UNIVERSITY OF MURCIA

International Journal of English Studies

IJES

http://revistas.um.es/ijes

A Corpus Analysis of the *Aktionsarten* of English-Speaking Patients with Alzheimer's Disease: A Role and Reference Grammar Account

ALEJANDRO SUÁREZ RODRÍGUEZ*
Research Institute of Text Analysis and Applications (IATEXT)
University of Las Palmas de Gran Canaria (Spain)

Received: 14/06/2023. Accepted: 04/12/2023.

ABSTRACT
In this paper, we aim to describe the frequency and relative distribution of verb use by English-speaking patients with Alzheimer's disease along its three stages. To do so, we apply the semantic representation of Role and Reference Grammar by means of the lexical aspect or *Aktionsart* to samples of verbs taken from the Pitt corpus of American patients with Alzheimer's disease. Then we apply descriptive statistical measures and hypothesis testing to the samples. Our results show that patients systematically use states as the preferred type of verbs in the three stages when compared to the rest of *Aktionsarten*. We also show that there exists a statistically significant relation between the lexical aspect of verbs and the stages of Alzheimer's disease. Among the explanations for these results, we propose that states may be used as the default Aktionsart because of its easier cognitive processing.

KEYWORDS: Corpus linguistics; *Aktionsarten*; Alzheimer's disease; Pitt corpus; Role and Reference Grammar.

1. INTRODUCTION

Alzheimer's disease is the most common type of dementia in the world (World Health Organization, 2023) in people who are older than 60 years and it is especially found in women (Beam et al., 2018; Martinkova et al., 2020). It affects brain areas dedicated to calculating, language, memory or judgement, although consciousness is mostly unimpaired (Atri, 2019: 266; WHO, 2023). There exist three phases of the disease: early stage, intermediate stage and

Address for correspondence: Calle Pérez del Toro, 1, Facultad de Fiolología, Universidad de Las Palmas de Gran Canaria, 35003, Las Palmas, España; e-mail: alejandro.suarezrodriguez@ulpgc.es

Print ISSN: 1578-7044; Online ISSN: 1989-6131
doi: 10.6018/ijes.573611

late stage, where symptoms appear progressively even before diagnosis. These symptoms include a gradual memory loss up to complete dependence on another person to fulfill basic needs (WHO, 2023).

In the same vein, we can summarize the aims of clinical linguistics as to analyze and describe the language deficits in a patient in the clinical sphere (Crystal, 1981; Garayzábal Heinze 2009: 143; Perkins & Howard, 2011). This broad application of linguistics has resulted in a greater focus on aphasia and phonetic aspects (Perkins, 2011: 923-925). Hence, there are studies that ignore syntax and semantics in the analysis of language pathology (Gallardo Paúls & Valles González, 2008: 38). Since clinical linguistics attempts to improve people's life by means of linguistic tools, we agree with those who consider it an essential part of language sciences, following Gallardo Paúls and Valles González (2008: 47). Moreover, there are studies that show how Alzheimer's disease is present in the lexical semantic content of language (Ivanova, García Meilán, Martínez Nicolás & Llorente, 2020; Pérez Cabello de Alba, 2017; Pérez Cabello de Alba & Teomiro García, 2018).

For this reason, in this article we use the tools of Role and Reference Grammar (RRG; Van Valin and LaPolla 1997; Van Valin 2005), a functionalist theory that studies syntax, semantics and the interaction of the pragmatic content. These tools will be applied to a corpus of English-speaking patients with Alzheimer's disease, namely the Pitt Corpus (Becker et al. 1994), in order to describe the frequency and relative distribution of verbs uttered by these patients in the three stages. Another aim is to verify whether lexical aspect may be regarded as a linguistic marker of Alzheimer's. Until now, there are no studies that show lexical aspect and its distribution in an English-speaking corpus as a way to better understand this disease, although there exists some research related to *Aktionsart* with no mention of RRG (Jensen, 2007; Shain, 2011). Furthermore, this linguistic theory has been used as a descriptive tool for the same corpus we intend to analyze in this paper (Suárez-Rodríguez, in press).

The remainder of the paper is organized as follows. In Section 2, we present the theoretical background, in which we briefly describe RRG's semantic representation. In Section 3, we describe the corpus to be analyzed and we show the method that we have followed to study the predicates of English-speaking patients with Alzheimer's disease from an RRG point of view. In section 4, we show the results of the analysis and we discuss them. Lastly, we provide conclusions and future work regarding this research.

2. THEORETICAL FRAMEWORK

Functionalism is one of the main branches of linguistic analysis and it sees language as "a system of communicative social action" (Van Valin, 2005: 1). Role and Reference Grammar (RRG; Van Valin, 2005; Van Valin & LaPolla, 1997) is part of these functionalist approaches, where it studies not only the morpho-syntactic structure of sentences, but also its semantic

Print ISSN: 1578-7044; Online ISSN: 1989-6131

content, its discourse-pragmatics and, to a certain extent, the cognitive aspects (Van Valin & LaPolla, 1997: 15).

RRG attempts to represent language production and comprehension of utterances, and it achieves this by means of a bi-directional linking algorithm, i.e., a series of rules that permit to go from the syntactic structure to semantic content (comprehension) or vice versa (production), where discourse-pragmatics play an important role (Van Valin, 2005: 1). This can be schematized as in Figure 1. Even though RRG attempts to capture these three elements in a sole representation, we will only focus on the semantic content in this paper, given that it is a universal part of the theory, i.e., it can be applied to any language of the world.

