Open Journal of Modern Linguistics, 2017, 7, 8-32
http://www.scirp.org/journal/ojml
ISSN Online: 2164-2834
ISSN Print: 2164-2818

The Impact of Depression and Apathy on Sensory Language

David Kernot[1,2*], Terry Bossomaier[3], Roger Bradbury[1]

[1]National Security College, The Australian National University, Canberra, Australia
[2]Defence Science Technology Group, Edinburgh, Australia
[3]The Centre for Research in Complex Systems, Charles Sturt University, Bathurst, Australia
Email: *u5604766@anu.edu.au

How to cite this paper: Kernot, D., Bossomaier, T., & Bradbury, R. (2017). The Impact of Depression and Apathy on Sensory Language. *Open Journal of Modern Linguistics, 7,* 8-32.
https://doi.org/10.4236/ojml.2017.71002

Received: December 19, 2016
Accepted: February 3, 2017
Published: February 6, 2017

Abstract

The debate that language strongly influences thought is equally met by those who suggest language does not influence thought. While historically, the ability to communicate with words was believed to be intimately tied to an ability to form thoughts, we would argue that thought and language are linked together through our sensory and motor systems and severely impacted by depression and apathy. We test this by conducting parts of speech analysis from the comparative longitudinal studies of two highly creative and prolific writers, where one is diagnosed with Alzheimer's disease (AD), and the other lives a long and healthy life. We calculate function and contents word ratios, measure lexical repetition, and the use of sensory-based words in textual language to test for depression and apathy in AD which is supported by Mann-Whitney U-Testing and Principal Component Analysis. Our results support the hypothesis that thought and language are impacted by depression and apathy and revealed in a person's writing style 12 years before a formal diagnosis of Alzheimer's disease presents. We find higher lexical repetition in language 12 years prior to one author being diagnosed with AD while not apparent in the other. We identify low olfactory word use and also find that an increased use of sensory-based adjectives might be a sign of the early onset of Alzheimer's disease.

Keywords

Depression, Apathy, Thought, Sensory Language, Word Richness, Alzheimer's Disease

1. Introduction

The debate over the notion that language strongly influences thought is met equally by those who argue that language does not influence it, but historically, language was thought to be tied to an ability to form thoughts (Wicklund, Johnson, & Weintraub, 2004). Some believe that language and thought are combined to modify language and thought further (Ammar & Gohar Ayaz, 2016). Others suggest human language is an instrument of thought and communicates the attributes of human culture (Lieberman, 2016), or that language allows us to share the knowledge and experiences of others to increase our mental resources (Corballis, 2016). Through this embodied cognition, our concepts are grounded in our sensory and motor systems to develop new abstract representations (Jamrozik et al., 2016). We would argue that the way we think comes through clearly in the multimodal sensory elements of our language, and that these aspects of language (such as through sound and visual body language cues) impact on thought, but that disease impacts these.

While thought strongly influences language, depression and apathy severely impact thinking, as seen in dementia. There is a close link with depression in dementia, and apathy and depression are the most frequent neuropsychiatric symptoms in one type of dementia, Alzheimer's disease (AD) (Robert, Bremond, & David, 2016). In a different kind of language-based dementia, known as Primary Progressive Aphasia (PPA), patients who could not find the right words to express their thoughts, could still demonstrate they could think clearly (Fedorenko & Varley, 2016; Wicklund, Johnson, & Weintraub's, 2004). While language in PPA is a prominent dysfunction for the first two years of the disease, Alzheimer's disease comes to medical attention because of forgetfulness, usually accompanied by apathy, but not from language dysfunction (Mesulam, 2003). Personal identity persists far into the end stage of the disease (Sabat, & Harré, 1992), while apathy is characterized by reduced motivation, social disinterest, and emotional blunting in the absence of mood-related changes (Chau et al., 2016).

If writing is the ability to think and put language on paper or some other visual mediums, then the impact of depression and apathy on thought and language might be measurable, but this concept is not new. The idea that cognitive decline in Alzheimer's disease is visible in writing appeared in Snowdon et al.'s (1996) findings of a longitudinal study of 678 Catholic sisters. Known as the Nun Study, researchers were able to correlate post-mortem markers for AD in the sister's brains to the density of ideas (from Kintsch & Keenan, 1973 and Turner & Greene, 1977) expressed in sentences using Parts of Speech (POS) Tag analysis. Idea density uses elements of language, verbs, adjectives, adverbs, prepositions, and conjunctions divided by the number of words to create a measure of cognitive ability (Brown et al., 2008). Garrard et al. (2005) were also instrumental in highlighting Alzheimer's disease through changes in writing and used a different approach which included some other elements of language (nouns, verbs, adverbs and adjectives and function words, e.g., conjunctions, and pronouns) to create word lists. As Arefin et al. (2014) and Ferguson et al. (2014) point out: the study of the subtle language changes over the lifespan of well-known writers

(Lancashire, 2010), including Iris Murdoch and Agatha Christie (e.g. Garrard et al, 2005; Van Velzen & Garrard, 2008; Lancashire & Hirst, 2009; Le, 2010; Le et al., 2011) and political figures (Garrard, 2009) has highlighted that Alzheimer's disease may be apparent years or even decades before anyone becomes aware of any symptoms of cognitive deterioration. A recent study suggests that Alzheimer's disease can be seen in people's writing 10 - 12 years before the disease is diagnosed (Rajan et al., 2015). To test the hypothesis that thought and language are impacted by depression and apathy and revealed in a person's writing style 12 years before a formal diagnosis of Alzheimer's disease presents, we draw on earlier studies of AD.

We use the novels of Iris Murdoch and P.D. James, however, we use a larger, more complete set than previously used by Garrard et al. (2005) and Le et al. (2011). We use broader Parts of Speech analysis techniques, and we also use a new analytical technique from sensory adjectives, to determine what can be seen in language from the impact of depression and apathy in the early onset of Alzheimer's disease.

2. Methodology

In this section, we describe the two underlying concepts of the method, Richness and Sensory Adjectives, and what is done with the data in the pre-processing stage to visualise markers for depression, apathy, and Alzheimer's disease in language.

2.1. Richness (R)

Richness Equation (1) is based on Menhinick's (1964) species diversity equation, and for two documents of the same length, the one with more different (unique) words - a larger vocabulary - has greater richness. The Richness score can be determined by:

$$\text{Richness}(R) = \frac{w}{N} \quad (1)$$

where w = number of unique words in the document, and N = the total document word count.

There are theoretical limitations to this equation, and the size of documents must be carefully controlled to avoid artifacts. Eventually, the value will reach an asymptote as no new words are found. Near that point, the larger the document size, the smaller the Richness score will be (0 as $N \to \infty$).

The type—token ratio (TTR) can also be considered a variant of the species diversity equation, and is text size dependent, however, it is a popular metric, and we mitigate any size impacts and avoid issues near the asymptote of word counts by keeping all the samples of each novel of equal size and at 4000 words, well below 10,000 where other techniques perform better; further, we use Richness as part of a larger multivariate technique (Juola & Mikros, 2016; Kimura & Tanaka-Ishii, 2014; Kubát & Milička, 2013; Tanaka-Ishii & Aihara, 2015; Van

Gijsel, Speelman, & Geeraerts, 2005; Vermeer, 2000).

2.2. Sensory Adjectives (S)

While apathy is characterised by reduced motivation, social disinterest, and emotional blunting in the absence of mood-related changes, it has been associated with low norepinephrine levels in the brain (Chau et al., 2016). In the following paragraphs, we describe the link between apathy and depression in people to different levels of norepinephrine in the brain, and how it might be apparent in sensory processing and impact the sensory language of Adjectives.

