A longitudinal linguistic analysis of written text production in a case of semantic variant primary progressive aphasia

Nari Heitkamp a,1, Rahel Schumacher b,1,2, Karen Croot c, Ernst G. de Langen d, Andreas U. Monsch e, Thomas Baumann f,2, Adrian Danek a,*

a Neurologische Klinik, Ludwig-Maximilians-Universität, Postfach 701260, 81366 München, Germany
b Division of Cognitive and Restorative Neurology, Department of Neurology, Inselspital, Bern University Hospital, and University of Bern, Freiburgstrasse, 3010 Bern, Switzerland
c School of Psychology, Brennan MacCallum Building (A18), University of Sydney, NSW 2006, Australia
d Reha-Zentrum Passauer Wolf, Bgm.-Hartl-Platz 1, 94086 Bad Griesbach, Germany
e Memory Clinic, University Center for Medicine of Aging, Felix Platter Hospital, Schanzenstrasse 55, 4031 Basel, Switzerland
f Praxisgemeinschaft für Neurologie, Thunstrasse 95, 3006 Bern, Switzerland

ARTICLE INFO

Article history:
Received 21 February 2015
Received in revised form 17 December 2015
Accepted 18 December 2015
Available online xxx

Keywords:
Semantic dementia
Semantic variant of primary progressive aphasia
Written text production
Semantics
Syntax
Morphology

ABSTRACT

The semantic variant of primary progressive aphasia (svPPA) presents with a degradation of semantic knowledge due to atrophy of the anterior temporal lobe and is characterized by impaired confrontation naming and impaired single-word comprehension. So far, little is known about the development of symptoms and their order of occurrence in the pre-clinical phase, and information regarding written text production is scarce.

We had the unique opportunity to analyze the diary of a man written over a time span of 12 years before he was diagnosed with svPPA. We sought to identify the earliest indicators of cognitive change in his diary entries, and to track the important changes over time. Based on transcripts of the entries (one week every six months) we assessed the overall structure, vocabulary, surface dysgraphia and semantic paraphasia, syntax, and morphology. We found changes in all domains up to seven years before the clinical diagnosis. The earliest changes concerned the vocabulary, with decreased variety and increased use of high frequency words. This was followed by syntactic and morphological errors. We found no increase of surface dysgraphia. Semantic paraphasias increased only during the last three years but characterized the entries of the last year.

We were therefore able to further corroborate recent findings regarding difficulties in the morpho-syntactic domain in this patient group. In this natural context for written text production, such errors seem, in addition to changes in vocabulary, to be the first error types to appear, possibly as a result of compensating for the degradation of semantic representations.

© 2015 Elsevier Ltd. All rights reserved.
1. Introduction

Dementia, a decline of mental abilities that interferes with activities of daily life, is caused by various neurodegenerative pathologies and can feature language impairments. Whereas Alzheimer’s dementia (AD), the most common form of dementia, is characterized by memory impairments, and language impairments usually emerge only during disease progression, the latter are the core symptoms in a group of syndromes termed primary progressive aphasia (PPA) first proposed by Mesulam (1982).

The current consensus (Gorno-Tempini et al., 2011) differentiates three forms of PPA: 1) the non-fluent/agrammatic variant (nfvPPA), 2) the logopenic variant (LvPPA) and, 3) the semantic variant (svPPA, also known as semantic dementia). The prominent abnormalities in svPPA concern semantic memory function and result from atrophy of the anterior temporal lobes that causes a degradation of semantic representations (Jefferies & Lambon Ralph, 2006; but see Bright, Moss, Stamatakis, & Tyler, 2008). This degradation leads to impaired confrontation naming and impaired single-word comprehension, which are core and obligatory symptoms of svPPA (Gorno-Tempini et al., 2011). In naming, patients may make coordinate or superordinate errors, which is qualitatively different to individuals with stroke aphasia (Jefferies & Lambon Ralph, 2006). Other indicators of svPPA are surface dyslexia or dysgraphia (i.e., regularization of orthographically irregular words), spared repetition, and spared speech production.