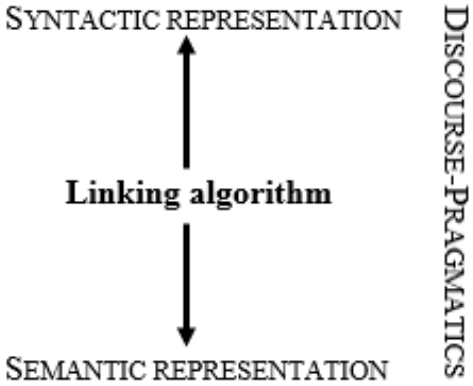

Figure 1. General structure of Role and Reference Grammar (taken from Van Valin, 2005: 2).

In the following subsection, we introduce the semantic representation according to RRG, where we focus only on the classification of predicates. A very brief account of the rest of the semantic representation is also offered at the end of the subsection.

2.1. The semantic representation in RRG

It has been traditionally assumed that predicates can only be verbs, but RRG considers nominal, adjectival and prepositional predicates too (Van Valin, 2005: 28). These lexical units capture the meaning of happenings in the (real or fictional) world and the entities that take part in that space-time, also known as *states of affairs* (Van Valin & La Polla, 1997: 82)

When lexical representation is taken into account, it implies a theory of verb classes (Van Valin 2005: 31), and thus RRG is based on Vendler's (1957) classification of *Aktionsarten*, a German word which means "form of action" and which refers to "the inherent temporal properties of verbs" (Van Valin and LaPolla 1997: 91-92); we will use both "*Aktionsart*" and "lexical aspect" interchangeably throughout this paper. As such, Vendler's classification distinguishes between states, activities, accomplishments and achievements, whereas RRG extends it by modifying the representational schema proposed by Dowty in 1979 (Van Valin 2005: 31). Moreover, RRG adds active accomplishments (Van Valin 2005: 32)

Print ISSN: 1578-7044; Online ISSN: 1989-6131

and Comrie's (1976) semelfactives to its verb classes, then used in Smith (1997). In doing so, we have at least six types of verb classes and their causative counterparts, which leaves twelve minimum possibilities of classifying verbs. These verb classes can be coded in Boolean values by means of semantic features, so that each verb class is defined according to the absence or presence of four semantic features: [±static], [±dynamic], [±telic] and [±punctual] (Van Valin, 2005: 33). We explain these features considering Van Valin (2005: 32-42):

1. The [±static] feature refers to the difference from those verbs that express a happening in a state of affairs and those that code a 'non-happening.'
2. In RRG, the [±dynamic] feature has a distinct usage, since it refers exclusively to "whether a situation involves action or not", i.e., participants do something or causes something to happen. That means that not every non-static verb is dynamic and in the case of semelfactives, some are derived from activities and thus are dynamic.
3. Telicity, i.e. the [±telic] feature, has to do with how a verb depicts a state of affairs that has an inherent end point or not. Achievements, accomplishments and active accomplishments have an inherent terminal point and so they are telic.
4. The [±punctual] feature deals with the internal duration of verbs and it helps to distinguish those verbs that are instantaneous from those that are not. Only achievements and semelfactives are instantaneous, and only achievements have a result state.

Verb classes represent different states of affairs and they are defined by the combination of the previous semantic features (adapted from Van Valin, 2005: 33):

State	[+static] [–dynamic] [–telic] [–punctual]
Activity	[–static] [+dynamic] [–telic] [–punctual]
Achievement	[–static] [–dynamic] [+telic] [+punctual]
Semelfactive	[–static] [±dynamic] [–telic] [+punctual]
Accomplishment	[–static] [–dynamic] [+telic] [–punctual]
Active accomplishment	[–static] [+dynamic] [+telic] [–punctual]

Each verb class is mapped to a situation that has been induced by another state of affairs, that is, the causative counterparts. Consider the following examples (1)-(12), taken from Van Valin (2005: 34), to have a better understanding of the lexical aspect and their causative versions:

(1)	State:	The boy is afraid.
(2)	Causative state:	The dog frightens the boy.
(3)	Activity:	The soldiers marched in the park.
(4)	Causative activity:	The sergeant marched the soldiers in the park.
(5)	Achievement:	The balloon popped.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

(6) Causative achievement: The cat popped the balloon.
(7) Semelfactive: The pencil tapped on the table.
(8) Causative semelfactive: The teacher tapped the pencil on the table.
(9) Accomplishment: The ice melted.
(10) Causative accomplishment: The hot water melted the ice.
(11) Active accomplishment: The soldiers marched to the park.
(12) Causative active accomp.: The sergeant marched the soldiers to the park.

In order to obtain the *Aktionsarten*, RRG provides eight tests that are not infallible, but when applied sequentially, they allow us to distinguish the verb classes by using the previous semantic features, as well as their causative counterparts, in order to avoid inconsistencies and to make the determination more effective. We will show the tests from Cortés Rodríguez, González Vergara and Jiménez Briones (2012: 62-65), since they slightly modify the tests from Van Valin (2005: 35-41). All these tests can be summarized as in Table 1, where the asterisks refer to exceptions for that *Aktionsart* and test; we explain the tests afterwards.

Table 1. Tests to determine the verb classes (adapted from Van Valin 2005: 39, and Cortés Rodriguez, González Vergara and Jiménez Briones 2012: 65).