Many mental disorders have also been associated with alterations of neurotransmitters in the brain (Heilman, Nadeau, & Beversdorf, 2003; Nonen et al., 2016; Sun, Hunt, & Sah, 2015; Szot, 2016), and the neurotransmitter, norepinephrine, has been seen to be lacking in depressed suicide victims (Khan et al., 2016; Klimek et al., 1997; Ramirez, 2016). Norepinephrine levels have also been linked to studies on creative people, where a reduction in their aural sensory processing, known as sensory gating is tied to the neurotransmitter, while creative achievers have shown "leaky" sensory gating because they simultaneously focus on a large range of stimuli (Zabelina et al., 2015). They have low levels of norepinephrine which increase the size and distribution of the brain's concept representations. Their ability to modulate the frontal lobe-locus coeruleus system and reduce norepinephrine levels leads to the discovery of novel orderly relationships, or creative innovation (Heilman, Nadeau, & Beversdorf, 2003). Other creative people, divergent thinkers, on the other hand, reduce sensory gating, which is also a marker of psychosis, including schizophrenia (Zabelina et al., 2015). They have high levels of norepinephrine that restricts their concept representations (Heilman, Nadeau, & Beversdorf, 2003), and therefore their sensory processing narrows to focus tightly on the task at hand. Here, thought influences ideas and modifies aspects of the sensory cortex that feeds language.

The brain processes many Sensory (S) words as sight/feel and smell/taste word categories (Lynott, & Connell, 2009). Adjectives are used over verbs or nouns because they appear more frequently, and their context is not necessary. These 774 sensory-based adjectives (van Dantzig et al., 2011) are recorded in two different contexts to assess the dominant visual (V), auditory (A), haptic (H), olfactory (O), or gustatory (G) modality. These sensory words are allocated a modality exclusivity score that reflects the brain's Representational System (Fernandino et al., 2015), and can be used to capture the sensory gating biomarker characteristics of a person. There are five sensory categories, one each for V, A, H, O, G. If we let the number of words in each sensory category, i, be φ_i and ϑ_i is the exclusivity value (see van Dantzig et al., 2011 for values) and N_k be the number of discrete word score, then the weight, or exclusivity score for each category then the Sensory score Equation (2) can be determined by:

$$\text{Sensory adjectives}\left(S_{k1-5}\right)=\sum^{N_k}\frac{\varphi_i\vartheta_i}{D} \quad (2)$$

where $\sum^{N_k}$ = 774, and D is the number of words in the document, and k = 1 - 5.

The overall Sensory adjectives ($\mathcal{S}$) score is calculated by summing the five sensory categories.

2.3. Preparing the Text

A collective corpus of 180,000 words contains a 104,000-word sample from 26 Iris Murdoch novels and a 76,000-word sample from 19 P.D. James novels (see Table 1 and Table 2). The 4,000-word novel sample is from the first 3,000 words and the last 1,000 words. Generally, in a novel, this is where characters, rich in setting and plot, are introduced and at the end, a conclusion of the general novel 'problem' has been resolved and summarised. We process the files with the Stanford Parts Of Speech (POS) Tagger (Toutanova & Manning, 2000) and remove all stop words (punctuation) before achieving each 4000-word sample per

Table 1. Iris Murdoch's novels by year published.

ID	Iris Murdoch Novels	Published
B1	Under the Net	1954
B2	The Flight from the Enchanter	1956
B3	The Sandcastle	1957
B4	The Bell	1958
B5	A Severed Head	1961
B6	An Unofficial Rose	1962
B7	The Unicorn	1963
B8	The Italian Girl	1964
B9	The Red and the Green	1965
B10	The Time of the Angels	1966
B11	The Nice and the Good	1968
B12	Bruno's Dream	1969
B13	A Fairly Honourable Defeat	1970
B14	An Accidental Man	1971
B15	The Black Prince	1973
B16	The Sacred and Profane Love Machine	1974
B17	A Word Child	1975
B18	Henry and Cato	1976
B19	The Sea, the Sea	1978
B20	Nuns and Soldiers	1980
B21	The Philosopher's Pupil	1983
B22	The Good Apprentice	1985
B23	The Book and the Brotherhood	1987
B24	The Message to the Planet	1989
B25	The Green Knight	1993
B26	Jackson's Dilemma	1995

Table 2. P.D. James' novels by year published.

ID	P.D. James Novels	Published
B1	Cover Her Face	1962
B2	A Mind to Murder	1963
B3	Unnatural Causes	1967
B4	Shroud for a Nightingale	1971
B5	An Unsuitable Job for a Woman	1972
B6	The Black Tower	1975
B7	Death of an Expert Witness	1977
B8	Innocent Blood	1980
B9	The Skull Beneath the Skin	1982
B10	A Taste for Death	1986
B11	Devices and Desires	1989
B12	The Children of Men	1992
B13	Original Sin	1994
B14	A Certain Justice	1997
B15	Death in Holy Orders	2001
B16	The Murder Room	2003
B17	The Lighthouse	2005
B18	The Private Patient	2008
B19	Death Comes to Pemberley	2011

novel, and then we aggregate the works by individual word frequency. We also aggregate the data from the 45 POS types into 12 more general POS types, representing higher classes of Nouns, Verbs, Adverbs, Adjectives, Modal Verbs, Conjunctives, Prepositions, Determiners, Pronouns, Existential There, Articles, and Other categories, comprising Cardinal Numbers, Interjections, and Foreign words.

3. Results

In this section, we begin by testing for markers within the Richness (R) of language that can highlight Alzheimer's disease (AD), and we support this with Mann-Whitney U-Testing. We use Parts of Speech (POS) analysis to group the data by Content and Function Words and use ratios to support the presence of AD in the data further, and we back up this claim with Mann-Whitney U-Testing. We test for markers within the Sensory (S) aspects of language to see if the variables can identify additional markers for AD, and support these results with Mann-Whitney U-Testing and Principal Component Analysis.

3.1. Testing for Alzheimer's Disease Markers in Richness

Alzheimer's disease is apparent through lexical repetition, marked by smaller, higher frequency vocabulary (Garrard et al., 2005), and can be seen in people's writing 10 - 12 years before the disease is diagnosed (Rajan et al., 2015). Iris Murdoch was diagnosed with Alzheimer's disease in 1997 (aged 78), and she died two years later in 1999, just four months before her eightieth birthday, and a post-mortem confirmed the Alzheimer's disease (Garrard et al., 2005). P.D. James was not diagnosed with Alzheimer's disease or dementia and died in 2014 (aged 94).

From Richness, the total mean of Iris Murdoch's writing is 29.5, while P.D. James' is slightly higher at 31.9, suggesting there is less lexical repletion than Iris Murdoch's. Separating the last 12 years of works highlights that this period in Iris Murdoch's case is slightly lower (29.48 versus 29.52), while P.D. James' is slightly higher (32 versus 31.86) and these do not appear to be significantly different.

To test this, we conduct a Mann-Whitney U-Test on the Richness scores. In this hypothesis testing for differences in a person's writing style, we use the Mann-Whitney U-Test because it is a non-parametric independent groups test. In this case, the total sample size for Iris Murdoch is 26, with the group sizes of 21 and 5, and for P.D. James, the total sample size is 19 with the group sizes being 15 and 4. This test is ideal for unequal group sizes that are small, have dissimilar variances, and a distribution that is not normal (Burns & Burns, 2012).

We find significant differences in the lexical repetition of Iris Murdoch's works (see **Table S1** and **Table S2** in Supplementary Data) during the writing period 1984-1996, 12 years before her diagnosis when compared to the earlier period of writing 1954-1983 ($U = 12.5$, $p = 0.009$). There are no significant differences in the lexical repetition of P.D. James works (see **Table S3** and **Table S4** in Supplementary Data) during the writing period 1962-2001, 12-years before her death when compared to the period 2002-2011 ($U = 29.0$, $p = 0.920$).

3.2. Content and Function Word Analysis

It is generally understood that while the ratios of different word types is relatively uniform across age, sex, and level of education in normal speakers, that there is a lower use of Function Words over Content Words in people with Dementia and different aphasia types that impact speech and language (Bird et al., 2000). Function Words contain little meaning and tend to hold sentence structure together. They are word types such as pronouns, articles, prepositions, and conjunctions. Content Words, on the other hand, tend to describe the message of a sentence through verbs, nouns, adverbs, and adjectives.

