The Well-Formedness of Two Psychoanalytic Word Categories in Portuguese Texts

Andrew Wilson (Lancaster, United Kingdom)

1 INTRODUCTION

Freud (1900) proposed the existence of two different modes of thought, which he called respectively primary process and secondary process. Primary process thought is, by nature, sensation-oriented, concrete, and ignorant of time, space, and social institutions. It is most closely associated with early childhood, but is also hypothesized to predominate in dreams and other altered states of consciousness such as hypnosis, mystical religious experiences, and under the influence of certain psychotropic drugs. In contrast, secondary process thought – which is the normal conscious mode of cognition for adults – is logical and oriented to time, space, and society. Very similar distinctions in thought processes have also been suggested by such theoretically diverse scholars such as Lévy-Bruhl (1910/1966), Goldstein (1939), Werner (1948), Piaget (1954), Sorokin (1957), Aulagnier (1975/2001), Bucci (1997), and Kristeva (1996/1997).

As a measure of primary and secondary process thought in texts, Martindale (1975) developed the Regressive Imagery Dictionary (hereafter "RID"), a computer-readable lexicon for content analysis. Originally built for the analysis of English-language texts, the RID now exists in versions for several other languages, including Portuguese. It contains two main summary categories – primary and secondary process – as well as a few further categories, such as a set of emotion types. Each of the two main summary categories is made up of several subcategories, and, in the case of primary process, these subcategories also contain further, smaller subcategories. Table 1 shows, in English, the composition of the RID, together with some sample words for each subcategory.

The RID has demonstrated good construct validity in many studies over the past thirty years or so. That is to say that frequency patterns of primary and secondary process words, which were predicted on the basis of the theoretical literature, have been found to hold empirically when operationalized using the RID. For example, in relation to the hypothesis linking primary process thought with early childhood, it has been found that stories composed by older children contain fewer primary process words than those composed by younger children (West, Martindale & Sutton-Smith, 1985). An increase in the use of primary process words has been strongly linked to drug-induced altered states (Martindale & Fischer, 1977; West, Martindale, Hines & Roth, 1983), and fluctuations in the frequency of primary process across the text of the bible have been shown to correlate with a five-stage theory of mystical development (West, 1991; Wilson, 2007). A study of language produced under hypnosis did not find an increase in primary process over the normal waking state, but it did find a significant decrease in secondary process (Elter-Nodvin, 2000). In accordance with psycho-analytic theory, primary process lexis has also been found to be elevated in the speech of schizophrenics (West & Martindale, 1988), in fetish fantasies (Wilson, 2002), and in the folk tales of more primitive societies (Martindale, 1976).

The present study aims to examine the validity of the RID from a different angle, i.e., that of the rank-frequency distributions of words within its two main component categories. Although research on the rank-frequency distributions of words in texts has a long history (cf. Prün, 1999; Baayen, 2001; Altmann, 2002), and is equally well established in contemporary quantitative linguistics, there has so far been relatively little research on the distribution of word frequencies within other linguistic categories that have been identified in a text. One exception to this observation is the work by Uhliřova (1995), who examined the rank-frequency distribution of word forms within word-length categories and found that they could be modelled by the well-known Zipf-Mandelbrot distribution. Another, more recent, exception is the work by Popescu, Best & Altmann (2007), who have examined the distribution of word forms within part-of-speech categories. This study has perhaps an even wider import than Uhliřova's, since it deals with interpretive categories constructed by scientists rather than with categories which derive from purely formal properties of the text (such as word lengths).

Popescu, Best & Altmann (2007) base their approach on word-frequency spectra. A word-frequency spectrum for a text is obtained when the number of different words that fall into each frequency category is listed in the natural ascending order of the frequency categories – i.e., first, the number of words which occur only once, then the number which occur twice, then the number which occur three times, and so on. Proceeding from these frequency spectra, Popescu, Best & Altmann (2007) found that the right-truncated Zeta distribution, which is the simplest and most commonly used model for the frequency spectra of words within a text, could also be fitted to the frequency spectra of words within discrete part-of-speech categories in texts – for example, the number of nouns that occur once, the number that occur twice, and so on.[1]

This observation leads straightforwardly to the following general hypothesis about the frequencies of words within categories: "if a linguistic class is constructed "naturally", then its elements abide by a proper rank-frequency distribution of the Zipf type" (Strauss, Fan & Altmann, 2008, 94). The present study aims to test whether this hypothesis holds for the primary and secondary process categories in the RID when they are applied in the content analysis of individual texts. If the right-truncated Zeta distribution can be fitted adequately to the words within both categories, then – in addition to possessing psychoanalytic construct validity – the categories would also appear to constitute well-formed natural properties of texts.

2 DATA AND METHOD

The main set of data for this study consisted of six complete books selected from a Portuguese translation of the Bible (Bíblia Católica v2.0, retrieved April 20, 2006 from http://www.bibliacatolica.com.br). The books chosen were the four gospels of Matthew, Mark, Luke and John, together with the Apocalypse (also known as the Book of Revelation) and the Epistle to the Romans. The latter two books were particularly chosen for this preliminary experiment as they were known from previous studies to have very different overall frequencies of primary and secondary process words (West, 1991; Wilson, 2007).