Verb class	**Test 1: What happened?**	**Test 2: Progressive aspect**	**Test 3: Dynamic**	**Test 4: Duration**	**Test 5: *for*-PP**	**Test 6: *in*-PP**	**Test 7: Stative modifier**	**Test 8: Causative**
State	No	No*	No	No	Yes*	No	Yes	No
Activity	Yes	Yes	Yes	Yes	Yes	No	No	No
Achievement	Yes	No*	No	No*	No*	No*	Yes	No
Accomplishment	Yes	Yes	No	Yes	Irrel.*	Yes	Yes	No
Semelfactive	Yes	No*	No*	No*	Yes*	No*	No	No
Active accomp.	Yes	Yes	Yes	Yes	Irrel.*	Yes	Yes	No
Causative state	No*	Yes*	Yes*	No	Yes	No	Yes	Yes
C. activity	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
C. achievement	Yes	No	Yes*	No*	No	No*	Yes	Yes
C. accomp.	Yes	Yes	Yes*	Yes	Irrel.*	Yes	Yes	Yes
C. semelfactive	Yes	No*	Yes*	No*	No*	No*	No	Yes
C. active accomp.	Yes	Yes	Yes	Yes	Irrel.*	Yes	Yes	Yes

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Test 1 deals with the [±static] feature to distinguish static verbs from non-static verbs by posing questions such as 'What has happened?' or 'What is happening?'. For causative verbs, these questions are problematic, since there are instances where they can be answered and the answer makes sense.

The second test is only possible in languages like English or Icelandic, where the progressive aspect of verbs exists (Van Valin, 2005: 35). Here staticity and punctuality are tested to determine whether a verb is an activity, an accomplishment or an active accomplishment. Virtually any state and achievement with singular subject cannot be used in this aspect, but a semelfactive can and, if so, the only way is an iterative reading. In fact, Van Valin (2005: 36) states that the progressive aspect "with a semelfactive verb (or an achievement verb with a plural subject) yields a verb which patterns like an activity verb".

Test 3 has to do with the co-occurrence of verbs with adverbs that signal dynamic action. Adverbs like *vigorously* or *actively* make that only activities and active accomplishments be compatible with them. Van Valin (2005: 36) does not recommend adverbs that "require a controlling subject, e.g., *deliberately*, *carefully*". Also, those semelfactives derived from activity verbs are found with this kind of adverbs (Cortés Rodríguez, González Vergara & Jiménez Briones, 2012: 63).

In test 4, punctual verbs are distinguished from non-punctual verbs. Thus, only activities, accomplishments and active accomplishments can co-occur with 'pace adverbs' such as *quickly* or *slowly*. These adverbs involve a temporal duration, but not necessarily a dynamic action. If we are to use these adverbs with semelfactives, the only possible reading is iterative. In fact, adverbs like *slowly* or adverbials like *in a slow pace* are preferred, since adverbs like *quickly* and adverbials like *in a quick pace* are compatible with achievements in every case.

Regarding tests 5 and 6, they show the difference between telic and non-telic verbs. Test 5 shows that states, accomplishments and active accomplishment have an internal duration. When determining the lexical aspect, we use the *for*-phrases as in *He read a book for an hour*, so semelfactives and achievements are only used with these phrases when they express very short duration, because they are [+punctual] and lack internal duration. On the other hand, test 6 refers to terminal points, i.e., whether verbs have an end and thus are telic. This test distinguishes those verbs with telic and internal duration features, so only accomplishments and active accomplishments can be used with *in*-phrases such as *He read a book in an hour*.

As for test 7, it attempts to differentiate between the two types of punctual predicates — achievements and semelfactives— by means of stative modifiers. Given that semelfactives have no result state, they cannot be used with a stative modifier. Achievements do have a result state and, consequently, we have examples like *the tapped window*, but not **the flashed light*. As Van Valin (2005: 38) states, this distinction is based on the fact that "semelfactives can have an iterative interpretation with a singular subject, while achievements can only have such a reading with a plural subject".

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Finally, test 8 is not a linguistic test per se, for there is "no simple syntactic test to determine whether a verb is inherently causative or not" (Van Valin, 2005: 38). Thus, we paraphrase sentences to know whether the underlying semantic structure is causative or not. Paraphrases need to have the same number of arguments as the sentence being paraphrased, so that we can have examples like *Pat causes Chris to come to have the book* as a paraphrase of *Pat gives the book to Chris*, but we cannot use **Leslie causes Leslie to run* as a paraphrase of *Leslie runs*.

As stated in the introduction, we will only deal with the semantic representation of verbs by means of the lexical aspect. Thus, next we briefly mention how RRG provides the rest of the semantic representation. Each *Aktionsart* is assigned a logical structure stemmed from formal semantics and then modified to make states and activities the basic logical structures from which the rest are derived by utilizing operators in uppercase. For instance, the clause *John has died* would be transcribed as BECOME **dead′** (John). After that, they are assigned at least one of the two macroroles that RRG proposes: actor generalizes roles such as agent or instrument and undergoer generalizes roles such as patient or theme. For a thorough description and discussion, see Van Valin and LaPolla (1997: 82-198) and Van Valin (2005: 31-67).

3. CORPUS AND METHOD

In this section, we briefly describe the corpus that we have analyzed, namely the Pitt corpus, and the method we have used to analyze it: descriptive statistical measures and a hypothesis testing.

3.1.The Pitt corpus

The Alzheimer Research Program of the University of Pittsburgh conducted a longitudinal study from 1983 to 1988, supported by the National Institute of Aging grants AG03705 and AG05133. This study is now part of DementiaBank, "a shared database of multimedia interactions for the study of communication in dementia", which is also part of TalkBank (Lanzi et al., 2023). The researchers recruited 319 healthy individuals and patients with Alzheimer's disease (AD) "to evaluate the full spectrum of behavioral neurologic factors" in these patients (Becker, Boller, Lopez, Saxton & McGonigle, 1994: 585). All these volunteers were subject to neuropsychiatric and laboratory tests so that they "provide a carefully screened and uniformly evaluated cohort of AD patients" (Becker et al., 1994: 586). The exclusion of patients was due to their not understanding what the study was evaluating. Thus, from the 319 individuals, 102 were assigned to the control group and 204 were included in the experimental group. There were 13 people that were considered "special" and thus were not part of the study (Becker et al., 1994: 586).