Given these differences in Function and Content Words (Bird et al., 2000), we would expect that as a person develops Dementia, there would be an increase in the use of Content Words, and in their Content to Function Word ratios. We test the use of Content Words by aggregating the 12 tagged Parts of Speech groups into Content Words and Function words (see **Table S5** for Iris Murdoch and **Table S6** for P.D. James in Supplementary Data) and separate the last 12 years of works to compare the later writing to the earlier period. In Iris Murdoch, there is a mean increase in Content Words use of 3.16%, or 75.4 (2380.0 - 2456.3), and a mean decrease in Function Words use of 4.65%, or 75.4 (1619.1 - 1543.67) in the later 12 Years period (see **Figure S1** and **Figure S2** in Supplementary Data). In contrast, in P.D. James', there is a mean increase in Content Words use of 1.77%, or 42.28 (2344.46 - 2386.75), and a mean decrease in Function Words use of 2.55%, or 42.29 (1655.53 - 1613.25) in the later 12 Years period (see **Figure S3** and **Figure S4** in Supplementary Data). Iris Murdoch's use of Content Words was approximately 44%, or 33.11 (42.29 - 75.4) larger in the period 12 years before her diagnosis of Alzheimer's disease.

Having seen differences in Content Words, we test their Content to Function Word ratios for signs of Dementia. Drawing on a technique from Garrard et al. (2005) and used in Le et al. (2011), we compare the first work of Iris Murdoch to the period 12 years before her diagnosis of Alzheimer's disease. We plot the Function to Content Word ratio (from **Table S5** in Supplementary Data) and do the same for P.D. James (from **Table S6** in Supplementary Data).

These results in Iris Murdoch's writing (**Figure 1**), reflect the findings of Garrard et al.'s (2005) observations that by her final book Jackson's Dilemma (B26), she was suffering from cognitive decline caused by Alzheimer's disease. Here, we extend the data from three of her novels to seven and find the Function Word to Content Word ratios are all lower for the six works 12 years before the diagnosis of AD. There is a steady decline in Iris Murdoch's work until the fourth work (B24) where the ratio is approximately level.

As we can see in P.D. James' writing (**Figure 2**), the Function Word to Content Word ratios are different and appear as a sawtooth pattern, and while two are much higher (B16 and B18), the other two are lower (B17 and B19). This suggests that there is neither a steady incline nor decline in the 12 years before P.D. James' death.

These results are supported by the Mann-Whitney U-Tests, which show that there are significant differences in the use of Function Words and Content

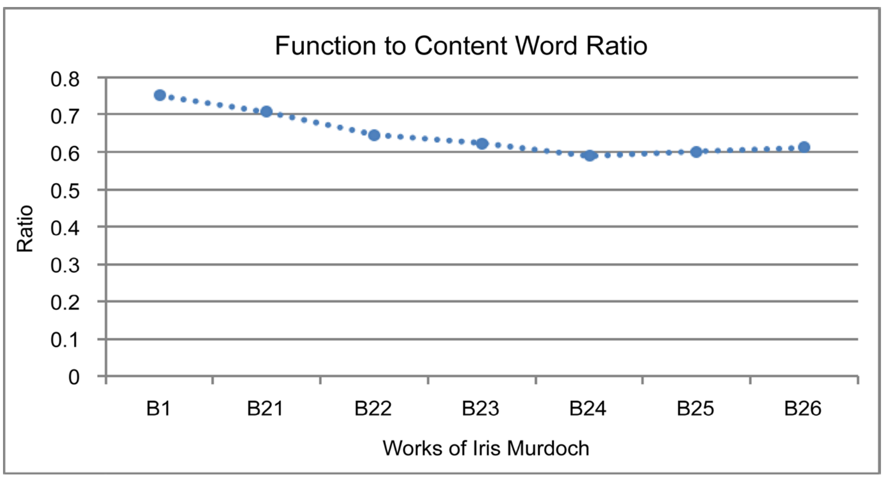

Figure 1. Iris Murdoch content to function word ratio comparison of her first work to the six works 12 years before her diagnosis with AD. All variables are lower than her early work.

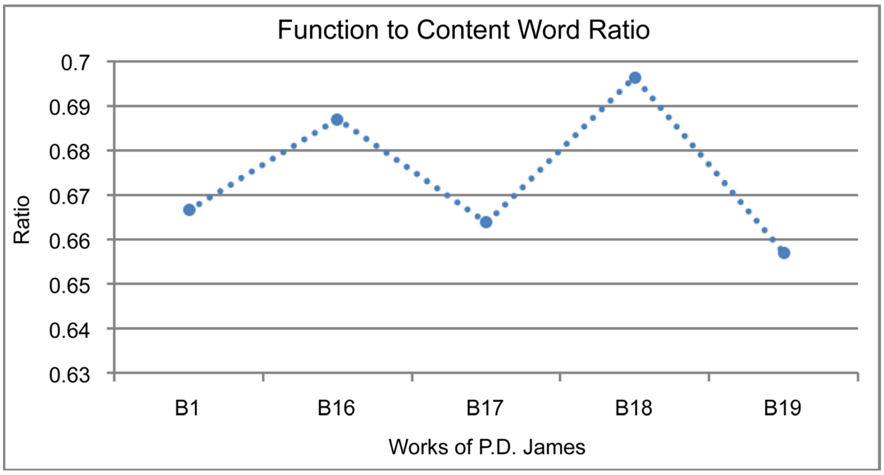

Figure 2. P.D. James content to function word ratio comparison of her first work to the four works 12 years before her death. The increasing sawtooth pattern is neither higher or lower overall than her early work, and the results are very different to the Iris Murdoch results.

Words by Iris Murdoch (U = 24, $p = 0.028$) 12 years before her diagnosis of AD (see Table S7 and Table S8 in Supplementary Data). As we would expect in a person without dementia or AD, there is no significant difference for P.D. James (U = 17, $p = 0.194$) 12 years before her death (see Table S9 and Table S10 in Supplementary Data).

3.3. Testing for Sensory Alzheimer's Disease Markers

Many sensory words are processed by the brain as sight/feel and smell/taste word categories (Lynott & Connell, 2009), and we use a group of 387 sensory adjectives and allocate them a modality exclusivity score that reflects the brain's Representational System (van Dantzig et al., 2011). These sensory words can be used to capture the sensory gating biomarker characteristics of a person (Fernandino et al., 2015) to create a unique signature of their inner self.

Examining the Sensory variable, we find that there is no overlap in the Standard Errors, and the mean is higher in the last 12 years (see Figure 3). P.D. James' Sensory variable 12 years before her death show a lower sensory score, and the Standard Errors also do not overlap (see Figure 4). Overall, Iris Murdoch's Sensory mean is slightly higher than P.D. James (0.022 versus 0.018).

These results are supported by the Mann-Whitney U-Tests, which show that there is significant differences in the use of Sensory Words by Iris Murdoch (U = 18, $p = 0.011$) 12 years before her diagnosis of AD. This is also true of P.D. James (U = 7, $p = 0.021$) 12 years before her death (refer detailed results Table S11-S14 in Supplementary Data), but in P.D. James' case, her use of Sensory Words is lower, not higher.

Overall, these observations are further supported in the underlying five sensory variables for Iris Murdoch. Except for the Olfactory element, all of the other means of the sensory elements (V, A, H, & G) were higher in the period 12 years before her diagnosis of AD. However, there were overlaps in the Haptic and

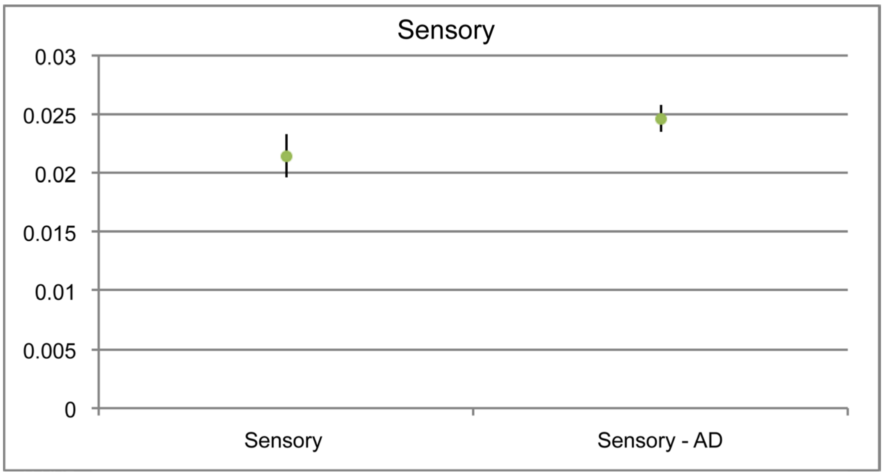

Figure 3. Iris Murdoch sensory mean with standard error bars highlighting a higher use of sensory adjectives during the period 12 years before her diagnosis with AD. Note the error variance is smaller in the latter period, and there are no overlaps in the standard error bars.