Since translation texts can sometimes be problematic for quantitative modeling, as they are not spontaneously composed texts in the target language, a further sample was also included. This was chosen at random (subject to availability as an electronic text) from the reading list in Portuguese literature for the University of Cambridge's degree course in Portuguese. It was the 19th century romance *A Dama-Pé-De-Cabra* by Alexandre Herculano (retrieved March 19, 2007, from http://www.gutenberg.org/1/7/0/0/17005/, as part of his *Lendas e Narrativas, Tomo II*).

[1] This is the original distribution proposed by Zipf (1935), simply truncated at the maximum word frequency for a given text. A fair amount of recent work, such as that of Uhlířová (1995), has tended to neglect this distribution in favour of other related distributions such as the Zipf-Mandelbrot distribution; however, since it has only one parameter, Zipf's original zeta distribution has the advantage of being simpler than its relatives, and this makes it much easier to interpret.

Table 1
Categories and subcategories in the RID, based on Martindale (1975)

Major subcategory	Minor subcategory	Example words
PRIMARY PROCESS		
Drive		
	Oral	*breast, drink, lip*
	Anal	*sweat, rot, dirty*
	Sex	*lover, kiss, naked*
Sensation		
	General sensation	*fair, charm, beauty*
	Touch	*touch, thick, stroke*
	Taste	*sweet, taste, bitter*
	Odor	*breath, perfume, scent*
	Sound	*hear, voice, sound*
	Vision	*see, light, look*
	Cold	*cold, winter, snow*
	Hard	*rock, stone, hard*
	Soft	*soft, gentle, tender*
Perceptual Disinhibition		
	Passivity	*die, lie, bed*
	Voyage	*wander, desert, beyond*
	Random movement	*wave, roll, spread*
	Diffusion	*shade, shadow, cloud*
	Chaos	*wild, crowd, ruin*
Regressive Cognition		
	Unknown	*secret, strange, unknown*
	Timeless	*eternal, forever, immortal*
	Altered consciousness	*dream, sleep, wake*
	Brink passage	*road, wall, door*
	Narcissism	*eye, heart, hand*
	Concreteness	*at, where, over*
Icarian Imagery		
	Ascend	*rise, fly, throw*
	Height	*up, sky, high*
	Descend	*fall, drop, sink*
	Depth	*down, deep, beneath*
	Fire	*sun, fire, flame*
	Water	*sea, water, stream*
SECONDARY PROCESS		
Abstract Thought		*know, may, thought*
Social Behavior		*say, tell, call*
Instrumental Behavior		*make, find, work*
Restraint		*must, stop, bind*
Order		*simple, measure, array*
Temporal Reference		*when, now, then*
Moral Imperative		*should, right, virtue*

Table 2 shows, for the seven texts, their chi-square ratios of primary to secondary process, calculated according to the formula pattern proposed by Ziegler, Best & Altmann (2002, 76): X^2 = (primary process – secondary process)^2 / (primary process + secondary process). This equation is distributed as the chi-square with one degree of freedom. A book leans significantly towards primary or secondary process content if the chi-square value is greater than 3.84 ($p < 0.05$); the direction of the trend can be determined simply from the process type with the higher frequency.

Table 2
Chi-square ratios of primary to secondary process for the seven books

Book	F(PP)	F(SP)	X^2	Trend
Matthew	1657	1546	3.85	PP > SP
Mark	1078	863	23.82	PP > SP
Luke	1592	1608	0.08	not significant
John	1338	1276	1.47	not significant
Romans	596	768	21.69	SP > PP
Apocalypse	1143	492	259.21	PP > SP
Herculano	779	349	163.92	PP > SP

To identify the primary and secondary process words in the texts and measure their frequencies, a computerized content analysis was performed using the Portuguese version of the RID by Cardoso e Cunha, Detry, Hogenraad & Martindale (version of 28 August 1996). The RID was applied to the texts using the PROTAN suite of programs for content analysis (Hogenraad, Daubies, Bestgen & Mahau, 2003). PROTAN first divides the input file into the segments premarked by the analyst (in this case, the individual books). The majority of words in these segments are then reduced by another procedure to their basic, uninflected forms. The reduced text is finally matched against the entries in the RID. For each text segment, PROTAN produces a frequency count of each requested RID category, which shows how many word occurrences fell into that category. PROTAN can also produce a listing of each individual word form that fell into a given category and of its frequency in each text segment: this listing was used as the basis for the frequency spectrum calculations in this paper. Note that the frequencies studied in this paper are those for the actual word forms in the texts, not the reduced forms used by PROTAN for its internal dictionary matching procedure.

For each text, two frequency spectra were produced: one for the frequencies of words in the category of primary process and one for the frequencies of words in the category of secondary process. Using the Altmann Fitter software, the right-truncated Zeta distribution was then fitted to the empirical frequencies for each spectrum. The right-truncated Zeta distribution is given by:

$$P_x = x^{-a} / F(R), x = 1,2,...,R$$

where a = a parameter and $F(R) = \sum_{i=1}^{R} i^{-a}$.

The probability of the chi-squared test between the empirical and estimated frequencies was used as a measure of goodness of fit. The distribution was considered to be a good model if $P(X^2) \geq 0.05$.

