production density in both populations, with 163 units in the cognitively healthy population, and 67 for the population with dementia. From there, the relationships with higher production density are different and their density is considerably lower than the ratio already mentioned.

Among them, the best ranked in the healthy population were *Restatement*, *Evidence* and *Concession* (40, 37 and 35 relationships, respectively). On the contrary, *Justify* and *Restatement*, with 17 and 15 repetitions were the most productive for the population with dementia. Relations with a quite low production density were *Otherwise*, *No Conditional* and *Unless*.

Relations that did not appear at all were *Enablement* and *Multinuclear Restatement*.

5.1.1 Correlation Analysis

Although it is true that the counting of the frequency of relationships provides very prominent results in the production density and type of relations used and unused by the sample, the truth is that this is not a reliable indicator of the cognitive state of a subject. Therefore, it was necessary to conduct a correlation analysis between the use of each of the relationships and their association with the disease.

Table 1 presents the most significant results of the correlation analysis. The first column includes each of the already normalized relationships and the total of each type. In the second column the values of correlation between the healthy condition or with dementia are registered regarding to the production of rhetorical relations and their total production. In addition, the *p-value* that indicated an acceptable confidence interval ($p\text{-value} > .05$) in the obtained data was included.

We then corroborate two aspects of the previous graph. The first is that cognitively healthy patients tend to produce a greater number of connections.

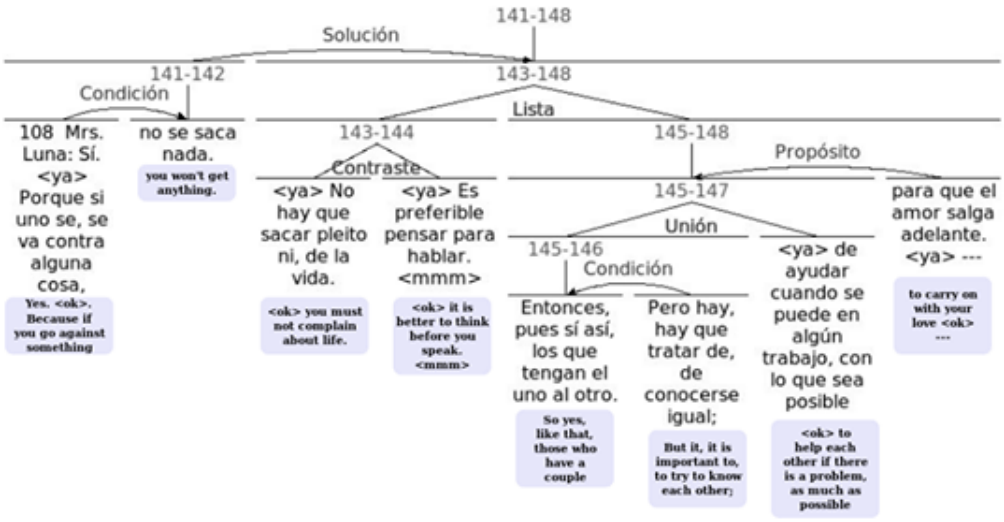

Fig. 1. A complete discourse tree. Solutionhood is the most general relation in the turn and Condition, the most specific

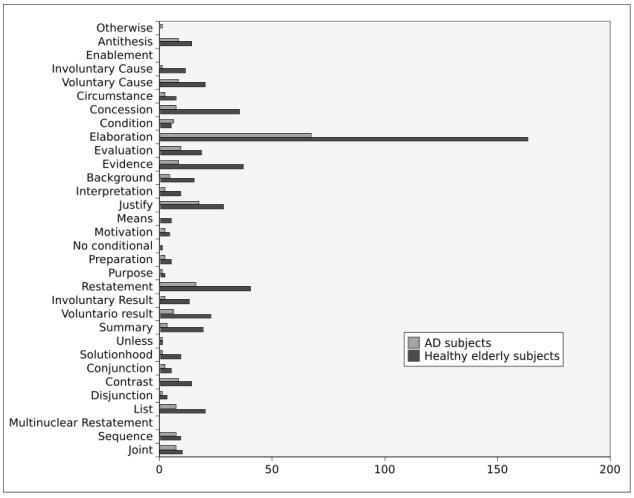

Fig. 2. Rhetorical Relations production density by AD patients and healthy elderly subjects

Second, we verified that the *Elaboration* and *Concession* structures are more frecuently produced by healthy population.