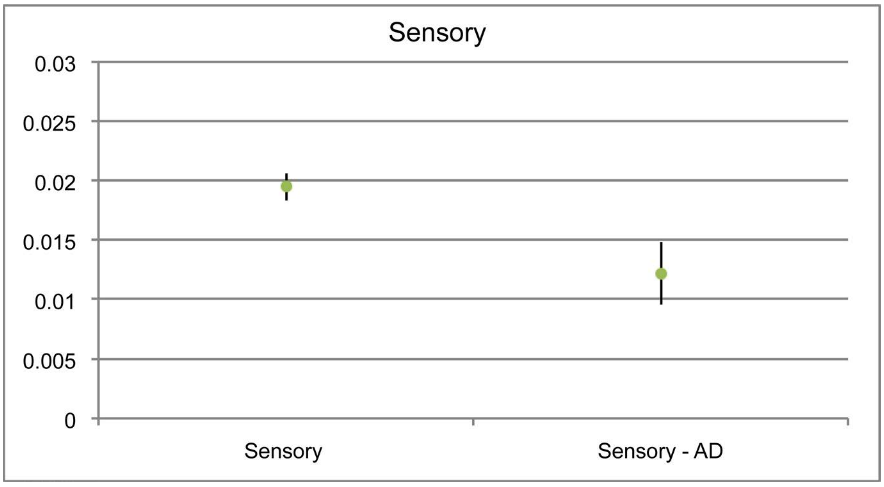

Figure 4. P.D. James sensory mean with standard error bars highlighting a lower use of sensory adjectives during the period 12 years before her death. Note that these results are the reverse of the Iris Murdoch results in that the error variance is larger in the latter period with a lower mean. Again, there are no overlaps in the standard error bars.

Olfactory Standard Errors. In contrast, all of P.D. James Visual, Auditory, Haptic, Olfactory, and Gustatory (VAHOG) elements were lower in the 12 years before her death, and there were no overlaps in the Standard Errors (see Figure S5-S14 in Supplementary Data).

To support the sensory observations, we conduct Principal Component Analysis on the VAHOG elements. We measure the total percentage of variance these elements contribute to the overall extracted components and then iteratively remove and replace them to determine the impact that each one has on the total component's variation. We find that except for the Olfactory element, the results are relatively similar across V, A, H, & G for both Iris Murdoch and P.D. James (see Table 3). In the case of the Olfactory element, we see a large, negative impact

Table 3. A comparison of the sensory contribution to PCA variance. Here it is clear that the impact on Iris Murdoch's Olfactory variable is significantly larger than any other result for either author.

Impact on Variance (%)		
Sensory Element	Iris Murdoch	P.D. James
Visual	−0.86	−1.13
Auditory	2.93	5.1
Haptic	−2.05	−1.3
Olfactory	−16.19	3.87
Gustatory	6.74	0.36

from its removal, highlighting the significant contribution, and therefore difference, which this element plays within the Iris Murdoch data. Alzheimer's disease impacts normal olfactory function with suggestions that olfactory loss may be a biomarker for AD and cognitive decline (Wesson et al., 2010; Woodward et al., 2015).

4. Discussion

Analysis of the Richness of each of the writer's novels using a Mann-Whitney U-Test highlighted significant differences in the lexical repetition of Iris Murdoch's works in the last 12 years of her writing (1984-1996). However, there were no significant differences in the works of P.D. James 12 years before her death. When using Parts of Speech analysis to group each work into Content and Function words, there was an increase in the use of Content Words in the later 12 Years of their writing. In Iris Murdoch's case, they were approximately 44% larger than P.D. James'. A decrease in Content to Function Word ratios as an indication of Dementia was observed in the overall works of Iris Murdoch 12 years before the diagnosis of AD, with four of the six works declining before the ratios levelled. A Mann-Whitney U-Test supported the significant differences in the later period of her writing. No such decrease or significant difference was observed in the works of P.D James, and the rising and falling sawtooth variation showed a pattern of neither steady incline nor decline in the 12 years before her death.

While there are no prior documented links to the Sensory variable and Dementia or AD, we test this and the five Sensory elements (VAHOG) for indications. We found that overall, Iris Murdoch's Sensory mean in the period 12 Years before her diagnosis of AD is higher, while P.D. James is lower than their earlier work, and these two groups are significantly different for both writers. Comparisons of the underlying five sensory variables for Iris Murdoch and P.D. James (see Table 4) highlight the means of the sensory elements (V, A, H, & G) were higher in the period 12 years before Iris Murdoch's diagnosis of AD. The Olfactory element was the exception. However, all of P.D. James Visual, Auditory, Haptic, Olfactory, and Gustatory (VAHOG) elements were lower in the 12

Table 4. Summary of sensory means of Iris Murdoch and P.D. James showing the overall higher sensory component in Iris Murdoch's results during the 12 year period before her diagnosis of AD, where P.D. James' results were all lower. Note that there was overlap in Murdoch's Haptic and Olfactory scores and that the Olfactory value was equal, or only slightly higher than the earlier period.

Variable	SE Overlap?		12 Years	
	Iris Murdoch	P.D. James	Iris Murdoch	P.D. James
Sensory	No	No	Higher	Lower
Visual Sensory	No	No	Higher	Lower
Auditory Sensory	No	No	Higher	Lower
Haptic Sensory	**Yes**	No	Higher	Lower
Olfactory Sensory	**Yes**	No	**Almost Equal**	Lower
Gustatory Sensory	No	No	Higher	Lower

years before her death. These two points support the overall Sensory result. While there were overlaps in Iris Murdoch's Haptic and Olfactory Standard Errors, Principal Component Analysis only supports the exceptionally low Olfactory element that is quite different in Iris Murdoch's writing (see Table 3).

Both Richness and POS analysis supported by Mann-Whitney U-Test highlight the evidence of dementia and AD in Iris Murdoch's writing in the latter 12 years before her diagnosis. These can be seen in Richness, a higher level use of Content Words, and a lower and lower Content to Function Word ratios. We also find that a higher Sensory mean might suggest the presence of AD in the period 12 Years before Iris Murdoch's diagnosis. While this is supported by the analysis of the different means in the last 12 Years of writing and the earlier works, and the comparative differences in Principal Component Analysis variances, the exceptionally low Olfactory element is quite different in Iris Murdoch's writing.

Murdoch's depression and apathy have been well documented through her prolific habit of writing about herself throughout her life (Dooley & Nerlich, 2014; Martin & Rowe, 2010; Murdoch, 2016; Wilson, 2004). Her sad decline into Alzheimer's disease has also been recorded by her husband (Bayley, 1998; 1999). We have stated earlier that there is a strong link between depression and apathy in dementia and particularly AD. It is known that while a depressed mood and apathy alter brain function in the prefrontal limbic network, and that it overlaps regions dealing with olfaction, such that depression can reduce olfactory ability (Croy et al., 2014).

A limitation to this longitudinal study is that it is the writing of only two authors and is not sufficient enough to suggest that AD, or indeed depression or apathy can be determined from the sensory writing of individuals. However, in this new approach to identifying the style of a person's writing using sensory adjectives, there were clear differences between both author's works in their last 12 years that warrants further study of other known authors who developed dementia.

5. Conclusion

In a study of two highly creative and prolific authors, we have been able to draw on a more complete set of novels than that used by Garrard et al. (2005) and Le et al. (2011) to characterise a person's use of language through writing. In doing so, we have applied both known techniques to identify linguistic markers for Alzheimer's disease, depression and apathy (through lexical repetition and function to content word ratios) and test a new technique based on sensory adjectives. Our results support the hypothesis that thought and language are impacted by depression and apathy and revealed in a person's writing style 12 years before a formal diagnosis of Alzheimer's disease presents. Using Richness to measure lexical repetition (a form of the type-token ratio) the writing of Iris Murdoch is statistically significantly lower in the last 12 years of her novel writing. This result is also reflected in Iris Murdoch's use of Function Words and Content Words). In contrast, a healthy P.D. James' writing during the same period showed no decline in lexical repetition and function to content word ratios and was not different from her earlier writing. There were clear differences in their use of sensory adjectives, with higher Iris Murdoch's use and lower P.D. James' during their latter 12 years of writing, but in Iris Murdoch's case, her use of olfactory words, a biological sensory marker for Alzheimer's disease, depression, and apathy, was low. It is possible that olfactory sensory words in language could be used to help identify depression and apathy in people. We suggest that cognitive diseases such as dementia impact on thinking, as seen through depression and apathy and can influence language use.