3 RESULTS

Tables 3 to 16 show the results of fitting the right-truncated Zeta distribution to the fourteen frequency spectra. In the tables, the following abbreviations are used:

x = the frequency class (e.g. 1 = words with a frequency of 1)
g(x) = the number of words falling into frequency class x in the observed data
NP_x = the estimated value of g(x) generated by the right-truncated Zeta distribution
a = the parameter of the right-truncated Zeta distribution
R = the value of x at which the right-truncated Zeta distribution is truncated
DF = the number of degrees of freedom of the chi-squared test
X^2 = value of the chi-squared test
$P(X^2)$ = probability of the chi-squared test

It will be seen that the right-truncated Zeta distribution could be fitted successfully in all cases. In the majority of cases, the value of $P(X^2)$ was very high, demonstrating an excellent quality of fit. In three cases – Matthew (primary process), Mark (primary process), and Apocalypse (secondary process) – $P(X^2)$ was rather low, but even here the quality of fit for the theoretical model was still acceptable.

Table 3: Matthew. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	275	265.87	26	1	0.47	51	0	0.13	76	0	0.06	101	0	0.03	126	0	0.02	151	0	0.02
2	61	69.09	27	1	0.44	52	0	0.12	77	0	0.06	102	0	0.03	127	0	0.02	152	0	0.02
3	23	31.41	28	0	0.41	53	0	0.12	78	0	0.06	103	0	0.03	128	0	0.02	153	0	0.02
4	21	17.95	29	1	0.38	54	1	0.11	79	0	0.05	104	0	0.03	129	0	0.02	154	0	0.01
5	10	11.63	30	0	0.36	55	0	0.11	80	0	0.05	105	0	0.03	130	0	0.02	155	0	0.01
6	7	8.16	31	0	0.34	56	0	0.11	81	1	0.05	106	0	0.03	131	0	0.02	156	0	0.01
7	7	6.05	32	0	0.32	57	0	0.1	82	0	0.05	107	0	0.03	132	0	0.02	157	0	0.01
8	6	4.67	33	0	0.3	58	0	0.1	83	0	0.05	108	0	0.03	133	0	0.02	158	0	0.01
9	2	3.71	34	0	0.28	59	0	0.1	84	0	0.05	109	0	0.03	134	0	0.02	159	0	0.01
10	6	3.02	35	1	0.26	60	0	0.09	85	0	0.05	110	0	0.03	135	0	0.02	160	0	0.01
11	8	2.51	36	0	0.25	61	0	0.09	86	0	0.05	111	0	0.03	136	0	0.02	161	0	0.01
12	4	2.12	37	0	0.24	62	0	0.09	87	0	0.05	112	0	0.03	137	0	0.02	162	0	0.01
13	4	1.82	38	0	0.23	63	0	0.08	88	0	0.04	113	0	0.03	138	0	0.02	163	0	0.01
14	0	1.57	39	0	0.21	64	0	0.08	89	0	0.04	114	0	0.03	139	0	0.02	164	0	0.01
15	1	1.37	40	0	0.2	65	0	0.08	90	0	0.04	115	0	0.03	140	0	0.02	165	0	0.01
16	2	1.21	41	0	0.19	66	0	0.08	91	0	0.04	116	0	0.03	141	0	0.02	166	0	0.01
17	2	1.08	42	0	0.19	67	0	0.07	92	0	0.04	117	0	0.03	142	0	0.02	167	0	0.01
18	0	0.96	43	0	0.18	68	0	0.07	93	0	0.04	118	0	0.02	143	0	0.02	168	0	0.01
19	0	0.87	44	0	0.17	69	0	0.07	94	0	0.04	119	0	0.02	144	0	0.02	169	0	0.01
20	1	0.79	45	0	0.16	70	0	0.07	95	0	0.04	120	0	0.02	145	0	0.02	170	0	0.01
21	0	0.71	46	0	0.16	71	0	0.07	96	0	0.04	121	0	0.02	146	0	0.02	171	0	0.01
22	0	0.65	47	0	0.15	72	0	0.07	97	0	0.04	122	0	0.02	147	0	0.02	172	0	0.01
23	0	0.6	48	0	0.14	73	1	0.06	98	0	0.04	123	0	0.02	148	0	0.02	173	0	0.01
24	0	0.55	49	0	0.14	74	0	0.06	99	0	0.04	124	0	0.02	149	0	0.02	174	0	0.01
25	2	0.51	50	0	0.13	75	0	0.06	100	0	0.03	125	0	0.02	150	0	0.02	175	0	0.01
	a = 1.94			**R = 176**			**X^2 = 37.71**				**DF = 28**			**P(X^2) = 0.10**				176	1	0.01