Print ISSN: 1578-7044; Online ISSN: 1989-6131

In 1992, there were 43 people out of the 75 that were diagnosed with "definite AD" after they had died, so the final group under study consisted of 282 people, with 181 individuals that were "probable and definite patients with AD" (Becker et al., 1994: 586). Ages varied, but there were more people closer to 60 and to 80 years. At the same time, there were 121 women in the final group, so it agrees with what was expected (Beam et al., 2018; Martinkova et al., 2021). Patients in the group were given an identification code: three numbers followed by a dash and then a number between 0 and 3 to signal the visits of researchers. For example, patient 325-0 refers to patient number 325, first visit, while patient 325-1 refers to the same patient but to second visit.

As noted before, patients were given neuropsychiatric tests, based on the Mini-Mental State Examination (MMSE; Folstein, Folstein & McHugh, 1975). These tests only measure "the cognitive aspects of mental functions", but they do not consider "mood, abnormal mental experiences and the form of thinking" (Folstein, Folstein & McHugh, 1975: 189). Points are assigned depending on the performance of patients in two sections: the first one "requires vocal responses only and covers orientation, memory and attention", while the second one records "the ability to name, follow verbal and written commands, write a sentence spontaneously, and copy a complex polygon" (Folstein, Folstein & McHugh, 1975: 190). The first section has a maximum score of 21 points and the second one has a maximum of 9 points; thus, the MMSE assigns a maximum score of 30 points. A score less than 20 points is "found essentially only in patients with dementia, delerium [*sic*], schizophrenia or affective disorder" (Folstein, Folstein & McHugh, 1975: 196). We also note that the Cochrane Collaboration sets 24 points for the detection of cognitive problems (Creavin et al., 2016).

The transcriptions of the Pitt corpus show that researchers visited the patients at least once and gave them four different tasks. Interviews were guided by using these tasks to test the cognitive abilities of patients and all patients received at least one of the four tasks. The first task (labelled "cookie" in the transcriptions) is a picture of two children trying to steal cookies from a jar. The second task ("fluency" in the transcriptions) is an entity-naming task to test how fluent patients are. The third task ("recall" in the transcriptions) tests patients' memory by making them remember a story. And in the fourth task ("sentence" in the transcriptions) patients have to utter a sentence based on 3 or 4 words that researchers give them. Last, there are 1047 individual transcriptions in the corpus: 309 in the cookie test, 238 in the fluency test, 262 in the recall test and 238 in the sentence test.

3.2.Method

The method we have followed is, first, the selection of those patients that have the "probable AD" tag in the transcriptions. This condition is set to ensure that patients are not diagnosed with mild cognitive impairment, vascular or other, and we consider the "probable AD" tag to

Print ISSN: 1578-7044; Online ISSN: 1989-6131

be stronger than the "possible AD" one as shown in the transcriptions. We have manually identified the verbal predicates for each patient in the four tasks, because using computational tools —e.g., regular expressions— to perform this task would require manual check in most cases. To carry out the analysis, we employ Microsoft Excel 2021 as the simplest way to achieve our objectives regarding statistics.

Second, we only considered verbal predicates because there are studies suggesting that lexical access of AD patients may revolve around verb usage (Davis et al., 2010; De Almeida et al., 2021). Furthermore, Paek, Murray and Newman (2020) suggest that verbs are cognitive markers. Since RRG is based on verb predicates, it seems reasonable to study how the analysis of the lexical aspect agrees with those lines of research. Thus, we counted the number of verb predicates in each task (cookie, fluency, recall and sentence; see Table 2) and found that the cookie test has 2632 predicates, the fluency test has 985, the recall test has 1979 and the sentence test has 1049, which amounts to 6645 predicates. We analyzed the three stages of Alzheimer's disease —early, intermediate and late— to describe the behavior of verbs. In doing so, we found that the early stage has 1024 predicates, the intermediate stage has 1009 and the late stage has 177, that is, a total of 2210 predicates (see Table 3). Given that the distribution is unbalanced, we took samples with a 95% confidence interval for each test, so we need to analyze 663 predicates in the cookie test, 468 in fluency, 591 in recall and 488 in sentence (see Table 4).

Table 2. Number of predicates per test.

Tests	Cookie	Fluency	Recall	Sentence	Total
Number of predicates	2632	985	1979	1049	6645

Table 3. Number of predicates per stage.

Stage	Early	Intermediate	Late	Total
Number of predicates	1024	1009	177	2210

Table 4. Samples per test with a 95% confidence interval.

Tests	Cookie	Fluency	Recall	Sentence	Total
Number of predicates	663	468	591	488	2210

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Third, we applied the 8 tests provided by RRG to obtain the lexical aspect of verbs, which are based on the Boolean values of four semantic features: static, dynamic, telic and punctual. When used sequentially, these tests ensure that verbs are mapped to one of the 12 possible classes of verbs of RRG: states, activities, achievements, accomplishments, active accomplishments, semelfactives and their respective causative counterparts.

Finally, we analyzed each test and each stage by means of standard measures used in descriptive statistics to know the frequency and distribution of *Aktionsarten* in the samples, that is, we have analyzed the range, mean, median, mode, standard deviation, sample variance, coefficient of variation and interquartile range of the verbs in the Pitt corpus. In this way, we would have a bird's-eye view of how lexical aspect is distributed in the samples. Furthermore, we applied Pearson's chi-squared test to assess whether both variables (*Aktionsarten* and stages) show independence or not.

4. RESULTS AND DISCUSSION

4.1. Results

In this section, we present the results of the analysis after we have applied the tests to determine the *Aktionsarten* of samples from the Pitt corpus and then the statistical measures to know their frequency and distribution. In Table 5, we show the counting of the three stages of Alzheimer's disease and, in Table 6, the counting of each task —cookie, fluency, recall and sentence. Percentages are shown with respect to the total of the stage or task. We give decimal numbers with two decimal places, so the sum of percentages may be slightly altered.