Acknowledgements

This research is supported by the Defence Science Technology Group, the Australian Government's lead agency dedicated to providing science and technology support for the country's defence and security needs.

References

Ammar, D. Z. A., & Gohar Ayaz, A. (2016). Language versus Thought, and Theory of Formation of Meanings. *Global Journal of Human-Social Science Research, 15,* 77-80.

Arefin, A. S., Vimieiro, R., Riveros, C., Craig, H., & Moscato, P. (2014). An Information Theoretic Clustering Approach for Unveiling Authorship Affinities in Shakespearean Era Plays and Poems. *PLoS One, 9,* e111445. https://doi.org/10.1371/journal.pone.0111445

Bayley, J. (1998). *Elegy for Iris*. London: Macmillan.

Bayley, J. (1999). *Iris: A Memoir of Iris Murdoch*. New York: Gerald Duckworth & Co Ltd.

Bird, H., Ralph, M. A. L., Patterson, K., & Hodges, J. R. (2000). The Rise and Fall of Frequency and Imageability: Noun and Verb Production in Semantic Dementia. *Brain and Language, 73,* 17-49. https://doi.org/10.1006/brln.2000.2293

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic Measurement of Propositional Idea Density from Part-of-Speech Tagging. *Behavior Research Methods, 40,* 540-545.https://doi.org/10.3758/BRM.40.2.540

Burns, R. B., & Burns, R. A. (2012). *Business Research Methods and Statistics Using SPSS*. SAGE Publications.

Corballis, M. C. (2016). The Evolution of Language: Sharing Our Mental Lives. *Journal of Neurolinguistics,* In Press. https://doi.org/10.1016/j.jneuroling.2016.06.003

Croy, I., Symmank, A., Schellong, J., Hummel, C., Gerber, J., Joraschky, P., & Hummel, T. (2014). Olfaction as a Marker for Depression in Humans. *Journal of Affective Disorders, 160,* 80-86. https://doi.org/10.1016/j.jad.2013.12.026

Chau, S. A., Chung, J., Herrmann, N., Eizenman, M., & Lanctôt, K. L. (2016). Apathy and Attentional Biases in Alzheimer's Disease. *Journal of Alzheimer's Disease, 51,* 837-846. https://doi.org/10.3233/JAD-151026

Dooley, G., & Nerlich, G. (2014). *Never Mind about the Bourgeoisie: The Correspondence between Iris Murdoch and Brian Medlin 1976-1995*. Cambridge: Cambridge Scholars Publishing.

Fedorenko, E., & Varley, R. (2016). Language and Thought Are Not the Same Thing: Evidence from Neuroimaging and Neurological Patients. *Annals of the New York Academy of Sciences, 1369,* 132-153. https://doi.org/10.1111/nyas.13046

Fernandino, L., Binder, J. R., Desai, R. H., Pendl, S. L., Humphries, C. J., Gross, W. L., Seidenberg, M. S. et al. (2015). Concept Representation Reflects Multimodal Abstraction: A Framework for Embodied Semantics. *Cerebral Cortex, 26,* 2018-2034. https://doi.org/10.1093/cercor/bhv020

Ferguson, A., Spencer, E., Craig, H., & Colyvas, K. (2014). Propositional Idea Density in Women's Written Language over the Lifespan: Computerized Analysis. *Cortex, 55,* 107-121. https://doi.org/10.1016/j.cortex.2013.05.012

Garrard, P. (2009). Cognitive Archaeology: Uses, Methods, and Results. *Journal of Neurolinguistics, 22,* 250-265. https://doi.org/10.1016/j.jneuroling.2008.07.006

Garrard, P., Maloney, L. M., Hodges, J. R., & Patterson, K. (2005). The Effects of Very Early Alzheimer's Disease on the Characteristics of Writing by A Renowned Author. *Brain, 128,* 250-260. https://doi.org/10.1093/brain/awh341

Heilman, K. M., Nadeau, S. E., & Beversdorf, D. O. (2003). Creative Innovation: Possible Brain Mechanisms. *Neurocase, 9,* 369-379. https://doi.org/10.1076/neur.9.5.369.16553

Jamrozik, A., McQuire, M., Cardillo, E. R., & Chatterjee, A. (2016). Metaphor: Bridging Embodiment to Abstraction. *Psychonomic Bulletin & Review, 23,* 1080-1089. https://doi.org/10.3758/s13423-015-0861-0

Juola, P., & Mikros, G. K. (2016). *Cross-Linguistic Stylometric Features: A Preliminary Investigation*.

Khan, M. I., Sameem, B., Nikoui, V., & Dehpour, A. R. (2016). Is the War on Terror Induced-Post Traumatic Stress Disorder. The Cause of Suicide Attack? An Approach from Psycho-Cognitive and Neurobiological Perspective. *Advancements in Life Sciences, 3,* 109-111.

Kimura, D., & Tanaka-Ishii, K. (2014). Study on Constants of Natural Language Texts. *Information and Media Technologies, 9,* 771-789. https://doi.org/10.5715/jnlp.21.877

Kintsch, W., & Keenan, J. (1973). Reading Rate and Retention as a Function of the Number of Propositions in the Base Structure of Sentences. *Cognitive Psychology, 5,* 257-274. https://doi.org/10.1016/0010-0285(73)90036-4

Klimek, V., Stockmeier, C., Overholser, J., Meltzer, H. Y., Kalka, S., Dilley, G., & Ordway, G. A. (1997). Reduced Levels of Norepinephrine Transporters in the Locus Coeruleus in Major Depression. *The Journal of Neuroscience, 17,* 8451-8458.

Kubát, M., & Milička, J. (2013). Vocabulary Richness Measure in Genres. *Journal of Quantitative Linguistics, 20,* 339-349. https://doi.org/10.1080/09296174.2013.830552

Lancashire, I. (2010). *Forgetful Muses: Reading the Author in the Text*. Toronto: University of Toronto Press. https://doi.org/10.3138/9781442686328

Lancashire, I., & Hirst, G. (2009). Vocabulary Changes in Agatha Christie's Mysteries as an Indication of Dementia: A Case Study. In *19th Annual Rotman Research Institute Conference, Cognitive Aging: Research and Practice* (pp. 8-10).

Le, X. (2010). *Longitudinal Detection of Dementia through Lexical and Syntactic Changes in Writing*. Master's Thesis. Toronto: Department of Computer Science, University of Toronto. http://ftp.cs.toronto.edu/pub/gh/Le-MSc-2010.pdf

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal Detection of Dementia through Lexical and Syntactic Changes in Writing: A Case Study of Three British Novelists. *Literary and Linguistic Computing, 26,* 435-461. https://doi.org/10.1093/llc/fqr013

Lieberman, P. (2016). The Evolution of Language and Thought. *Journal of Anthropological Sciences, 94,* 1-20.