However, as pointed out by Sajjadi, Patterson, Tomek, and Nestor (2012), the diagnostic examination is mostly conducted at the single word level and inaccurate conclusions might be drawn about a patient’s performance at the discourse/text level. The spoken language performance of svPPA patients is described as fluent but empty, full of overlearned phrases and semantic paraphasias (i.e., replacement of the target by a semantically related item) (Reilly, Rodriguez, Lamy, & Neils-Strunjas, 2010; Warren, Rohrer, & Rosser, 2013). To a certain extent, connected speech can sound quite normal (Sajjadi et al., 2012) but patients tend to lose their differentiated vocabulary over time and use more general and more frequent instead of semantically precise words (Ash et al., 2006; Bird, Lambon Ralph, Patterson, & Hodges, 2000; Bozeat et al., 2003; Fraser et al., 2014; Laisney et al., 2011). Apart from the changes in vocabulary, errors of syntax and morphology may also occur in the spoken language of individuals with svPPA, as has — somewhat contrary to initial expectations — been pointed out recently (Auclair-Ouellet, 2015; Benedet, Patterson, Gomez-Pastor, & La Rocha, 2006; Gorno-Tempini et al., 2011; Kavé, Leonard, Cupit, & Rochon, 2007; Leyton & Hodges, 2014; Meteyard & Patterson, 2009; Meteyard, Quain, & Patterson, 2014; Silveri et al., 2014; Thompson et al., 2012; Wilson et al., 2010, 2014). Individuals with svPPA seem to show syntactic abnormalities in the sense of paragrammatic constructions such as sentence inversions or duplications and substitutions on the morphological level. Error rates are usually higher for atypical, less familiar or infrequent constructions.

Despite the wide-ranging investigation of language impairments associated with dementia that started more than 30 years ago (Appell, Kertesz, & Fisman, 1982; Mesulam, 1982; Snowden, Goulding, & Neary, 1989; Warrington, 1975) two aspects are still relatively neglected. First, there is a lack of knowledge about the development of symptoms and their order of occurrence in the preclinical phase of different types of dementia, and this is especially true of the less well-known syndromes. Second, written text production has barely been explored in patients with dementia.

Only three studies have, to the best of our knowledge, investigated the written text production of individuals later diagnosed with AD. Two studies conducted computer-based analyses of prose texts by professional writers. The analysis of renowned Irish author Iris Murdoch’s work (Garrard, Maloney, Hodges, & Patterson, 2005) revealed an increased use of higher frequency words, a less varied vocabulary, and a shallower and more banal plot in her last novel, which was written around the time AD was diagnosed, four years before her death. Word length, use of different word classes, the overall structure of the texts, and syntactic composition did not differ between three novels from the beginning, prime and end of her writing career. Similar results and additional evidence for difficulties in finishing the novel (writer’s block) were found in the analysis of texts from the Dutch writer Gerard Reve, who also suffered from AD (van Velzen & Garrard, 2008). Abnormalities of language use in written text production were detectable years before the clinical diagnosis of dementia. This was also the case in a study by Snowden et al. (1996), the so-called nun study. The analyses of a single, hand written page produced about 58 years earlier, enabled the researchers to identify the nuns later affected by AD. Their texts showed a lower rate of idea density (i.e. the quantity of propositions/ideas per sentence) and a lower degree of syntactic complexity. The nuns were not professional writers, which might explain why lower syntactic complexity was found very early in their texts but not in Murdoch’s and Reve’s work.

To date there do not appear to have been any longitudinal investigations of written text in svPPA. In the present study, we had the opportunity to analyze the written texts of a man, HK, who was diagnosed with svPPA 12 years after he began to write daily diary entries. We studied the texts to identify the onset of linguistic abnormalities, and to document the “linguistic fingerprint” of the disease in this individual over time. We also considered whether HK’s written language would parallel the patterns found in the spoken language production of individuals with svPPA, and to what extent the changes in his written texts would be comparable to changes reported in the written language of individuals with probable AD. Our analyses included the overall structure, vocabulary, surface dysgraphia and semantic paraphasia, syntax, and morphology.

On the basis of the findings regarding spoken language production in individuals with svPPA and written language production by individuals with AD cited above, we expected various changes on different levels. Regarding the overall structure, we expected no change in the composition of the diary entries but a decrease in the number and length of entries, and a trivialization of their topics. Furthermore, we expected that HK’s diary entries would show a decreasing variety of vocabulary over time as well as an increased use of higher frequency words. As surface dysgraphia and semantic paraphasia
are regarded as typical abnormalities in written text production of individuals with svPPA (Gorno-Tempini et al., 2011), we expected that these abnormalities would arise early and increase over time. The semantic paragraphias should also become increasingly semantically distant from the target item over time. Regarding syntax and morphology, we expected to find a high rate of simple constructions and paragrammatic errors.