As can be seen in Table 5, states are the most used *Aktionsart* in the three stages (37.37%), followed by active accomplishments (23.71%), activities (20.76%), accomplishments (10.13%) and achievements (5.61%). Only in the late stage are activities (36 cases out of 177) more used than active accomplishments (32 cases). Semelfactives and causatives verbs are rarely used, and when ~~so~~ they are, they tend to be causative accomplishments (1.62%). Data show a strong asymmetry in the samples of *Aktionsarten*, where states are generally most used in the three stages with respect to what we would expect from a healthy individual, that is, more homogeneous data. This non-uniformity in the distribution of the lexical aspect may be related to the decline in cognitive abilities of patients over time. As examples of the verbs that we have analyzed from the corpus, we find states like *know*, *want* or *think*; activities like *do*, *fish* or *read* (intransitive); achievements like *lose*, *topple* or *drop* (intransitive); accomplishments like *get*, *fall down* or *slant*; active accomplishments like *dry*, *take* or *try*; and causative accomplishments are mainly verbs like *open* or *give*.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Table 5. Results for the early, intermediate and late stages.

Aktionsart	**Early**	**%**	**Intermediate**	**%**	**Late**	**%**	**Total**	**%**
State	352	34.37	399	39.54	75	42.37	826	37.37
Activity	215	20.99	208	20.61	36	20.33	459	20.76
Achievement	63	6.15	56	5.55	5	2.82	124	5.61
Semelfactive	0	0.00	2	0.19	1	0.56	3	0.13
Accompllishment	113	11.03	90	8.91	21	11.86	224	10.13
Active accomp.	254	24.80	238	23.58	32	18.07	524	23.71
Causative state	0	0.00	3	0.29	1	0.56	4	0.18
Causative activity	1	0.09	1	0.09	0	0.00	2	0.09
Causative achiev.	0	0.00	2	0.19	3	1.69	5	0..22
Causative semelf.	0	0.00	0	0.00	0	0.00	0	0.00
Causative accomp.	26	2.53	8	0.79	2	1.12	36	1.62
Causative act. acc.	0	0.00	2	0.19	1	0.56	3	0.13
Total	1024		1009		177		2210	

In Table 6, we observe that *Aktionsarten* are more irregular in the tasks than in the stages, since states are prevalent only in the fluency and recall tasks. In the cookie task, activities are the most used type of lexical aspect (160 cases out of 663), whereas in the sentence task active accomplishments represent 34.22% of verbs. Once more, semelfactives and causative verbs are rare, and the most relevant causative verbs are causative accomplishments (36 cases out of 2210). Again, we observe that data are not evenly distributed and states are generally the type of verbs that patients prefer, which may be linked to the symptoms of Alzheimer's disease.

Table 6. Results for the cookie, fluency, recall and sentence tasks.

Aktionsart	**Cookie**	**%**	**Fluency**	**%**	**Recall**	**%**	**Sentence**	**%**	**Total**
State	156	23.53	230	49.14	305	51.60	135	27.66	826
Activity	160	24.13	68	14.52	110	18.61	121	24.79	459
Achievement	39	5.88	30	6.41	35	5.92	20	4.09	124
Semelfactive	1	0.15	0	0.00	1	0.16	1	0.20	3
Accomplishment	127	19.16	25	5.34	42	7.10	30	6.14	224
Active accomp.	151	22.78	113	24.14	93	15.73	167	34.22	524
Causative state	1	0.15	1	0.21	1	0.16	1	0.20	4
Causative activ.	1	0.15	0	0.00	1	0.16	0	0.00	2
Causative ach.	3	0.45	0	0.00	1	0.16	1	0.20	5
Causative sem.	0	0.00	0	0.00	0	0.00	0	0.00	0
Causative acc.	23	3.46	0	0.00	1	0.16	12	2.45	36
Causative a. acc.	1	0.15	1	0.21	1	0.16	0	0.00	3
Total	663		468		591		488		2210

Print ISSN: 1578-7044; Online ISSN: 1989-6131

After we have identified the lexical aspect of verbs in the samples, we now apply descriptive statistical measures to know the relative frequency and distribution of the the lexical aspect. These statistical measures are range, mean, median, mode, standard deviation, sample variance, coefficient of variation and interquartile range. We summarize the results of the analysis per stage in Table 7 and per task in Table 8. Note that the mode we are looking for is a nominal value, not a number. Thus, although the most repeated value is 0 for those non-existent *Aktionsarten*, we show the most repeated *Aktionsart* in each case.

Table 7. Results of the statistical analysis per stage.

	Early	**Intermediate**	**Late**	**Total**
Range	352	399	75	399
Mean	85.33	84.08	14.75	275.33
Median	13.5	5.5	2.5	352
Mode	State	State	State	State
Standard deviation	122.35	129.89	22.99	175.07
Sample variance	14,970.78	16,871.90	528.75	30,652.33
Coefficient of variation	1.43	1.54	1.55	0.63
Interquartile range	138.5	117.5	22.75	162

In the three stages, states are the most repeated *Aktionsart*, as expected. We observe that the mean is between 4 and 5 times lower than the range, but the median is considerably lower: 13.5 in the early stage, 5.5 in the intermediate stage and 2.5 in the late stage; thus, most verbs revolve around a few classes of verbs (see Table 5) and thus verbs are not uniformly distributed. This heterogeneity is further established in both the standard deviations and sample variances —14,970.78, 16,871.90 and 528.75 in each stage, respectively. The coefficient of variation reflects that the standard deviations are much greater than the mean —43%, 54% and 55% greater, respectively—, as the interquartile range (IR) shows in the early and intermediate stages: the greater the IR, the more scattered the data. In the late stage, the IR is relatively small, which may be due to the utterance of fewer verbs.