Table 4: Matthew. Secondary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	276	271.5	26	0	0.39	51	0	0.1	76	0	0.05	101	0	0	126	0	0.02	151	0	0
2	64	67.49	27	1	0.36	52	0	0.1	77	0	0.04	102	0	0	127	0	0.02	152	0	0
3	28	29.89	28	0	0.34	53	1	0.1	78	0	0.04	103	0	0	128	0	0.02	153	0	0
4	21	16.77	29	1	0.31	54	0	0.1	79	0	0.04	104	0	0	129	0	0.02	154	0	0
5	9	10.72	30	1	0.29	55	1	0.1	80	0	0.04	105	0	0	130	0	0.02	155	0	0
6	9	7.43	31	0	0.27	56	0	0.1	81	0	0.04	106	0	0	131	0	0.02	156	0	0
7	4	5.45	32	0	0.26	57	0	0.1	82	0	0.04	107	0	0	132	0	0.02	157	0	0
8	5	4.17	33	0	0.24	58	0	0.1	83	0	0.04	108	0	0	133	0	0.01	158	0	0
9	2	3.29	34	0	0.23	59	1	0.1	84	0	0.04	109	0	0	134	0	0.01	159	0	0
10	2	2.66	35	1	0.22	60	0	0.1	85	0	0.04	110	0	0	135	0	0.01	160	0	0
11	2	2.2	36	0	0.2	61	0	0.1	86	0	0.04	111	0	0	136	0	0.01	161	0	0
12	3	1.85	37	0	0.19	62	0	0.1	87	0	0.03	112	0	0	137	0	0.01	162	0	0
13	0	1.57	38	0	0.18	63	0	0.1	88	0	0.03	113	0	0	138	0	0.01	163	0	0
14	0	1.36	39	0	0.17	64	0	0.1	89	0	0.03	114	0	0	139	0	0.01	164	0	0
15	2	1.18	40	1	0.16	65	0	0.1	90	0	0.03	115	0	0	140	0	0.01	165	0	0
16	0	1.04	41	0	0.16	66	0	0.1	91	0	0.03	116	0	0	141	0	0.01	166	0	0
17	2	0.92	42	0	0.15	67	0	0.1	92	0	0.03	117	0	0	142	0	0.01	167	1	0
18	1	0.82	43	0	0.14	68	0	0.1	93	0	0.03	118	0	0	143	0	0.01			
19	1	0.73	44	0	0.14	69	0	0.1	94	0	0.03	119	0	0	144	0	0.01			
20	1	0.66	45	0	0.13	70	0	0.1	95	0	0.03	120	0	0	145	0	0.01			
21	0	0.6	46	0	0.12	71	0	0.1	96	0	0.03	121	0	0	146	0	0.01			
22	1	0.55	47	0	0.12	72	0	0.1	97	0	0.03	122	0	0	147	0	0.01			
23	0	0.5	48	0	0.11	73	1	0.1	98	0	0.03	123	0	0	148	0	0.01			
24	0	0.46	49	0	0.11	74	0	0.1	99	0	0.03	124	0	0	149	0	0.01			
25	0	0.42	50	0	0.11	75	0	0.1	100	0	0.03	125	0	0	150	0	0.01			
	a = 2.01			**R = 167**			**X^2 = 10.62**				**DF = 25**			**P(X^2) = 0.99**						

Table 5: Mark. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	202	193	21	1	0.6	41	0	0.2	61	0	0.1	81	0	0	101	0	0
2	46	51.1	22	2	0.5	42	0	0.2	62	0	0.1	82	0	0	102	0	0
3	19	23.5	23	0	0.5	43	0	0.1	63	0	0.1	83	0	0	103	0	0
4	13	13.6	24	0	0.4	44	0	0.1	64	0	0.1	84	0	0	104	0	0
5	8	8.84	25	0	0.4	45	1	0.1	65	0	0.1	85	0	0	105	0	0
6	7	6.24	26	0	0.4	46	0	0.1	66	0	0.1	86	0	0	106	0	0
7	5	4.64	27	0	0.4	47	0	0.1	67	0	0.1	87	0	0	107	0	0
8	4	3.6	28	0	0.3	48	0	0.1	68	0	0.1	88	0	0	108	0	0
9	8	2.87	29	1	0.3	49	0	0.1	69	0	0.1	89	0	0	109	0	0
10	3	2.35	30	0	0.3	50	0	0.1	70	0	0.1	90	0	0	110	0	0
11	2	1.96	31	0	0.3	51	0	0.1	71	0	0.1	91	0	0	111	0	0
12	0	1.66	32	0	0.3	52	0	0.1	72	0	0.1	92	0	0	112	0	0
13	0	1.42	33	0	0.2	53	0	0.1	73	0	0.1	93	0	0	113	0	0
14	0	1.23	34	1	0.2	54	0	0.1	74	0	0.1	94	0	0	114	0	0
15	4	1.08	35	0	0.2	55	0	0.1	75	0	0.1	95	0	0	115	0	0
16	1	0.95	36	0	0.2	56	0	0.1	76	0	0.1	96	0	0	116	1	0
17	1	0.85	37	0	0.2	57	0	0.1	77	0	0.1	97	0	0			
18	0	0.76	38	0	0.2	58	0	0.1	78	0	0.1	98	0	0			
19	0	0.69	39	0	0.2	59	0	0.1	79	0	0	99	0	0			
20	1	0.62	40	0	0.2	60	0	0.1	80	0	0	100	0	0			
	a = 1.91			**R = 116**			**X^2 = 26.23**				**DF = 23**			**P(X^2) = 0.29**			