2. Materials and methods

2.1. Patient HK

HK was a right-handed accountant of Swiss origin with 12 years of education. He was admitted to the Basel Memory Clinic in 2009, 12 years after his retirement at the age of 65. HK had consulted his general practitioner about 2.5 years earlier than that, reporting memory problems (see Czarnecki et al., 2008). After this, he started to take the medication “Symfona”, capsules made from ginkgo extract. His partner reported that the main problem was word-finding difficulty that had been increasing for three years. Furthermore, she had to repeat things several times and she noticed that he had difficulties in recalling his friends’ names. Moreover, he had given up answering the phone and playing the organ. She had also started to support him in financial matters in order to prevent him giving away large sums to strangers.

Detailed general medical, neurological, neuropsychological and neurolinguistic examinations were performed. Magnetic resonance imaging (see Fig. 1) showed an asymmetric cerebral atrophy, predominantly affecting the left anterior temporal lobe with so-called knife blade appearance of the temporal gyri. Temporal lobe atrophy included lateral and medial structures, i.e. superior, middle and inferior temporal gyrus, amygdala, entorhinal cortex and hippocampus. Mild atrophy was present in the left frontal lobe. Consequently, a marked widening of the left lateral ventricle was seen.

Fig. 1. Transverse and coronary brain magnetic resonance images of patient HK taken at the age of 77 years (after he stopped writing diary). Images are oriented according to radiological convention, i.e. left is right and right is left.
impairment. However, together with the impairments in some areas of executive functioning (flexibility, behavior), they may also be attributed to the progression of the neurodegenerative pathology towards medial temporal and orbitofrontal structures (see Fig. 1), as recently described (Bonner, Ash, & Grossman, 2010; Brambati et al., 2015; Guo et al., 2013; Iaccarino et al., 2015; Karageorgiou & Miller, 2014; Murre, Graham, & Hodges, 2001).

A diagnosis of svPPA was made on the basis of the neurological results in combination with the neuropsychological and neurolinguistic findings.

2.2. Diary

HK started his diary after retirement and wrote entries over 12 years. The almanac format changed over time (see Fig. 2). In years 1–2, he dedicated 2 pages to 1 week (format A), in years 3–6, 4 days were described on 2 pages (format B), and for years 7–12, 1 day per page was typical (format C).

HK developed a consistent structure for his entries, starting out with a short comment about the weather, followed by the temperature, which was documented up to 6 times in a single entry. These elements can be seen as habits that constitute a so called “barrier ritual” (Surd-Büchele, 2011), i.e. a ritual supporting the diving in and out of the writing situation. Next, HK...
reported on the activities of the day and finally, he wrote down the time of sunrise and sunset. He additionally documented extreme weather conditions, and in the summer entries he mentioned the number of flowering lilies and frogs in his pond. His writing style can be described as “objective” (Surd-Büchele, 2011). HK recorded events and activities of the day, often with details about time and costs. Emotional states or thoughts were rarely included. For a few years HK went on a daily walk for about an hour, which he described in increasing detail over time. In most cases, he encircled the time he needed for his walk and marked it with color. He did the same with the number of flowers in his garden, frogs in his pond, and with descriptions of extreme weather conditions. HK’s diary started on January 1, 1997 and ended on December 29, 2008. An example of a typical diary entry and HK’s last entry are shown in Fig. 2, together with their transcription and translation.

2.3. Procedure

In contrast to other studies that examined texts (Garrard et al., 2005; van Velzen & Garrard, 2008) the diary entries were not available in digital format in our study. Thus, in order to make an analysis feasible, it was necessary to analyze only a
sample of the entries. We chose to transcribe one whole week every 6 months, the last seven days of June and December each year. In order to have a complete set of 7 days, we added single days from the previous or following weeks in December of year 12, where HK had left some days blank. In total, 168 entries were transcribed and two raters examined various features relating to overall structure, vocabulary, surface dysgraphia and semantic paraphasia, syntax, and morphology based on the transcripts. An overview of the error types with examples and translations is shown in Table 2. One of the raters speaks Swiss German and was able to identify dialectal habits so these were not counted as errors.