On the other hand, the correlation analysis shed light on other relations that had not been considered: according to the data, *Involuntary Cause* and *Solutionhood* have a close correlation with a subject in the process of normal aging and are significant regardless of the total number rhetorical relations that were recorded (both relations appeared two times in the correlation analysis).

Other relations that turned out to be highly significant were *Involuntary Result*, *Interpretation*, *List*, *Circumstance*, *Summary*, *Background* and *Elaboration*.

6 Conclusion and Future Work

The estimation of the production density and the correlation analysis yielded interesting data about the use of rhetorical relations. The registers suggest that the use of rhetorical relations as a possible measurement index for coherence and discourse cohesion turns out to be highly productive, due to the quantitative difference that appears in both populations. This parameter would even serve to detect other linguistic indices quite used in the methodology of the works of psycholinguistics or clinical linguistics since we start from the identification of discourse structures that integrate syntactic and semantic characteristics.

Regarding the use of relations, there are several things to be noted. It is not strange that in both samples the *Elaboration* ratio is the most used

Table 1. Correlation analysis data and *p-value*

Feature	*p-correlation*	*p-value*
Concession	0.79820	0.0010
Total_nuclear	0.71298	0.0062
Total	0.70301	0.0073
Involuntary Cause	0.69850	0.0079
Solution	0.69532	0.0083
Involuntary Result	0.64246	0.0178
Involuntary Cause	0.63921	0.0186
Interpretation	0.63829	0.0188
List	0.61226	0.0261
Circumstance	0.60860	0.0272
Solution	0.60598	0.0281
Summary	0.60230	0.0293
Background	0.60090	0.0298
Voluntary Result	0.58955	0.0339
Elaboration	0.57447	0.0400

since it allows to add complementary blocks of information.

Likewise, the correlation of certain rhetorical relationships over others has a discourse explanation. Basically, all the relations that were found to be relevant for the healthy population involve pragmatic or discourse skills that a subject with a linguistic disorder might find difficult to produce. For example, *Concession* apparently involves the contraposition of two ideas that, in reality, are the consideration of two different arguments. Such consideration of arguments can hardly be expressed by a patient with dementia.

In general, both the *Involuntary/ Voluntary Cause* and *Result* imply that the speaker considers his inclusion or his decision in a triggering event or in an action that gives rise to another. It is known that patients with dementia are personalize their speech as the disease progresses and often lose consideration of the other during communication. We speculate that this could be the reason why their production density is minimal in the population with AD.

In the analysis, we noticed that the relation of *Solutionhood* in a conversation frequently required the formulation of other relationships to complement it: *Evidence*, *Justify*, *Elaboration*, etc. Because of this recurrence of units and relations, subjects with dementia would show a lower production density.

Finally, the relations of *Interpretation* and *Summary* share a characteristic: their use implies a recapitulation of previous arguments. While *Summary* requires the ability to synthesize, *Interpretation* requires a new and critical evaluation regarding what has previously been said, so that a person with Alzheimer's would have several disadvantages to produce this kind of relationship.

The negative correlations for the cognitively impaired population did not obtain the significance indices necessary to be reliable data. A quantitatively higher population would have contributed to validate such data.

6.1 Future Work

The purpose of this work was to apply the methodology of the RST to spoken language in search of patterns that allow us to identify particular characteristics within the discourse of cognitively healthy older adults and the elderly who suffer from dementia of the Alzheimer type. Despite the small size of our sample, we consider that the methodology of the RST was opportunely applied.

Likewise, the initial data of simple statistical processes yielded interesting results that are relevant from the perspective of the RST. However, a lot of tasks should be performed in the future. A first one would be to apply the same process to a quantitatively superior corpus to obtain more objective data in the profiling of people. Another valuable experiment would consist in repeating Kong's methodology to infer whether the relations shown by our analysis are typical of the cognitive state of the patients or whether the gender in question has an impact.

As a newly discovered method in the clinical field, the RST opens the door to more rigorous application proposals. For example, its contribution in identifying the degree of progress of dementia

in which a patient is. Methodologically, there are several applications to be made. For example, analyzing the spoken discourse structure through the number of nodes and the levels of each scheme will help us in this. Additionally, there is a lack of analysis of the conversation as a whole, which would be the second part of this work.

Acknowledgements

This research was supported by CONACYT's projects: *Fronteras de la Ciencia*, and *Ampliación del corpus multiétnico de conversaciones con personas de edad avanzada*. It was also supported by the FRQNT 177 601, the Ministère des Relations Internationales et de la Francophonie; the Ramón y Cajal contract (RYC-2014-16935) associated with the Department of Foreign Philologies and their Linguistics (UNED) and the PAPIITT's project IA400117.