These data per stage contrast with the data taken as a whole. The mean is almost 1.5 times lower than the range and the median is greater than in the three stages, although the standard deviation and the sample variance are still high. The coefficient of variation shows that the data from this sample are much less heterogeneous than shown in the stages separately.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Finally, the IR shows that the data are scattered in the sample. In other words, when analyzed separately, the data in the three stages seem disproportionately spread, but if we analyze the samples as a whole, the scattering and heterogeneity are less pronounced.

Table 8. Results of the statistical analysis per task.

	Cookie	**Fluency**	**Recall**	**Sentence**	**Total**
Range	160	230	305	167	305
Mean	55.25	39.00	49.25	40.66	46.04
Median	13	1	1	6.5	1
Mode	Activity	State	State	Act. accomp.	State
Standard deviation	70.25	69.78	89.27	62.05	71.45
Sample variance	4,936.20	4,869.81	7,969.29	3,850.60	5,105.70
Coefficient of variation	1.27	1.78	1.81	1.52	1.55
Interquartile range	132	39.5	53.75	52	25.25

Now we present the results of the analysis per task. First, the means are between 39.00 and 55.25, but the medians are much lower. This signals that most of the verbs are concentrated on a few *Aktionsarten* (see Table 5). Mode varies in each task: activities are mostly used in the cookie task, active accomplishments are mostly used in the sentence task and states are mostly used in both fluency and recall tasks. Standard deviations and sample variances are high in each task, which points at a great spreading of data. This is confirmed by the coefficients of variation, with standard deviations between 27% and 81% greater than the mean. The interquartile range is high, but tasks differ: in cookie, IR is disproportionately high, whereas in recall and sentence is moderately high. In any case, the four tasks show heterogeneity and scattered data. When taken as a whole, the samples reflect what the tasks show, except for a lower interquartile range that signals slightly smaller scattering.

Last, in the independence of variables test (Pearson's chi-squared test) we find that there exists a statistically significant relation. We first state the null hypothesis (H_0) and the alternative hypothesis (H_1) in the following terms:

H_0: There no is relationship between the lexical aspect and the stages of AD.

H_1: There is a relationship between the lexical aspect and the stages of AD.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Second, after we have computed the expected data in the sample, we calculate the squared difference of observed data and expected data, divided by the expected data. This operation is applied to every entry in the 12 rows and 3 columns matrix (see Table 9). Those cases where there are no observed nor expected data have been filled with a zero, given that the previous operation (0/0) is not defined.

Table 9. Squared difference between observed and expected data, divided by the expected data.

Aktionsart	**Early stage**	**Intermediate stage**	**Late stage**
State	2,466711952	1,269513287	1,182657726
Activity	0,025375045	0,011635734	0,01577575
Achievement	0,535108484	0,006649887	2,44853537
Semelfactive	1,390045249	0,290066476	2,402230062
Accomplishment	0,817258209	1,472040613	0,52183929
Active accomp.	0,517151839	0,00640645	2,367300057
Causative state	1,853393665	0,754391024	1,441830918
Causative activity	0,005798395	0,008265879	0,160180995
Causative achiev.	2,316742081	0,03503536	16,87502876
Causative semelf.	0	0	0
Causative accomp.	5,206801667	4,330043385	0,270577441
Causative act. acc.	1,390045249	0,290066476	2,402230062

Third, if we add all the entries and then calculate the Pearson's test, we obtain a value of 55.08. Since there are 12 rows and 3 columns in this matrix, the degrees of freedom are 22:

$$df(r, c) = (\mathrm{r} - 1)\cdot(c - 1) = (12 - 1)\cdot(3 - 1) = 11\cdot 2 = 22$$

Last, the critical value of this matrix is 33.92, taking into account the degrees of freedom and a significance level of $\alpha = 0.05$. In this way, we obtain a p-value = $0.00011 < 0.05 = \alpha$. Therefore, we cannot accept the null hypothesis and, thus, we conclude that there exists a dependency relationship between lexical aspect and the stages of Alzheimer's disease in this corpus.

4.2.Discussion

We can state that the samples we have analyzed are not homogeneous, that patients tend to use three types of verbs, i.e., states, activities and active accomplishments, and that states are the preferred *Aktionsart*. Furthermore, we have shown that there is a statistically significant relationship between lexical aspect and the stage of Alzheimer's disease in this corpus. There are some questions we would like to address that concern the results of this analysis.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

First, states are the preferred *Aktionsart* and this systematic use by patients in the three stages needs explanation. Our approach is that states are cognitively easier to process, as Gennari and Poeppel (2002) have suggested. Thus, patients tend to use states as the default *Aktionsart*. As stated in the results, the distribution may be related to the states of affairs that patients try to convey, not only because they want to express those specific states of affairs, but because their ability to describe them is aggravated over time. This proposal is based on the assumption that a healthy individual uses a more stable or uniform distribution of *Aktionsarten*, where states are still the preferred verb class. When the symptoms of Alzheimer's disease appear, the patient starts to use more states at the expense of the other *Aktionsarten*. The confirmation or refutation of this hypothesis is beyond the aims of this paper; however, research in other kind of individuals, e.g., foreign language students or children with specific language impairment, may clarify whether there exist differences in cognitive processing of lexical aspect. So far, there are no studies on the lexical aspect applied to corpora of English-speaking patients to the same extend as studied in this paper.