2.3.1. Overall structure and vocabulary

The assessment of the overall structure included counting the number of missing or unfinished entries in the whole year 12 (final year) and the words per entry (referred to as tokens) of the transcribed weeks. Dividing the number of new words used (types) by the number of tokens results in the type-token-ratio (TTR), an indicator of the variety of the vocabulary. A higher TTR indicates a more variable vocabulary. Furthermore, word frequencies were analyzed by using the CELEX database (Baayen, Piepenbrock, & Van Rijn, 1993) and a computer-based program provided by Aichert, Marquardt, and Ziegler (unpublished). Before scanning the texts with CELEX, all punctuation marks and formatting were removed, and umlauts were rewritten. Additionally, all the orthographical mistakes were corrected and HK's idiosyncratic abbreviations written in full (e.g. Rest. standing for “restaurant”). Some dialectal expressions, such as Zmorge essen [having breakfast], were substituted by standard German forms in order to allow the program to also analyze these words. The median word frequencies per week were used for statistical analyses.

2.3.2. Surface dysgraphia and semantic paragraphia

Surface dysgraphias and semantic paragraphias were counted. If a word’s spelling deviated from orthographical convention without altering the sound, this was considered as surface dysgraphia. If a word was understandable, but not precise enough within the context, this was considered as semantic paragraphia. For each week and error type, the total number of these errors was computed and their percentage with respect to the total number of words was then calculated.

As a high use of quotation marks, not only for citing, was noted during the transcription process, these were also counted. The sum of inappropriate quotation marks was computed and their percentage with respect to the total number of words was then calculated for each week.

2.3.3. Syntax

On the syntactic level we analyzed the syntactic structure and (par)agrammatic errors. The syntactic structure was examined indirectly on the basis of HK’s use of subordinate clauses by analyzing his use of connectives (connectives refers to sentence-linking words here). Since these do not form a syntactically homogeneous class, but are defined by their function, the connectives analyzed are a selection of the ones used by HK, and hence no claim for completeness is made. For the analyses of the connectives, we classified these as simple versus complex (simple: and, then; complex: e.g. in spite of, because). The classification was based on results about cognitive complexity and the age of connector use in language acquisition (Reimann, 1996, n.d.; Van Veen, 2011). Syntactic errors comprised sentence inversions or repetition of (parts of) sentences, as well as word elisions (Klann, 2001; Tesak, 1990).

2.3.4. Morphology

The level of morphology comprised errors in inflectional morphemes (in German very often used to indicate number, gender, and case), and unusual use of function words such as prepositions and articles. For each week and error type, the sum of morphological errors was computed and their percentage with respect to the total number of words was then calculated.

2.3.5. Statistical analysis

After completion of the transcription and rating procedures, the obtained measures were statistically analyzed. The data were divided into 6 timespans of two years each, during which the same almanac style (A, B, or C) was used (span A = year 1&2; span B1 = year 3&4; span B2 = year 5&6; span C1 = year 7&8; span C2 = year 9&10; span C3 = year 11&12). Changes related to the different almanac styles would become visible in significant changes between timespans A and B1, and between timespans B2 and C1. Changes in these successive spans were identified using Mann–Whitney tests. To assess whether there were differences between the spans and whether these followed a trend over time, the scores from the different measures were subjected to the non-parametric Kruskal–Wallis and Jonckheere’s tests, once considering all 12 years and once considering the last six years (C1–C3) only. The level of statistical significance was set at p < 0.05, two-tailed.

3. Results

3.1. Overall structure and content

The analysis of the overall structure of the entries revealed that HK developed a consistent composition style for his daily texts over the years. In the first two years, his notes looked like a memory back-up for appointments or events. Progressively, his entries became more prosaic, narrating about each day (see Fig. 2 for an example page). Noting the temperature(s) was first seen in year 2, time of sunrise/sunset in year 6. Over the timespan of 12 years, we observed a change of topics in the texts.
of HK. Initially, he mostly noted distinctive and extraordinary events. Subsequently, he described his daily walks and the weekly grocery shopping in more and more detail. Later, he also described other daily routines in more detail, such as dinner, going to bed and Lichterlöschen [turning off the lights]. Missing or unfinished entries appeared only towards the end of the last year (see Fig. 3).

The analysis of the entry length per week (total number of words/tokens) revealed a median entry length of 358.5 tokens \((A = 70, B1 = 194.5, B2 = 288.5, C1 = 475.5, C2 = 638.5, C3 = 669)\). There was a significant effect of span \((H(5) = 20.49, p < 0.01)\), and Jonckheere’s test revealed a significant increase of entry length over the 12 years \((J = 227, p < 0.01)\). Post-hoc tests revealed significant differences between span A and B1, the almanac change thus lead to significantly longer entries \((U = 0, z = -2.3, p = 0.021)\). The second almanac change between span B2 and C1 yielded also a significant increase in entry length \((U = 0, z = -2.3, p = 0.021)\). A further increase was noted between span B1 and B2 \((U = 2, z = -1.7, p = 0.043)\), even without an associated change of almanac style. Within the last six years (spans C1 to C3) there was no significant change of entry length \((J = 36, p = 0.76)\). The change of the entry length over time is depicted in Fig. 4.