Table 6: Mark. Secondary Process

x	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	x	g(x)	NPx
1	205	201	13	1	0.9	25	0	0.2	37	0	0.1	49	0	0.1	61	0	0	73	0	0
2	48	46.4	14	0	0.8	26	0	0.2	38	0	0.1	50	0	0.1	62	0	0	74	0	0
3	16	19.7	15	1	0.7	27	0	0.2	39	0	0.1	51	0	0.1	63	0	0	75	0	0
4	9	10.7	16	1	0.6	28	0	0.2	40	1	0.1	52	0	0.1	64	0	0	76	0	0
5	5	6.66	17	2	0.5	29	0	0.2	41	0	0.1	53	0	0	65	0	0	77	0	0
6	4	4.53	18	0	0.4	30	0	0.2	42	0	0.1	54	0	0	66	0	0	78	0	0
7	4	3.27	19	3	0.4	31	0	0.1	43	0	0.1	55	0	0	67	0	0	79	0	0
8	2	2.46	20	0	0.4	32	1	0.1	44	0	0.1	56	0	0	68	0	0	80	0	0
9	0	1.92	21	0	0.3	33	1	0.1	45	0	0.1	57	0	0	69	0	0	81	0	0
10	3	1.54	22	1	0.3	34	0	0.1	46	0	0.1	58	0	0	70	0	0	82	1	0
11	1	1.25	23	0	0.3	35	0	0.1	47	0	0.1	59	0	0	71	0	0			
12	0	1.04	24	0	0.2	36	0	0.1	48	0	0.1	60	0	0	72	0	0			
	a = 2.12			**R = 82**			**X^2 = 11.32**				**DF = 16**			**P(X^2) = 0.79**						

Table 7: Luke. Primary Process

x	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	x	g(x)	NPx
1	271	262.9	31	0	0.33	61	0	0.09	91	0	0.04	121	0	0.02	151	0	0.02	181	0	0
2	58	68.29	32	0	0.31	62	0	0.09	92	0	0.04	122	0	0.02	152	0	0.02	182	0	0
3	28	31.04	33	0	0.29	63	0	0.08	93	0	0.04	123	0	0.02	153	0	0.01	183	0	0
4	18	17.74	34	0	0.28	64	0	0.08	94	0	0.04	124	0	0.02	154	0	0.01	184	0	0
5	17	11.49	35	1	0.26	65	0	0.08	95	0	0.04	125	0	0.02	155	0	0.01	185	0	0
6	6	8.06	36	0	0.25	66	0	0.08	96	0	0.04	126	0	0.02	156	0	0.01	186	0	0
7	6	5.97	37	0	0.23	67	0	0.07	97	0	0.04	127	0	0.02	157	0	0.01	187	0	0
8	5	4.61	38	0	0.22	68	0	0.07	98	0	0.04	128	0	0.02	158	0	0.01	188	0	0
9	6	3.66	39	0	0.21	69	0	0.07	99	0	0.03	129	0	0.02	159	0	0.01	189	0	0
10	4	2.99	40	0	0.2	70	0	0.07	100	0	0.03	130	0	0.02	160	0	0.01	190	0	0
11	1	2.48	41	0	0.19	71	0	0.07	101	0	0.03	131	0	0.02	161	0	0.01	191	0	0
12	5	2.09	42	0	0.18	72	0	0.06	102	0	0.03	132	0	0.02	162	0	0.01	192	0	0
13	2	1.79	43	0	0.18	73	0	0.06	103	0	0.03	133	0	0.02	163	0	0.01	193	0	0
14	1	1.55	44	0	0.17	74	0	0.06	104	0	0.03	134	0	0.02	164	0	0.01	194	0	0
15	1	1.36	45	0	0.16	75	0	0.06	105	0	0.03	135	0	0.02	165	0	0.01	195	0	0
16	1	1.2	46	0	0.15	76	0	0.06	106	0	0.03	136	0	0.02	166	0	0.01	196	1	0
17	2	1.06	47	0	0.15	77	0	0.06	107	0	0.03	137	0	0.02	167	0	0.01			
18	1	0.95	48	0	0.14	78	1	0.06	108	0	0.03	138	0	0.02	168	0	0.01			
19	0	0.86	49	0	0.14	79	0	0.05	109	0	0.03	139	0	0.02	169	0	0.01			
20	1	0.78	50	0	0.13	80	0	0.05	110	0	0.03	140	0	0.02	170	0	0.01			
21	1	0.71	51	0	0.13	81	0	0.05	111	0	0.03	141	0	0.02	171	0	0.01			
22	0	0.64	52	0	0.12	82	0	0.05	112	0	0.03	142	0	0.02	172	0	0.01			
23	2	0.59	53	0	0.12	83	0	0.05	113	0	0.03	143	0	0.02	173	0	0.01			
24	0	0.54	54	0	0.11	84	0	0.05	114	0	0.03	144	0	0.02	174	0	0.01			
25	2	0.5	55	0	0.11	85	0	0.05	115	0	0.03	145	0	0.02	175	0	0.01			
26	1	0.47	56	0	0.1	86	0	0.05	116	0	0.03	146	0	0.02	176	0	0.01			
27	0	0.43	57	1	0.1	87	0	0.04	117	0	0.03	147	0	0.02	177	0	0.01			
28	0	0.4	58	0	0.1	88	0	0.04	118	0	0.02	148	0	0.02	178	0	0.01			
29	1	0.38	59	0	0.09	89	0	0.04	119	0	0.02	149	0	0.02	179	0	0.01			
30	0	0.35	60	0	0.09	90	0	0.04	120	0	0.02	150	0	0.02	180	0	0.01			
	a = 1.94			**R = 196**			**X^2 = 19.56**				**DF = 28**			**P(X^2) = 0.88**						