Lynott, D., & Connell, L. (2009). Modality Exclusivity Norms for 423 Object Properties. *Behavior Research Methods, 41,* 558-564. https://doi.org/10.3758/BRM.41.2.558

Martin, P., & Rowe, A. (2010). *Iris Murdoch: A Literary Life*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230282964

Mesulam, M. M. (2003). Primary Progressive Aphasia—A Language-Based Dementia. *New England Journal of Medicine, 349,* 1535-1542. https://doi.org/10.1056/NEJMra022435

Murdoch, I. (2016). *Living on Paper: Letters from Iris Murdoch, 1934-1995*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400880300

Nonen, S., Kato, M., Takekita, Y., Wakeno, M., Sakai, S., Serretti, A., & Kinoshita, T. (2016). Polymorphism of rs3813034 in Serotonin Transporter Gene SLC6A4 Is Associated with the Selective Serotonin and Serotonin-Norepinephrine Reuptake Inhibitor Response in Depressive Disorder: Sequencing Analysis of SLC6A4. *Journal of clinical psychopharmacology, 36,* 27-31. https://doi.org/10.1097/JCP.0000000000000454

Ramirez, J. (2016). Suicide: Across the Life Span. *Nursing Clinics of North America, 51,* 275-286. https://doi.org/10.1016/j.cnur.2016.01.010

Rajan, K. B., Wilson, R. S., Weuve, J., Barnes, L. L., & Evans, D. A. (2015). Cognitive Impairment 18 Years before Clinical Diagnosis of Alzheimer Disease Dementia. *Neurology, 85,* 898-904. https://doi.org/10.1212/WNL.0000000000001774

Robert, P., Bremond, F., & David, R. (2016). Depression, Apathy and Alzheimer's Disease: New Perspectives. *Neurobiology of Aging, 39,* S29-S30. https://doi.org/10.1016/j.neurobiolaging.2016.01.127

Sabat, S. R., & Harré, R. (1992). The Construction and Deconstruction of Self in Alzheimer's Disease. *Ageing and Society, 12,* 443-461. https://doi.org/10.1017/S0144686X00005262

Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic Ability in Early Life and Cognitive Function and Alzheimer's Disease in Late Life: Findings from the Nun Study. *JAMA, 275,* 528-532. https://doi.org/10.1001/jama.1996.03530310034029

Sun, Y., Hunt, S., & Sah, P. (2015). Norepinephrine and Corticotropin-Releasing Hormone: Partners in the Neural Circuits That Underpin Stress and Anxiety. *Neuron, 87,* 468-470. https://doi.org/10.1016/j.neuron.2015.07.022

Szot, P. (2016). Elevated Cerebrospinal Fluid Norepinephrine in the Elderly Can Link Depression and a Reduced Glymphatic System as Risk Factors for Alzheimer's Disease. *Journal of Aging Science, 4,* 158. https://doi.org/10.4172/2329-8847.1000158

Tanaka-Ishii, K., & Aihara, S. (2015). Computational Constancy Measures of Texts—Yule's K and Rényi's Entropy. *Computational Linguistics, 41,* 481-502. https://doi.org/10.1162/COLI_a_00228

Toutanova, K., & Manning, C. D. (2000). Enriching the Knowledge Sources Used in a Maximum Entropy Part-of-Speech Tagger. In *Proceedings of the 2000 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora: Held in Conjunction with the 38th Annual Meeting of the Association for Computational Linguistics* (Volume 13, pp. 63-70). Association for Computational Linguistics. https://doi.org/10.3115/1117794.1117802

Turner, A., & Greene, E. (1977). *The Construction and Use of a Propositional Text Base* (Technical Report 63). Boulder, CO: University of Colorado, Institute for the Study of Intellectual Behavior.

Van Dantzig, S., Cowell, R. A., Zeelenberg, R., & Pecher, D. (2011). A Sharp Image or a Sharp Knife: Norms for the Modality-Exclusivity of 774 Concept-Property Items. *Behavior Research Methods, 43,* 145-154. https://doi.org/10.3758/s13428-010-0038-8

Van Gijsel, S., Speelman, D., & Geeraerts, D. (2005). *A Variationist, Corpus Linguistic Analysis of Lexical Richness*.

Van Velzen, M., & Garrard, P. (2008). From Hindsight to Insight-Retrospective Analysis of Language Written by a Renowned Alzheimer's Patient. *Interdisciplinary Science Reviews, 33,* 278-286. https://doi.org/10.1179/174327908X392852

Vermeer, A. (2000). Coming to Grips with Lexical Richness in Spontaneous Speech Data. *Language Testing, 17,* 65-83. https://doi.org/10.1191/026553200676636328

Wesson, D. W., Wilson, D. A., & Nixon, R. A. (2010). Should OLFACTORY dysfunction Be Used as a Biomarker of Alzheimer's Disease? *Expert Review of Neurotherapeutics, 10,* 633-635. https://doi.org/10.1586/ern.10.33

Wicklund, A. H., Johnson, N., & Weintraub, S. (2004). Preservation of Reasoning in Primary Progressive Aphasia: Further Differentiation from Alzheimer's Disease and the Behavioral Presentation of Frontotemporal Dementia. *Journal of Clinical and Experimental Neuropsychology, 26,* 347-355. https://doi.org/10.1080/13803390490510077

Wilson, A. N. (2004). *Iris Murdoch as I Knew Her*. New York: Random House.

Woodward, M. R., Dwyer, M. G., Amrutkar, C. V., Zivadinov, R., & Szigeti, K. (2015). Olfactory Identification Deficit as a Predictor of White Matter Tract Integrity in Alzheimer's Disease. *The American Journal of Geriatric Psychiatry, 23,* S103-S104. https://doi.org/10.1016/j.jagp.2014.12.104

Zabelina, D. L., O'Leary, D., Pornpattananangkul, N., Nusslock, R., & Beeman, M. (2015). Creativity and Sensory Gating Indexed by the P50: Selective versus Leaky Sensory Gating in Divergent Thinkers and Creative Achievers. *Neuropsychologia, 69,* 77-84. https://doi.org/10.1016/j.neuropsychologia.2015.01.034

Supplementary Data

Table S1. Iris Murdoch richness mann-whitney U-test 12-year ranks.

Ranks				
	AD	N	Mean Rank	Sum of Ranks
Rank of RICHNESS by AD	1	21	15.40	323.50
	2	5	5.50	27.50
	Total	26		

Table S2. Iris Murdoch richness mann-whitney U-test 12-year statistics.

Test Statistics[b]	
	Rank of RICHNESS by AD
Mann-Whitney U	12.500
Wilcoxon W	27.500
Z	−2.605
Asymp. Sig. (2-tailed)	0.009
Exact Sig. [2*(1-tailed Sig.)]	0.006[a]

a. Not corrected for ties. b. Grouping Variable: AD.

Table S3. P.D. James richness mann-whitney U-test 12-year ranks.

Ranks				
	AD	N	Mean Rank	Sum of Ranks
Rank of Richness	1	15	10.07	151.00
	2	4	9.75	39.00
	Total	19		

Table S4. P.D. James richness mann-whitney U-test 12-year statistics.

Test Statistics[b]	
	Rank of Richness
Mann-Whitney U	29.000
Wilcoxon W	39.000
Z	−0.100
Asymp. Sig. (2-tailed)	0.920
Exact Sig. [2*(1-tailed Sig.)]	0.961[a]

a. Not corrected for ties. b. Grouping Variable: AD.

Table S5. Iris Murdoch aggregated content and function word ratios.

Work	Content Words	Function Words	Ratio
B1	2284	1716	0.751313485
B2	2307	1693	0.733853489
B3	2424	1576	0.650165017
B4	2284	1716	0.751313485
B5	2281	1719	0.753616835
B6	2439	1561	0.6400164
B7	2368	1632	0.689189189
B8	2227	1773	0.796138303
B9	2321	1679	0.723395088
B10	2407	1593	0.661819693
B11	2487	1513	0.60836349
B12	2404	1596	0.663893511
B13	2451	1549	0.631986944
B14	2516	1484	0.589825119
B15	2320	1680	0.724137931
B16	2478	1522	0.614205004
B17	2341	1659	0.708671508
B18	2427	1573	0.648125258
B19	2391	1609	0.672940192
B20	2461	1539	0.625355547
B21	2343	1657	0.707212975
B22	2432	1568	0.644736842
B23	2465	1535	0.622718053
B24	2517	1483	0.589193484
B25	2500	1500	0.6
B26	2481	1519	0.612253124

Table S6. P.D. James aggregated content and function word ratios.

Work	Content Words	Function Words	Ratio
B1	2400	1600	0.666666667
B2	2445	1555	0.63599182
B3	2317	1683	0.726370306
B4	2425	1575	0.649484536
B5	2373	1627	0.685630004
B6	2375	1625	0.684210526
B7	2267	1733	0.764446405
B8	2307	1693	0.733853489
B9	2382	1618	0.679261125
B10	2279	1721	0.75515577
B11	2308	1692	0.733102253
B12	2340	1660	0.709401709
B13	2328	1672	0.718213058
B14	2369	1631	0.68847615
B15	2252	1748	0.776198934
B16	2371	1629	0.687051877
B17	2404	1596	0.663893511
B18	2358	1642	0.696352841
B19	2414	1586	0.657000829

Table S7. Iris Murdoch function to content word ratio mann-whitney U-test statistics.