### 3.2. Vocabulary

The analysis of the type-token-ratio (TTR) as a measure of vocabulary variety revealed a median TTR of 0.5 \((A = 0.74, B1 = 0.67, B2 = 0.61, C1 = 0.48, C2 = 0.41, C3 = 0.35)\). There was a significant difference between the six spans \((H(5) = 20.87, p < 0.01)\) and Jonckheere’s test revealed a significant decrease of TTR over the 12 years \((J = 10, p < 0.01)\). Post-hoc tests revealed a significant decrease between B2 and C1 \((U = 0, z = -2.3, p = 0.021)\). Moreover, there was a significant decrease of TTR within the last six years \((J = 6, p < 0.01)\). The decrease in TTR might thus only partially be due to the change in almanac style.

Analysis of word frequencies revealed a median frequency of 91.75 \((A = 16.25, B1 = 45.5, B2 = 80.5, C1 = 91.75, C2 = 108.25, C3 = 237.5)\). There was a significant effect of timespan \((H(5) = 18.83, p < 0.01)\) and Jonckheere’s test revealed a significant increase \((J = 219.5, p < 0.01)\), thus an increased use of high frequency words, over the 12 years. The increase of high frequency word use was also significant within the last six years \((J = 46, p < 0.01)\). Post-hoc tests showed no significant differences associated with the almanac style change. Median word frequencies and the TTR are also depicted in Fig. 4.

Furthermore, we observed an inappropriate use of quotation marks that became more inadequate over time. Dialect words like Zmorge [breakfast] or Znacht [dinner] as well as common words (e.g., post, bread) appeared more often with quotation marks. In total, 71 instances of inappropriate quotation marks were noted \((A = 0, B1 = 0, B2 = 3, C1 = 18, C2 = 22, C3 = 28)\), with the first occurrence in year 5. The percentage of inappropriate quotation marks differed significantly between timespans \((H(5) = 18.04, p < 0.01)\) and there was a significant increase over time \((J = 200, p < 0.01)\), but not within the last six years \((J = 27, p = 0.661)\).

### 3.3. Surface dysgraphia and semantic paragraphia

Analysis of the small number of surface dysgraphic errors (6 in total) revealed no significant effect of timespan \((H(5) = 7.20, p = 0.21)\) on the percentage of surface dysgraphic errors and no trend within the last six years \((J = 28, p = 0.51)\). There were 39 instances of semantic paragraphias in the transcribed entries, the first in year four \((A = 0, B1 = 2, B2 = 1, C1 = 2, C2 = 8, C3 = 26)\). There was a significant effect of timespan on the percentage of semantic paragraphias \((H(5) = 12.89, p = 0.03)\) which also showed a significant increase over the 12 years \((J = 177.5, p < 0.01)\). The increase was furthermore significant within the last six years \((J = 40, p = 0.02)\). The percentages of both types of errors are depicted in Fig. 5.

Qualitative analyses showed that the relation between the semantic paragraphias and the target items became more distant over time.

### 3.4. Syntax

The analyses of HK’s use of connectives revealed that und [and] and dann [then] outnumbered the other connectives \((A = 3, B1 = 18, B2 = 37, C1 = 89, C2 = 130, C3 = 130; A = 0, B1 = 0, B2 = 4, C1 = 9, C2 = 18, C3 = 37)\). The occurrence of these two simple connectives differed significantly between timespans \((H(5) = 19.01, p < 0.01; H(5) = 13.9, p = 0.16; \text{respectively})\) and their number in percent of total words increased markedly over time (and: \(1\% - 5.5\%, J = 213, p < 0.01;\text{ then: }0\% - 1.6\%, J = 196, p < 0.01)\). Post-hoc tests revealed a significant increase of und between span B2 \((\text{mdn} = 3.3\%)\) and C1 \((\text{mdn} = 4.7\%)\) \((U = 0, z = -2.3, p = 0.21)\), and no further significant increase within the last six years \((J = 29, p = 0.46)\). Separate analyses revealed, however, a significant increase of dann within the last six years \((J = 38, p = 0.04)\). The increase of und thus seems to be associated with the almanac style change, while this does not apply to the increase of dann. The use of connectives over time is depicted in Fig. 6.