Table 8: Luke. Secondary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	329	325	31	2	0.2	61	1	0.1	91	0	0	121	0	0	151	0	0
2	72	73.6	32	0	0.2	62	0	0.1	92	0	0	122	0	0	152	0	0
3	29	30.9	33	0	0.2	63	0	0.1	93	0	0	123	0	0	153	0	0
4	18	16.7	34	0	0.2	64	1	0	94	0	0	124	0	0	154	0	0
5	8	10.3	35	0	0.2	65	0	0	95	0	0	125	0	0	155	0	0
6	6	7	36	1	0.2	66	0	0	96	0	0	126	0	0	156	0	0
7	3	5.03	37	0	0.1	67	0	0	97	0	0	127	0	0	157	0	0
8	2	3.78	38	0	0.1	68	0	0	98	0	0	128	0	0	158	0	0
9	3	2.94	39	0	0.1	69	0	0	99	0	0	129	0	0	159	0	0
10	3	2.34	40	0	0.1	70	0	0	100	0	0	130	0	0	160	0	0
11	2	1.91	41	0	0.1	71	0	0	101	0	0	131	0	0	161	0	0
12	1	1.59	42	0	0.1	72	0	0	102	0	0	132	0	0	162	0	0
13	3	1.34	43	0	0.1	73	0	0	103	0	0	133	0	0	163	0	0
14	0	1.14	44	0	0.1	74	0	0	104	0	0	134	0	0	164	0	0
15	2	0.98	45	2	0.1	75	0	0	105	0	0	135	0	0	165	0	0
16	3	0.86	46	0	0.1	76	0	0	106	0	0	136	0	0	166	0	0
17	0	0.75	47	0	0.1	77	0	0	107	0	0	137	0	0	167	0	0
18	0	0.67	48	0	0.1	78	0	0	108	0	0	138	0	0	168	0	0
19	1	0.59	49	0	0.1	79	0	0	109	0	0	139	0	0	169	0	0
20	0	0.53	50	0	0.1	80	0	0	110	0	0	140	0	0	170	0	0
21	0	0.48	51	0	0.1	81	0	0	111	0	0	141	0	0	171	0	0
22	0	0.43	52	0	0.1	82	0	0	112	0	0	142	0	0	172	0	0
23	0	0.39	53	0	0.1	83	0	0	113	0	0	143	0	0	173	0	0
24	0	0.36	54	0	0.1	84	0	0	114	0	0	144	0	0	174	0	0
25	0	0.33	55	0	0.1	85	0	0	115	0	0	145	0	0	175	0	0
26	0	0.3	56	0	0.1	86	0	0	116	0	0	146	0	0	176	0	0
27	2	0.28	57	0	0.1	87	0	0	117	0	0	147	0	0	177	0	0
28	0	0.26	58	0	0.1	88	0	0	118	0	0	148	0	0	178	0	0
29	0	0.24	59	0	0.1	89	0	0	119	0	0	149	0	0	179	1	0
30	0	0.22	60	0	0.1	90	1	0	120	0	0	150	0	0			
	a = 2.14			**R = 179**			**X^2 = 12.49**				**DF = 21**			**P(X^2) = 0.92**			