References

1. **Alegria, R., Gallo, C., Bolso, M., dos Santos, B., Prisco, C. R., Bottino, C., & Ines, N. M. (2013).** Comparative study of the uses of grammatical categories: Adjectives, adverbs, pronouns, interjections, conjunctions and prepositions in patients with Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, Vol. 9, No. 4, pp. P882.
2. **Appell, J., Kertesz, A., & Fisman, M. (1982).** A study of language functioning in Alzheimer patients. *Brain and language*, Vol. 17, No. 1, pp. 73–91.
3. **Asgari, M., Kaye, J., & Dodge, H. (2017).** Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, Vol. 3, No. 2, pp. 219–228.
4. **Bayles, K. A. & Tomoeda, C. K. (1983).** Confrontation naming impairment in dementia. *Brain and language*, Vol. 19, No. 1, pp. 98–114.
5. **Blair, M., Marczinski, C. A., Davis-Faroque, N., & Kertesz, A. (2007).** A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. *Journal of the International Neuropsychological Society*, Vol. 13, No. 2, pp. 237–245.
6. **Bucks, R. S., Singh, S., Cuerden, J. M., & Wilcock, G. K. (2000).** Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology*, Vol. 14, No. 1, pp. 71–91.
7. **Bunt, H., Alexandersson, J., Choe, J.-W., Fang, A. C., Hasida, K., Petukhova, V., Popescu-Belis, A., & Traum, D. R. (2012).** ISO 24617-2: A semantically-based standard for dialogue annotation. *LREC*, pp. 430–437.
8. **Cao, S., da Cunha, I., & Bel, N. (2016).** An analysis of the Concession relation based on the discourse marker aunque in a Spanish-Chinese parallel corpus. *Procesamiento del Lenguaje Natural*, Vol. 56, pp. 81–88.
9. **Carlson, L. & Marcu, D. (2001).** Discourse tagging reference manual. *ISI Technical Report ISI-TR-545*, Vol. 54, pp. 56.
10. **Chapman, S. B., Highley, A. P., & Thompson, J. L. (1998).** Discourse in fluent aphasia and Alzheimer's disease: Linguistic and pragmatic considerations. *Journal of Neurolinguistics*, Vol. 11, No. 1, pp. 55–78.
11. **Core, M. (1998).** Analyzing and predicting patterns of DAMSL utterance tags. *Proceedings of the AAAI spring symposium on Applying machine learning to discourse processing*.
12. **da Cunha, I. & Iruskieta, M. (2010).** Comparing rhetorical structures in different languages: The influence of translation strategies. *Discourse Studies*, Vol. 12, No. 5, pp. 563–598.
13. **da Cunha, I., Vivaldi, J., Torres-Moreno, J.-M., & Sierra, G. (2014).** SIMTEX: An approach for detecting and measuring textual similarity based on discourse and semantics. *Computación y Sistemas*, Vol. 18, No. 3, pp. 505–516.
14. **Ellis, D. G. (1996).** Coherence patterns in Alzheimer's discourse. *Communication Research*, Vol. 23, No. 4, pp. 472–495.
15. **Fawcett, R. P. & Davies, B. L. (1992).** Monologue as a turn in dialogue: Towards an integration of exchange structure and rhetorical structure theory. In *Aspects of automated natural language generation*. Springer, pp. 151–166.
16. **Feyereisen, P., Berrewaerts, J., & Hupet, M. (2007).** Pragmatic skills in the early stages of Alzheimer's disease: an analysis by means of a referential communication task. *International journal*

of language & communication disorders, Vol. 42, No. 1, pp. 1–17.