Errors of sentence composition (sentence inversion or repetition of [parts of] sentences) occurred 59 times in total \((A = 0, B1 = 0, B2 = 1, C1 = 9, C2 = 17, C3 = 32)\). Among them, two syntactic errors occurred regularly: heim nach Hause [instead of either heim or nach Hause, both meaning home] and wir schauen uns den TV an [we look at the TV, instead of wir schauen TV, we watch TV], which HK produced from the summer of year 7 onwards. HK here returned to the standard construction of “gehen nach x” [going to x] and “wir sehen uns x an” [we look at x] and ignored the special construction rule in the context of “going home” and “watching TV”. The sum of these error types differed significantly between timespans \((H(5) = 20.68, p < 0.01)\).
Table 2
Overview of error types with examples and translations.

<table>
<thead>
<tr>
<th>Level/Type</th>
<th>Example — correct form</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface dysgraphia</td>
<td>Blanschbecken — Planschbecken</td>
<td>Paddling pool</td>
</tr>
<tr>
<td>Semantic paraphasia</td>
<td>Wir fuhren dann im Auto nach R., wo ich unten im Keller vom COOP mein Auto hinstelle. — in der Garage Waschsalon — Badezimmer</td>
<td>We drove then in the car to R., where I park the car in the cellar of COOP (supermarket). — garage Wash-saloon — Bathroom</td>
</tr>
<tr>
<td>Syntax Inversion</td>
<td>Am Nachmittag entschliesse ich mich zu Wandern gehen. — Wandern zu gehen</td>
<td>In the afternoon I decide to go for a walk.</td>
</tr>
<tr>
<td>Perseveration</td>
<td>Wegen der Kalte lesen wir unsere Zeitungen + essen danach Zmorge, zur Mittagszeit. Wegen der Kalte + keinem Sonnenlicht bleibt ich heute zu Hause, weil ich keine Wanderung mache. Die windige + feucht/kalte Luft zwingt mich zu Hause zu bleiben.</td>
<td>Due to the cold we read our newspapers + eat breakfast afterwards, at lunchtime. Due to the cold + no sunlight I stay at home today. because I don't go for a walk. The windy + humid/cold air forces me to stay at home.</td>
</tr>
<tr>
<td>&quot;Home&quot;-construction</td>
<td>Mit Tram 16.09 h fahre ich heim nach Hause. — heim/nach Hause</td>
<td>With tram 16.09 I go home to home. — I go home</td>
</tr>
<tr>
<td>&quot;TV&quot;-construction</td>
<td>Mittags blieb ich zu Hause, las viel, spielte Orgel + schaute TV an. — schaute TV</td>
<td>At noon I stayed home, read a lot, played the organ and looked at the TV. — watched TV</td>
</tr>
<tr>
<td>Word elision</td>
<td>Um 11 Uhr fahren wir [in/with] unserem Auto nach R.</td>
<td>At 11 o'clock we drive [in/with] our car to R.</td>
</tr>
<tr>
<td>Morphology</td>
<td>Viele Sonnen. — Viel Sonne</td>
<td>A lot of suns.</td>
</tr>
<tr>
<td>Case</td>
<td>Anschliessend trinken wir im Rest. unsere Kaffee.</td>
<td>Afterwards we had our coffee in the restaurant.</td>
</tr>
<tr>
<td>Gender</td>
<td>... wo ich auch meine Tee trinke. — meinen</td>
<td>... where I also had my tea.</td>
</tr>
<tr>
<td>Function word</td>
<td>Wir haben mit Hrn. K. einen nächsten Termin um 30. September abgemacht. — am</td>
<td>We've fixed our next appointment with Mr. K. at September 30th — on</td>
</tr>
</tbody>
</table>

p < 0.01) and there was a significant increase over time (J = 212, p < 0.01). Post-hoc tests on the individual error types revealed a significant increase of the wrong “TV-construction” between B2 and C1 (from 0 to 3; U = 2, z = –2.05, p = 0.04). Furthermore, the “home”-construction and perseverations increased within the last six years (home: J = 36.5, p = 0.04; perseveration: J = 38.5, p = 0.03). Word elisions occurred only 4 times (in years 6 and 7) and their distribution did not differ between timespans (H(5) = 7.82, p = 0.17). The number of these diverse syntactic errors is depicted in Fig. 7.