Table 9: John. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	146	151	32	0	0.3	63	0	0.1	94	0	0	125	0	0	156	0	0	187	0	0
2	47	42.8	33	0	0.3	64	0	0.1	95	0	0	126	0	0	157	0	0	188	0	0
3	19	20.5	34	2	0.3	65	0	0.1	96	0	0	127	0	0	158	0	0	189	0	0
4	12	12.2	35	0	0.2	66	0	0.1	97	0	0	128	0	0	159	0	0	190	0	0
5	7	8.11	36	0	0.2	67	0	0.1	98	0	0	129	0	0	160	0	0	191	0	0
6	6	5.82	37	0	0.2	68	0	0.1	99	0	0	130	0	0	161	0	0	192	0	0
7	8	4.4	38	0	0.2	69	0	0.1	100	0	0	131	0	0	162	0	0	193	0	0
8	2	3.45	39	0	0.2	70	0	0.1	101	0	0	132	0	0	163	0	0	194	0	0
9	3	2.79	40	0	0.2	71	1	0.1	102	0	0	133	0	0	164	0	0	195	0	0
10	2	2.3	41	0	0.2	72	1	0.1	103	0	0	134	0	0	165	0	0	196	0	0
11	4	1.94	42	0	0.2	73	0	0.1	104	0	0	135	0	0	166	0	0	197	0	0
12	1	1.65	43	0	0.2	74	0	0.1	105	0	0	136	0	0	167	0	0	198	0	0
13	2	1.43	44	0	0.2	75	0	0.1	106	0	0	137	0	0	168	0	0	199	0	0
14	2	1.25	45	0	0.2	76	0	0.1	107	0	0	138	0	0	169	0	0	200	0	0
15	2	1.1	46	0	0.1	77	0	0.1	108	0	0	139	0	0	170	0	0	201	0	0
16	2	0.98	47	0	0.1	78	0	0.1	109	0	0	140	0	0	171	0	0	202	0	0
17	1	0.88	48	0	0.1	79	0	0.1	110	0	0	141	0	0	172	0	0	203	0	0
18	1	0.79	49	0	0.1	80	0	0.1	111	0	0	142	0	0	173	0	0	204	0	0
19	1	0.72	50	0	0.1	81	0	0.1	112	0	0	143	0	0	174	0	0	205	0	0
20	0	0.65	51	0	0.1	82	0	0.1	113	0	0	144	0	0	175	0	0	206	0	0
21	0	0.6	52	0	0.1	83	0	0.1	114	0	0	145	0	0	176	0	0	207	0	0
22	0	0.55	53	0	0.1	84	0	0.1	115	0	0	146	0	0	177	0	0	208	0	0
23	1	0.51	54	0	0.1	85	0	0.1	116	0	0	147	0	0	178	0	0	209	0	0
24	1	0.47	55	1	0.1	86	0	0.1	117	0	0	148	0	0	179	0	0	210	0	0
25	1	0.44	56	0	0.1	87	0	0.1	118	0	0	149	0	0	180	0	0	211	0	0
26	1	0.41	57	0	0.1	88	0	0	119	0	0	150	0	0	181	0	0	212	0	0
27	0	0.38	58	0	0.1	89	0	0	120	0	0	151	0	0	182	0	0	213	1	0
28	0	0.35	59	0	0.1	90	0	0	121	0	0	152	0	0	183	0	0			
29	0	0.33	60	0	0.1	91	0	0	122	0	0	153	0	0	184	0	0			
30	0	0.31	61	0	0.1	92	0	0	123	0	0	154	0	0	185	0	0			
31	0	0.29	62	0	0.1	93	0	0	124	0	0	155	0	0	186	0	0			
	a = 1.82			**R = 213**			**X^2 = 17.98**				**DF = 26**			**P(X^2) = 0.88**						

Table 10: John. Secondary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	166	156	26	1	0.4	51	0	0.1	76	0	0.1	101	0	0	126	0	0	151	0	0
2	41	43.1	27	0	0.4	52	0	0.1	77	0	0.1	102	0	0	127	0	0	152	0	0
3	15	20.3	28	0	0.3	53	0	0.1	78	0	0.1	103	0	0	128	0	0	153	0	0
4	16	11.9	29	0	0.3	54	0	0.1	79	0	0.1	104	0	0	129	0	0	154	0	0
5	5	7.88	30	0	0.3	55	0	0.1	80	0	0.1	105	0	0	130	0	0	155	0	0
6	4	5.62	31	1	0.3	56	0	0.1	81	0	0.1	106	0	0	131	0	0	156	0	0
7	3	4.23	32	0	0.3	57	0	0.1	82	1	0	107	0	0	132	0	0	157	0	0
8	5	3.3	33	0	0.2	58	0	0.1	83	0	0	108	0	0	133	0	0	158	0	0
9	2	2.65	34	0	0.2	59	0	0.1	84	0	0	109	0	0	134	0	0	159	0	0
10	2	2.18	35	1	0.2	60	0	0.1	85	0	0	110	0	0	135	0	0	160	0	0
11	1	1.83	36	1	0.2	61	1	0.1	86	0	0	111	0	0	136	0	0	161	0	0
12	1	1.56	37	1	0.2	62	0	0.1	87	0	0	112	0	0	137	0	0	162	0	0
13	0	1.34	38	0	0.2	63	0	0.1	88	0	0	113	0	0	138	0	0	163	0	0
14	0	1.17	39	1	0.2	64	0	0.1	89	0	0	114	0	0	139	0	0	164	1	0
15	0	1.03	40	0	0.2	65	0	0.1	90	0	0	115	0	0	140	0	0			
16	2	0.91	41	0	0.2	66	0	0.1	91	0	0	116	0	0	141	0	0			
17	3	0.82	42	1	0.2	67	0	0.1	92	0	0	117	0	0	142	0	0			
18	0	0.73	43	0	0.2	68	0	0.1	93	0	0	118	0	0	143	0	0			
19	0	0.66	44	0	0.1	69	1	0.1	94	0	0	119	0	0	144	0	0			
20	1	0.6	45	0	0.1	70	0	0.1	95	0	0	120	0	0	145	0	0			
21	0	0.55	46	0	0.1	71	0	0.1	96	0	0	121	0	0	146	0	0			
22	0	0.51	47	0	0.1	72	0	0.1	97	0	0	122	0	0	147	0	0			
23	1	0.47	48	0	0.1	73	0	0.1	98	0	0	123	0	0	148	0	0			
24	0	0.43	49	0	0.1	74	0	0.1	99	0	0	124	0	0	149	0	0			
25	0	0.4	50	0	0.1	75	0	0.1	100	0	0	125	0	0	150	0	0			
		a = 1.85			**R = 164**			**X^2 = 17.98**				**DF = 26**			**P(X^2) = 0.88**					