Test Statistics[b]	
	Rank of Ratio
Mann-Whitney U	24.000
Wilcoxon W	45.000
Z	−2.191
Asymp. Sig. (2-tailed)	0.028
Exact Sig. [2*(1-tailed Sig.)]	0.028[a]

a. Not corrected for ties. b. Grouping Variable: AD.

Table S8. Iris Murdoch function to content word ratio mann-whitney U-test ranks.

Ranks				
	AD	N	Mean Rank	Sum of Ranks
Rank of Ratio	1	20	15.30	306.00
	2	6	7.50	45.00
	Total	26		

Table S9. P.D. James function to content word ratio mann-whitney U-test ranks.

Ranks				
	AD	N	Mean Rank	Sum of Ranks
Rank of Ratio	1	15	10.87	163.00
	2	4	6.75	27.00
	Total	19		

Table S10. P.D. James function to content word ratio mann-whitney U-test statistics.

Test Statistics[b]	
	Rank of Ratio
Mann-Whitney U	17.000
Wilcoxon W	27.000
Z	−1.300
Asymp. Sig. (2-tailed)	0.194
Exact Sig. [2*(1-tailed Sig.)]	0.221[a]

a. Not corrected for ties. b. Grouping Variable: AD.

Table S11. Iris Murdoch sensory word mann-whitney U-test 12 year ranks.

Ranks				
	Group	N	Mean Rank	Sum of Ranks
Rank of Sensory by Group	1	20	15.60	312.00
	2	6	6.50	39.00
	Total	26		

Table S12. Iris Murdoch sensory word mann-whitney U-test 12-year statistics.

Test Statistics[b]	
	Rank of Sensory by Group
Mann-Whitney U	18.000
Wilcoxon W	39.000
Z	−2.559
Asymp. Sig. (2-tailed)	0.011
Exact Sig. [2*(1-tailed Sig.)]	0.009[a]

a. Not corrected for ties. b. Grouping Variable: Group.

Table S13. P.D. James sensory word mann-whitney U-test 12 year ranks.

Ranks				
	Group	N	Mean Rank	Sum of Ranks
Rank of Sensory	1	15	11.53	173.00
	2	4	4.25	17.00
	Total	19		

Table S14. P.D. James sensory word mann-whitney U-test 12-year statistics.

Test Statistics[b]	
	Rank of Sensory
Mann-Whitney U	7.000
Wilcoxon W	17.000
Z	−2.300
Asymp. Sig. (2-tailed)	0.021
Exact Sig. [2*(1-tailed Sig.)]	0.020[a]

a. Not corrected for ties. b. Grouping Variable: Group.

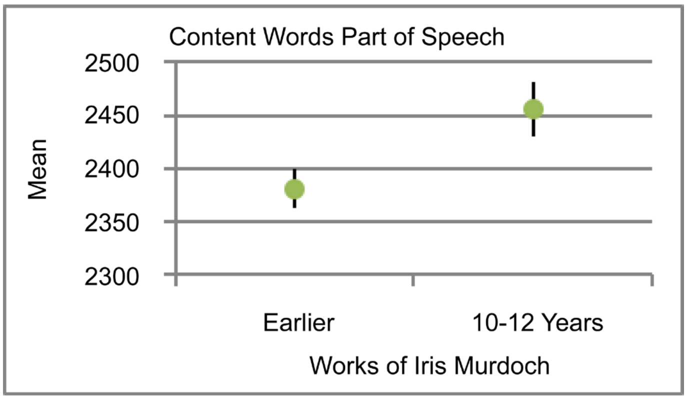

Figure S1. Iris Murdoch Content Words POS Mean with Standard Error bars. We see the aggregated Content Words part-of-speech is higher than the earlier work for the 12 year period before the diagnosis of AD, and there is no overlap between the Standard Error means. There is more variability in the 12 years period (25.6 versus 18.2).

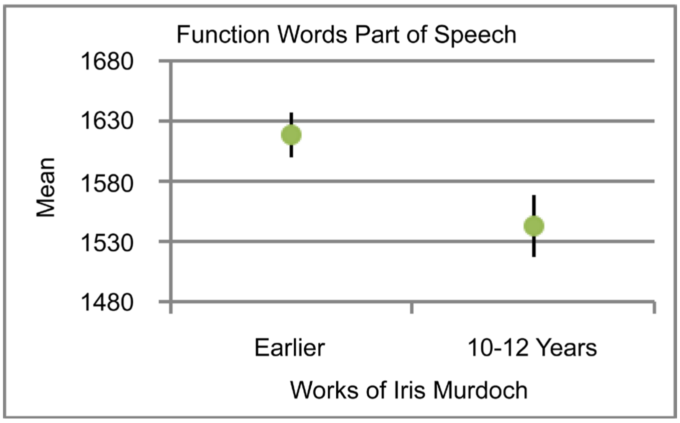

Figure S2. Iris Murdoch Function Words POS Mean with Standard Error bars. We see the aggregated Function Words part-of-speech is part-of-speech is lower than the earlier work for the 12 year period before the diagnosis of AD, and there is no overlap between the Standard Error means. There is more variability in the 12 years period (25.6 versus 18.2).

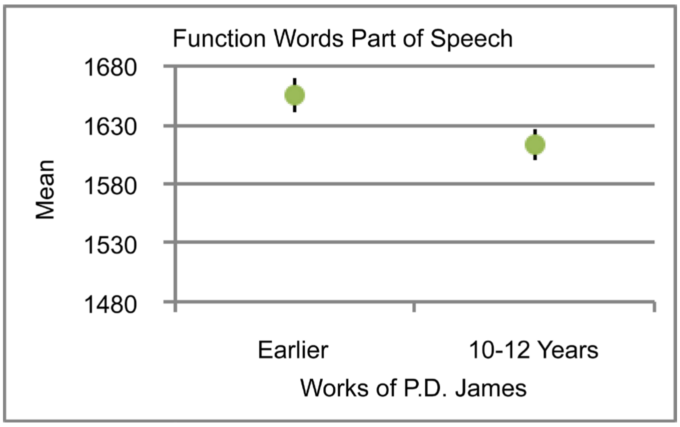

Figure S3. P.D. James Content Words POS Mean with Standard Error bars. We see the aggregated Content Words part-of-speech is higher than the earlier work for the 10 - 12 year period before death, and there is no overlap between the Standard Error means. There is less variability in the 10 - 12 years period (13.28 versus 14.81).

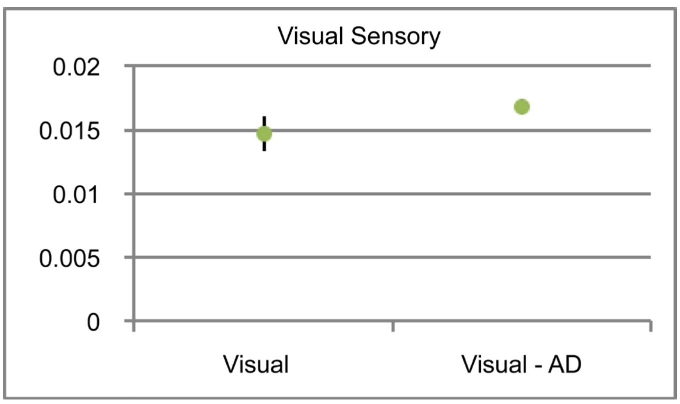

Figure S4. P.D. James Function Words POS Mean with Standard Error bars. We see the aggregated Function Words part-of-speech is lower than the earlier work for the 10 - 12 year period before the diagnosis of AD, and there is no overlap between the Standard Error means. There is less variability in the 10 - 12 years period (13.28 versus 14.81).

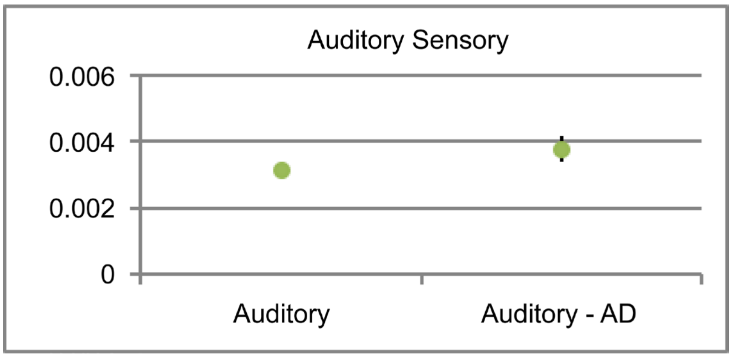

Figure S5. Iris Murdoch visual sensory mean with standard error bars.