3.5. Morphology

On the level of morphology we observed abnormalities of HK’s use of inflectional morphemes to indicate number, gender, or case, and unusual use of function words such as prepositions and articles. In total we found 54 mistakes of these kinds (A = 0, B1 = 1, B2 = 1, C1 = 11, C2 = 19, C3 = 22). Separate analyses revealed only significant differences between timespans for the percentage of errors in function words (H(5) = 17.06, p = 0.03) which significantly increased over the 12 years (J = 200, p < 0.01) and first occurred in year 5. The increase was also significant within the last six years (J = 38, p = 0.04). There was neither an effect of timespan in the percentage of errors in number morphology (H(5) = 8.88, p = 0.14) nor the percentage of errors in case or gender morphology (H(5) = 4.56, p = 0.47). The percentage of morphological errors is depicted in Fig. 8.

4. Discussion

We had the privilege of analyzing the diaries of a person diagnosed with svPPA 12 years after he had started writing them. This presented a unique opportunity to gain insights about written language production in the preclinical and early-to-mid
phases of svPPA. Qualitative and quantitative analyses were based on transcripts of two weeks' entries per year and encompassed the overall structure, vocabulary, surface dysgraphia and semantic paraphasia, syntax, and morphology. Our two main aims were to identify the earliest indicators of cognitive change in HK’s diary entries, and to track the important changes in the diary entries over time.

The first errors were detected in years 3 and 4 (errors in semantics and morphology). However, considering their small number, the lack of continuation in the following years, and the fact that healthy subjects also sometimes produce errors in writing or speaking (e.g. Sajjadi et al., 2012), these early errors cannot necessarily be interpreted as being the first signs of cognitive decline. Moreover, the fact that HK changed his almanac style twice, which for instance affected the entry length and the variety of vocabulary, further complicates the determination of an onset of cognitive decline. In our view, the first signs that do not represent single errors within the normal range, and that cannot be attributed to the different almanac styles, were observed between year 5 and 6, thus seven years before the clinical diagnosis of svPPA. These changes concerned a remarkable drop in the variety of the vocabulary, together with an increased use of higher frequency words.

Regarding the overall structure, we found an increase of incomplete or missing entries, an increase of recurring elements, a trivialization of the topics with day-to-day routines reported in a more detailed fashion, and a gradual increase of entry length over the first 11 years with a remarkable drop during the last months. The incomplete or missing entries might correspond to problems with the writing process itself and to writer’s block reported for professional writers with AD (Garrard et al., 2005; van Velzen & Garrard, 2008). The increase of recurring elements (e.g. noting the daily weather condition) considered as a “barrier ritual” (Surd-Büchele, 2011) seems to gain importance over time, and can be seen as an early strategy to deal with the increasingly demanding writing process. Another possible interpretation is that these elements are signs of increasing obsessiveness (Snowden et al., 2001). The change of reported topics towards more commonplace matters might be associated with the increase of mundane activities in his life, as is regarded typical for the majority of individuals with dementia (Shany-Ur & Rankin, 2011). Furthermore, it parallels the finding of increasingly shallow and banal stories with slightly chaotic and illogically composed passages produced by the writers with AD (Heumakers, 1998; Kakutani, 1996), even though one has to bear in mind the difference between composing a fictional story and the naturalistic setting of writing a diary about one’s daily life. The increase in entry length mirrors the tendency of svPPA patients to talk fluently but with less semantic content in the beginning (Warren et al., 2013). It is conceivable that this increase might also be related to the impairments in executive functioning that became apparent during the neuropsychological examination. The decrease of entry length during the last year is in line with the more pronounced difficulties in language processing overall.

Regarding vocabulary we found a decrease in variety paired with an increase of high frequency word use that fits well with the findings in spoken language production of individuals with svPPA (Bird et al., 2000; Bozeat et al., 2003; Laisney et al., 2011; Sajjadi et al., 2012). The decrease in variety can be partially explained by the nature of the diary entries, which, in
Association with the change in almanac style, became increasingly prosaic. The more prose-like writing style automatically triggered a higher number of repetitions in the vocabulary, e.g. caused by pronouns. However, statistical analyses showed also a significant decrease in variety within the last six years, where the same almanac style was used. Furthermore, an increase in high frequency word use was also observed in written text production in AD (Garrard et al., 2005; van Velzen & Garrard, 2008).