Second, we highlight the role that active accomplishments and activities play in these samples. They are the second and third types of verbs, but their use differs per stage and per task. On the one hand, active accomplishments are preferred in the early and intermediate stage, but not in the late stage. In general, active accomplishments are more used than activities, as Table 5 shows. On the other hand, in the cookie and recall tasks, activities are preferred, but in the fluency and sentence tasks, active accomplishments have more utterances. Given that these two *Aktionsarten* only differ in one semantic feature, i.e., [±telic], we wonder whether they are activated in the same brain region. Similarly, these results for activities and active accomplishments imply that the type of task is related to the use of these *Aktionsarten*, given that the answer to, e.g., the cookie task, is qualitatively different from the sentence task. In De Almeida et al. (2021), they show that patients with Alzheimer's disease perform better in the "dynamic scenes", which are related to events. In Lara, Beltrán, Rodríguez and Araque (2016), it is suggested that dynamic perception is similar both in patients with Alzheimer's disease and healthy individuals, but not in the static perception. With these two studies, we would expect a greater usage of dynamic verbs (activities, active accomplishments and activity-based semelfactives); however, our data show the opposite.

Third, achievements and accomplishments differ in the [±punctual] feature and in that achievements yield a result state. In the three stages, achievements are less than 6% of all *Aktionsarten*, whereas accomplishments are more than 10%. However, a closer look shows that achievements decrease from the early stage (6.15%) to the late stage (2.82%), while accomplishments have a slight increase from the early stage (11.03%) to the late stage (11.86%). If we look at the tasks, we observe that the pair achievement-accomplishment varies. In cookie, achievements represent less than 6%, but accomplishments are 19.16%; in fluency, achievements are 6.41%, while accomplishments are 5.34%; in recall, achievements are

Print ISSN: 1578-7044; Online ISSN: 1989-6131

5.92%, but accomplishments are 7.10%; and in sentence, achievements are 4.09%, but accomplishments are 6.14%. All this leads us to wonder whether patients with greater cognitive impairment tend to not perceive changes of state and thus they instead tend to use more states and activities/active accomplishments. This explanation may account for the low number of achievements and accomplishments, except for the cookie task, and also account for the decrease in achievements over the stages. As stated above, the use of different tasks may determine the kind of lexical aspect that patients utter, in general. Last, it seems that patients utter *Aktionsarten* in pairs, at least in the achievement-accomplishment case, in the light of the percentages of usage.

Fourth, causative verbs are less than 2.30% of all *Aktionsarten*, although most of them are causative accomplishments and they are mostly found in the early stage. In fact, causative verbs represent 1.22% of *Aktionsarten* in the early stage, while they are 0.72% in the intermediate stage and 0.31% in the late stage. Following the idea in the previous paragraph, we propose that the low number of causative verbs is due to the inability of patients to think of states of affairs that have been caused by other states of affairs; as a consequence, patients simply describe static states of affairs, i.e., they perceive reality as a picture and describe it by means of static verbs. Examples of this can be found in the fluency test for patient 018-0 —"There are a whole bunch more" of animals— and in the sentence test for patient 350-0 —"There's a tree in my yard". Patients appear to express states of affairs as if they were individual entities with no relation to other states of affairs and with no changes, either internal or external. Therefore, causative verbs would be more frequent in the early stage than in the late stage, where cognitive impairment is greater. This explanation is only based on the data of this corpus, but we consider it to be enough to account for the results in our analysis.

We highlight that the hypotheses we have provided stem from the data collected from the Pitt corpus, so further research in the same or similar line may confirm or refute what we have showed. Moreover, a limitation may reside in the use of four different tasks that patients of this corpus have received, in the sense that the distribution of the lexical aspect throughout the corpus is conditioned by the tasks. On the other hand, by studying other corpora —students of foreign languages, patients with Parkinson's disease, patients with Lewy's bodies…—, the cognitive aspects of lexical aspect may be better understood. In the same vain, the analysis of verbs has followed the tests in Van Valin (2005: 35-39) to establish six verb classes and their causative versions. However, Van Valin suggests the addition of a seventh class —processes. We have not considered this other class of verbs and it may better account for differences in semantic features. For instance, verbs that we have assigned the accomplishment *Aktionsarten* might actually be processes, because one of the features is different. This issue is also considered in Cortés Rodríguez, González Vergara and Jiménez Briones (2012: 67-68). A revision of our analysis that includes this new semantic feature may make our study complete.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

5. CONCLUSION

We have applied the eight tests provided by Role and Reference Grammar to determine the *Aktionsart* or lexical aspect of samples of verbal predicates from the Pitt corpus of patients with Alzheimer's disease. Then, we have analyzed the verb classes by means of descriptive statistics and hypothesis testing, finding that, in general, patients use states as the default *Aktionsart* and that activities and active accomplishments vary depending on the task — cookie, fluency, recall and sentence. The statistical analysis shows that the data from the sample is not homogeneous and that verbs revolve around a few types of *Aktionsart*, especially states. On the other hand, we have shown that in this corpus there is a statistically significant relations between lexical aspect and the stages of Alzheimer's disease. Given these results, we offer explanations for the variability of the data in the samples, such as the possibility that patients may have lost their ability to perceive changes of state or to conceptualize dynamic states of affairs. In the same vain, these results point to the opposite direction with respect to previous studies that relate Alzheimer's disease and perception of states of affairs. These hypotheses must be tested in the future in studies that go along the same or similar line of research.

REFERENCES

Atri, A. (2019). The Alzheimer's Disease Clinical Spectrum: Diagnosis and Management. *Medical Clinics*, *103*(2), 263-293. doi: 10.1016/j.mcna.2018.10.009

Beam, C. R., Kaneshiro, C., Jang, J. Y., Reynolds, C. A., Pedersen, N. L. & Gatz, M. (2018). Differences Between Women and Men in Incidence Rates of Dementia and Alzheimer's Disease. *Journal of Alzheimer's Disease*, *64*(4), 1077-1083. doi: 10.3233%2FJAD-180141

Becker, J., Boller, F., Lopez, O., Saxton, J. & McGonigle, K. (1994). The natural history of Alzheimer's disease: description of study cohort and accuracy of diagnosis. *Archives of Neurology*, *51*(6), 585-594. doi: 10.1001/archneur.1994.00540180063015

Cortés Rodríguez, F. J., González Vergara, C. & Jiménez Briones, R. (2012). Las clases léxicas. Revisión de la tipología de predicados verbales. In R. Mairal, L. Guerrero & C. González Vergara (Coords.), *El funcionalismo en la teoría lingüística: la Gramática del Papel y la Referencia* (pp. 59-84). Madrid: Akal.

Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., … Cullum, S. (2016). Mini-Mental State Examination (MMSE) for the Detection of Dementia in Clinically Unevaluated People Aged 65 and Over in Community and Primary Care Populations. *Cochrane Database of Systematic Reviews*, *1*.

Crystal, D. (1981). *Clinical Linguistics*. New York: Springer-Verlang Wien.

Davis, C., Heidler-Gary, J., Gottesman, R. F., Crinion, J., Newhart, M., Moghekar, A., … Hillis, A. E. (2010). Action versus animal naming fluency in subcortical dementia, frontal dementias, and Alzheimer's disease. *Neurocase*, *16*(3), 259-266. doi: 10.1080%2F13554790903456183

De Almeida, R. G., Mobayyen, F., Antal, C., Kehayia, E., Nair, V. P. & Schwartz, G. (2021). Category-specific verb-semantic deficits in Alzheimer's disease: Evidence from static and dynamic action naming. *Cognitive Neuropsychology*, *38*(1), 1-26. doi: 10.1080/02643294.2020.1858772

Dowty, David. (1979). *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Gallardo Paúls, B. & Valles González, B. (2008). Lingüística en contextos clínicos: la lingüística clínica. *Lengua y Habla*, *12*(1), 32-50.

Print ISSN: 1578-7044; Online ISSN: 1989-6131

Garayzábal Heinze, E. (2009). Lingüística Clínica: Teoría y práctica. In J. L. Jiménez Ruiz & L. Timofeeva (Eds.), *Estudios de Lingüística: Investigaciones Lingüísticas en el siglo XXI* (pp. 131-168). Alicante: Universidad de Alicante.

Ivanova, O., García Meilán, J. J., Martínez Nicolás, I. & Llorente, T. E. (2020). La habilidad léxico-semántica en la Enfermedad de Alzheimer: Un estudio de la fluidez verbal con categorías semánticas. *Revista Signos*, *53*(102), 319-342. doi: 10.4067/S0718-09342020000100319

Jensen, S. (2007). Aspect and child language acquisition: a Danish perspective. *Tidsskrift for Sprogforskning*, *4*(1), 209-219.

Lara, M. F., Beltrán, J. C., Rodríguez, S. R. & Araque, S. M. (2016). Análisis de la percepción de eventos estáticos y dinámicos en personas con enfermedad de Alzheimer. *Universitas Psychologica*, *15*(5). doi: 10.11144/Javeriana.upsy15-5.apee

Lanzi, A. M., Saylor, A. K., Fromm, D., Liu, H., MacWhinney, B., & Cohen, M. (2023). DementiaBank: Theoretical rationale, protocol, and illustrative analyses. *American Journal of Speech-Language Pathology*. doi.org/10.1044/2022_AJSLP-22-00281

Martinkova, J., Quevenco, F.-C., Karcher, H., Ferrari, A., Sandset, E. C., Szoeke, …, Ferretti, M. T. (2021). Proportion of women and reporting of outcomes by sex in clinical trials for Alzheimer disease. *JAMA Neurology*, *4*(9). doi: 10.1001/jamanetworkopen.2021.24124

Paek, E. J., Murray, L. L. & Newman, S. D. (2020). Neural Correlates of Verb Fluency Performance in Cognitively Healthy Older Adults and Individuals With Dementia: A Pilot fMRI Study. *Frontiers in Aging Neuroscience*, *12*. doi: 10.3389/fnagi.2020.00073

Pérez Cabello de Alba, M. B. (2017). A contribution of Natural Language Processing to the study of semantic memory loss in patients with Alzheimer's disease. *Revista de Lenguas para Fines Específicos*, *23*(2), 133-156.

Pérez Cabello de Alba, M. B. & Teomiro García, I. I. (2018). Characterizing semantic memory loss: Towards the location of language breakdown. *Círculo de Lingüística Aplicada a la Comunicación*, *76*, 219-240. doi: 10.5209/CLAC.62506

Perkins, M. (2011). Clinical Linguistics: Its Past, Present and Future. *Clinical Linguistics & Phonetics*, *25*(11-12), 922-927. doi: 10.3109/02699206.2011.599471

Perkins, M. & Howard, S. (2011). Clinical Linguistics. In J. Simpson (Ed.), *The Routledge Handbook of Applied Linguistics* (pp. 112-124). Nueva York: Routledge.

Shain, R. (2011). Exploring Aktionsart in Corpora: A Case Study of Koine Greek Erchomai and Eiserchomai. *Journal of Greek Linguistics*, *11*(2), 221-248. doi: 10.1163/156658411X600016

Smith, C. (1997). *The Parameter of Aspect*. Dordrecht: Reidel.

Suárez-Rodríguez, A. (in press). Language Impairment in English-Speaking Patients with Alzheimer's Disease: A Role and Reference Grammar Contribution. *Atlantis. Journal of the Spanish Association for Anglo-American Studies*.

Van Valin, R. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge UP.

Van Valin, R. & LaPolla, R. (1997). *Syntax: Structure, meaning and function*. Cambridge: Cambridge UP.

World Health Organization. (2023). *Dementia*. Retrieved 30 May, 2023 from https://www.who.int/news-room/fact-sheets/detail/dementia.

Print ISSN: 1578-7044; Online ISSN: 1989-6131