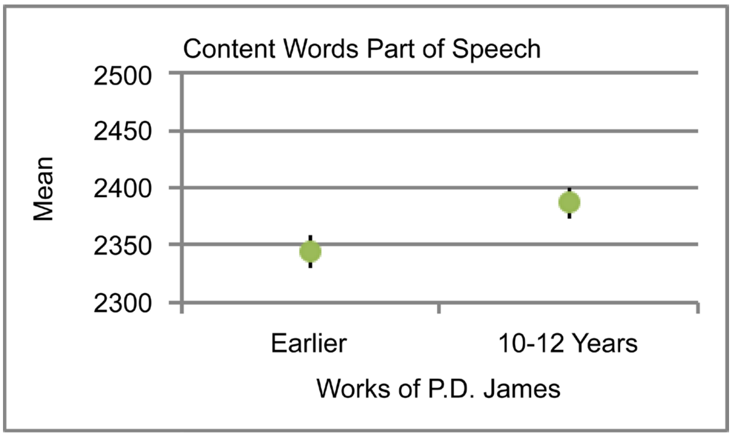

Figure S6. Iris Murdoch auditory sensory mean with standard error bars.

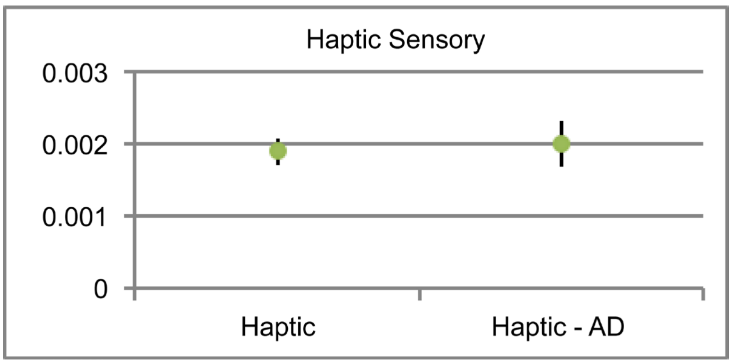

Figure S7. Iris Murdoch haptic sensory mean with standard error bars.

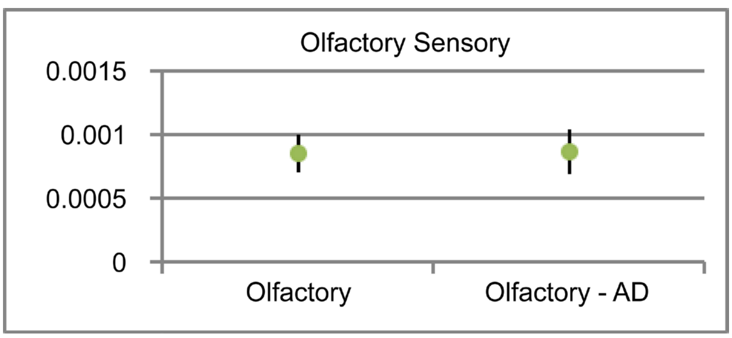

Figure S8. Iris Murdoch olfactory sensory mean with standard error bars.

Figure S9. Iris Murdoch gustatory sensory mean with standard error bars.

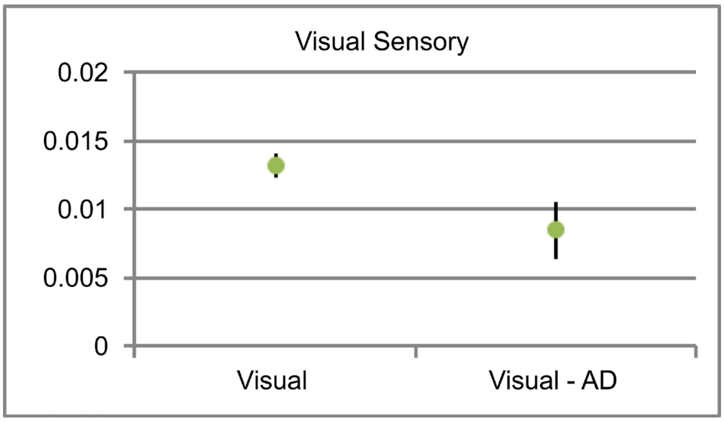

Figure S10. P.D. James visual sensory mean with standard error bars.

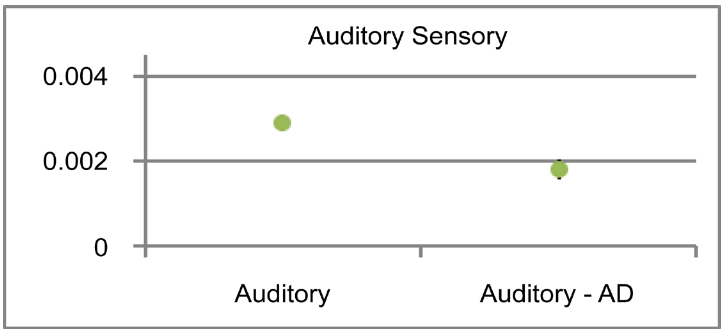

Figure S11. P.D. James auditory sensory mean with standard error bars.

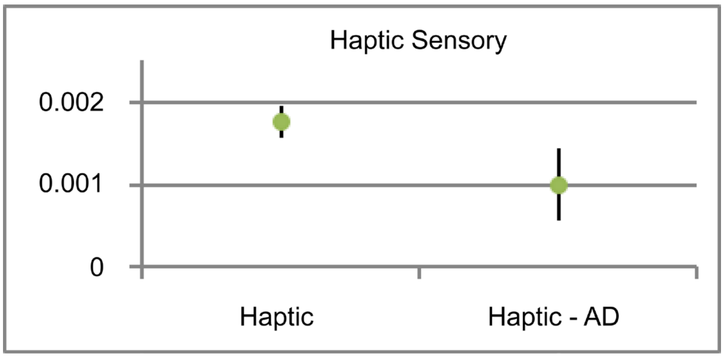

Figure S12. P.D. James haptic sensory mean with standard error bars.

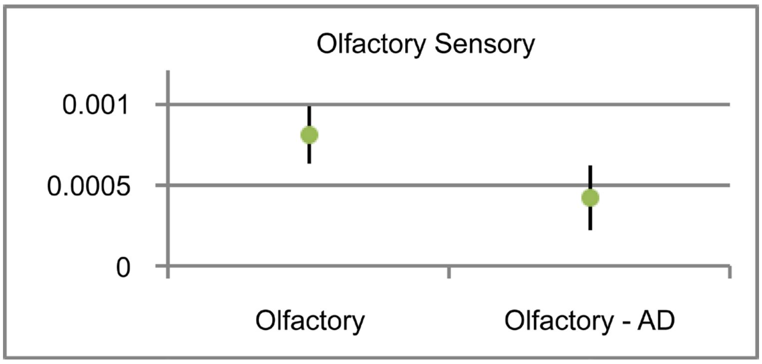

Figure S13. P.D. James olfactory sensory mean with standard error bars.

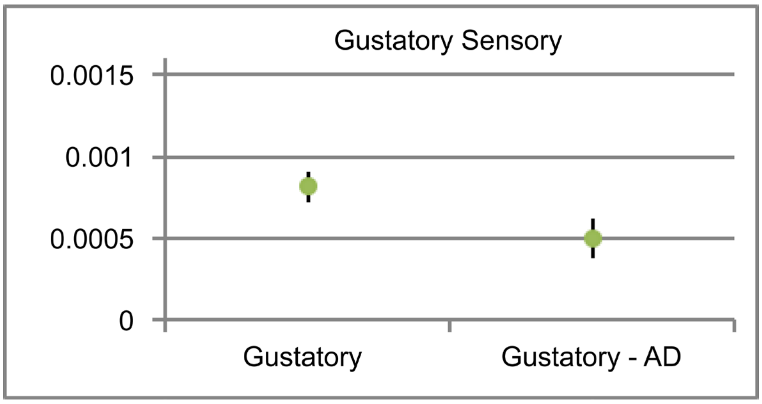

Figure S14. P.D. James gustatory sensory mean with standard error bars.