We expected to find errors related to semantic memory impairment such as surface dysgraphia and semantic paragraphia early and at an increasing rate over time. While we found few instances of surface dysgraphia, we did find semantic paragraphias in the pattern expected. However, semantic paragraphias occurred only in the last third of the whole timespan and therefore somewhat later than other linguistic abnormalities. Qualitatively, an increasing semantic distance between target and paragraphia was observed over time. We speculate that possible uncertainties or mild word finding difficulties may be more easily compensated in diary writing than in speech, as diary writing allows an individual to freely choose the rather high-frequency content without time-pressure. This might explain why semantic paragraphias only occur in later stages.

Another observation that is in our view related to semantic memory impairment is the increasing use of quotation marks for common words. While HK uses quotation marks adequately in the first years, he uses them excessively later. In our view, the most plausible explanation for this finding lies in the feeling of oddness that HK presumably developed for more and more words. We may speculate that his decreasing knowledge about word meanings, which is characteristic for svPPA, left HK with a strange feeling about the remaining word hulls. Therefore, he marked them by putting them between quotation marks, as if they were foreign, exotic words he was no longer familiar with or sure about.

The near absence of surface dysgraphia might be associated with the grapheme-phoneme and phoneme-grapheme correspondence of the German language. As Benedet et al. (2006) point out, the occurrence of surface dyslexia depends on the grapheme-phoneme correspondence in a given language. The fact that German is a so-called transparent language (i.e. one grapheme is usually associated with one phoneme) reduces the probability to observe such errors. We reason that the same might be applicable to the phoneme-grapheme correspondence in writing and therefore errors of surface dysgraphia are less probable in German than in English.

Concerning syntactic composition, we found, as expected, an increasing rate of simple connectives, such as “and” or “then”, in combination with a decrease of syntactically (and mentally) more complex connectives such as “despite” or “hence”. While the increase of “and” was partially associated with the almanac style change, this was not observed for “then”. We thus conclude that the cognitive decline lead to the use of simpler structures over time. Furthermore, errors at the syntactic level such as sentence inversions and perseverations were detected, as well as an increasing use of two remarkable incorrect constructions, the “home”-construction and the “TV”-construction. HK ignored the necessary but atypical rules for
These constructions, and we therefore interpret these two construction errors as regularization errors. Furthermore, we found evidence of changes at the morphological level which appeared mainly in the form of incorrect function words.

These changes at the syntactic and morphological levels are in line with recent findings in spoken language production in svPPA (e.g. Kave et al., 2007; Meteyard & Patterson, 2009; Wilson et al., 2010, 2014). In agreement with these authors, we interpret the errors as being tightly linked to increasing semantic difficulty. Their early occurrence speaks against the assumption that they are caused by the progression of atrophy to other, language-critical, cortical areas (Bright et al., 2008). However, as the writer is not under time pressure, his product, the text, is not as elusive as spoken language. This would give him enough time to think about the proper word or phrase, or to correct an incorrect word or phrase. The fact that we did not find many corrections in the diary entries also hints at a decreased monitoring ability. This could be further investigated in future studies.

In conclusion, we found changes in the number of entries made, their length, and the topics reported. Over time, HK’s vocabulary became less varied and words of higher frequency were used more often. Semantic paragraphias, as well as paragrammatic errors occurred. The sentence structure became progressively simpler and incorrect inflectional morphemes and function words were used. Quotation marks were used excessively, which we tentatively suggest may have been due to a feeling of increasing oddness about the world. Our analysis of the diary entries revealed the onset of clear abnormalities in the texts about seven years before the clinical diagnosis took place and initially showed a higher degree of syntactic and morphological than semantic errors. In this study, the characteristic semantic problems reported in svPPA patients emerged later in the course of disease and were less dominant in HK’s written language than were other features.

The analysis of text production in this single case reveals that a comparison of written texts from different time points might assist in the early detection of a decline in language functions and differential diagnosis. Written texts of similar form and content (e.g. e-mails, letters, diary entries) might be available for many patients and could thus provide a window into the premorbid status of a person. Word frequency and type-token ratio were the measures most sensitive to early changes in this study. Furthermore, the occurrence of regularization errors on different levels (syntax, spelling) might be an early indicator of semantic deficits. Even though the analysis of (hand)written texts might be too time consuming for a standard clinical diagnostic setting, it might be used in patients with higher levels of education and with diffuse symptoms.

**Acknowledgments**

We owe HK and his partner special thanks for their cooperation and their agreement to the study at hand.

**References**