17. **Gil i Saladié, D. (1988).** El lenguaje en la enfermedad de Alzheimer. *Revista de Logopedia, Foniatría y Audiología*, Vol. 8, No. 4, pp. 199–205.
18. **Guinn, C. I. & Habash, A. (2012).** Language analysis of speakers with dementia of the Alzheimer's type. *AAAI Fall Symposium: Artificial Intelligence for Gerontechnology*, pp. 8–13.
19. **Halliday, M. A. (1967).** Notes on transitivity and theme in english: Part 2. *Journal of linguistics*, Vol. 3, No. 2, pp. 199–244.
20. **Heerschop, B., Goossen, F., Hogenboom, A., Frasincar, F., Kaymak, U., & de Jong, F. (2011).** Polarity analysis of texts using discourse structure. *Proceedings of the 20th ACM international conference on Information and knowledge management*, ACM, pp. 1061–1070.
21. **Hernández, L., Ratté, S., Pope, C., & Davis, B. (2016).** Conversing with the elderly in latin america: a new cohort for multimodal, multilingual longitudinal studies on aging. *ACL 2016*, pp. 16.
22. **Hernández-Domínguez, L., García-Cano, E., Ratté, S., & Martínez, G. S. (2016).** Detection of Alzheimer's disease based on automatic analysis of common objects descriptions. *Proceedings of the 7th Workshop on Cognitive Aspects of Computational Language Learning*, pp. 10–15.
23. **Juliano, D. A. & Cassim, F. T. R. (2012).** Coherence relations in academic spoken discourse. *Linguistica*, Vol. 52, No. 1, pp. 323.
24. **Kong, A. P. H., Linnik, A., Law, S., & Shum, W. (2014).** Measuring the coherence of healthy and aphasic discourse production in Chinese using Rhetorical Structure Theory (RST). *Annual Meeting of the Academy of Aphasia*, Frontiers in Psychology.
25. **Marcu, D. (1997).** The rhetorical parsing of natural language texts. *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, Association for Computational Linguistics, pp. 96–103.
26. **Meilán, J. J. G., Martínez-Sánchez, F., Carro, J., López, D. E., Millian-Morell, L., & Arana, J. M. (2014).** Speech in Alzheimer's disease: Can temporal and acoustic parameters discriminate dementia? *Dementia and Geriatric Cognitive Disorders*, Vol. 37, No. 5-6, pp. 327–334.
27. **Müller, N. & Guendouzi, J. A. (2005).** Order and disorder in conversation: Encounters with dementia of the Alzheimer's type. *Clinical linguistics & phonetics*, Vol. 19, No. 5, pp. 393–404.
28. **Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985).** Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech and Hearing Research*, Vol. 28, No. 3, pp. 405–410.
29. **Ripich, N., Carpenter, B. D., & Ziol, E. W. (2000).** Conversational cohesion patterns in men and women with Alzheimer's disease: a longitudinal study. *International Journal of Language & Communication Disorders*, Vol. 35, No. 1, pp. 49–64.
30. **Rodríguez, J., Martínez, H., & Valles, B. (2015).** Las pausas en el discurso de individuos con demencia tipo Alzheimer. estudio de casos. *Revista de Investigación en Logopedia*, Vol. 5, No. 1.
31. **Sajjadi, S. A., Patterson, K., Tomek, M., & Nestor, P. J. (2012).** Abnormalities of connected speech in semantic dementia vs Alzheimer's disease. *Aphasiology*, Vol. 26, No. 6, pp. 847–866.
32. **Sidarenka, U., Bisping, M., & Stede, M. (2015).** Applying Rhetorical Structure Theory to twitter conversations.
33. **Soderland, S. & Lehnert, W. (1994).** Corpus-driven knowledge acquisition for discourse analysis. *A A*, Vol. 27, pp. 14neg.
34. **Stent, A. & Allen, J. (2000).** Annotating argumentation acts in spoken dialog. *University of Rochester, Rochester, NY*.
35. **Taboada, M. & Grieve, J. (2004).** Analyzing appraisal automatically. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Re# port SS# 04# 07), Stanford University, CA, pp. 158q161. AAAI Press*.
36. **Taboada, M. & Mann, W. C. (2006).** Rhetorical structure theory: Looking back and moving ahead. *Discourse studies*, Vol. 8, No. 3, pp. 423–459.
37. **Taboada, M. T. (2004).** *Building coherence and cohesion: Task-oriented dialogue in English and Spanish*, volume 129. John Benjamins Publishing.
38. **Taler, V., Baum, S. R., Chertkow, H., & Saumier, D. (2008).** Comprehension of grammatical and emotional prosody is impaired in Alzheimer's disease. *Neuropsychology*, Vol. 22, No. 2, pp. 188.
39. **Thomas, C., Keselj, V., Cercone, N., Rockwood, K., & Asp, E. (2005).** Automatic detection and

rating of dementia of Alzheimer type through lexical analysis of spontaneous speech. *Mechatronics and Automation, 2005 IEEE International Conference*, volume 3, IEEE, pp. 1569–1574.

40. **Tosto, G., Gasparini, M., Lenzi, G., & Bruno, G. (2011).** Prosodic impairment in Alzheimer's disease: assessment and clinical relevance. *The Journal of neuropsychiatry and clinical neurosciences*, Vol. 23, No. 2, pp. E21–E23.

Article received on 19/01/2018; accepted on 05/03/2018. Corresponding author is Anayeli Paulino.