Table 11: Apocalypse. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	165	163	13	0	1.5	25	1	0.5	37	0	0.2	49	0	0.1	61	1	0.1	73	0	0.1
2	46	45.6	14	3	1.3	26	1	0.4	38	0	0.2	50	0	0.1	62	0	0.1	74	0	0.1
3	23	21.7	15	0	1.1	27	0	0.4	39	0	0.2	51	0	0.1	63	0	0.1	75	1	0.1
4	9	12.8	16	1	1	28	0	0.4	40	0	0.2	52	0	0.1	64	0	0.1			
5	8	8.51	17	1	0.9	29	0	0.3	41	0	0.2	53	1	0.1	65	0	0.1			
6	7	6.1	18	0	0.8	30	0	0.3	42	0	0.2	54	0	0.1	66	1	0.1			
7	3	4.6	19	1	0.7	31	0	0.3	43	0	0.2	55	0	0.1	67	0	0.1			
8	3	3.6	20	0	0.7	32	0	0.3	44	0	0.2	56	0	0.1	68	0	0.1			
9	4	2.9	21	2	0.6	33	0	0.3	45	0	0.2	57	0	0.1	69	0	0.1			
10	5	2.39	22	1	0.6	34	1	0.3	46	0	0.2	58	0	0.1	70	0	0.1			
11	2	2.01	23	0	0.5	35	0	0.2	47	0	0.1	59	0	0.1	71	0	0.1			
12	0	1.71	24	2	0.5	36	0	0.2	48	0	0.1	60	0	0.1	72	0	0.1			
		a = 1.83			**R = 75**			**X^2 = 21.41**			**DF = 23**			**P(X^2) = 0.56**						

Table 12: Apocalypse. Secondary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	118	116.6	11	0	0.58	21	0	0.14	31	0	0.06	41	0	0.03	51	0	0.02
2	17	25.12	12	2	0.48	22	0	0.12	32	0	0.05	42	0	0.03	52	0	0.02
3	10	10.24	13	1	0.4	23	0	0.11	33	1	0.05	43	0	0.03	53	0	0.02
4	5	5.41	14	3	0.34	24	0	0.1	34	0	0.05	44	0	0.03	54	0	0.02
5	1	3.3	15	0	0.29	25	0	0.09	35	0	0.04	45	0	0.03	55	0	0.02
6	5	2.21	16	0	0.25	26	0	0.09	36	0	0.04	46	0	0.02	56	1	0.02
7	2	1.57	17	0	0.22	27	0	0.08	37	0	0.04	47	0	0.02			
8	2	1.17	18	1	0.19	28	0	0.07	38	0	0.04	48	0	0.02			
9	1	0.9	19	0	0.17	29	0	0.07	39	0	0.04	49	0	0.02			
10	1	0.71	20	1	0.15	30	0	0.06	40	0	0.03	50	0	0.02			
	a = 2.21			**R = 56**			X^2 = **14.94**			**DF = 10**				**P(X^2) = 0.13**			

Table 13: Romans. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)
1	113	113	21	0	0.3	41	0	0.1	61	0	0	81	0	0	101	0
2	36	28.8	22	0	0.3	42	0	0.1	62	0	0	82	0	0	102	0
3	11	12.9	23	0	0.2	43	0	0.1	63	0	0	83	0	0	103	0
4	8	7.33	24	0	0.2	44	0	0.1	64	0	0	84	0	0	104	0
5	6	4.72	25	0	0.2	45	0	0.1	65	0	0	85	0	0	105	0
6	4	3.29	26	2	0.2	46	0	0.1	66	0	0	86	0	0	106	0
7	1	2.43	27	0	0.2	47	0	0.1	67	0	0	87	0	0	107	0
8	2	1.86	28	0	0.2	48	0	0.1	68	0	0	88	0	0	108	0
9	0	1.48	29	0	0.2	49	0	0.1	69	0	0	89	0	0	109	0
10	0	1.2	30	0	0.1	50	0	0.1	70	0	0	90	0	0	110	0
11	0	0.99	31	0	0.1	51	0	0.1	71	0	0	91	0	0	111	0
12	1	0.84	32	0	0.1	52	0	0.1	72	0	0	92	0	0	112	0
13	1	0.71	33	0	0.1	53	0	0	73	0	0	93	0	0	113	0
14	0	0.62	34	0	0.1	54	0	0	74	0	0	94	0	0	114	0
15	0	0.54	35	0	0.1	55	0	0	75	0	0	95	0	0	115	0
16	0	0.47	36	0	0.1	56	0	0	76	0	0	96	0	0	116	1
17	0	0.42	37	0	0.1	57	0	0	77	0	0	97	0	0		
18	1	0.38	38	0	0.1	58	1	0	78	0	0	98	0	0		
19	0	0.34	39	0	0.1	59	0	0	79	0	0	99	0	0		
20	0	0.3	40	0	0.1	60	0	0	80	0	0	100	0	0		
	a = 1.98			**R = 116**			**X^2 = 10.25**				**DF = 16**			**P(X^2) = 0.85**		

Table 14: Romans. Secondary Process

x	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**	**x**	**g(x)**	**NPx**
1	197	190.9	13	2	0.71	25	0	0.17	37	1	0.07	49	0	0.04	61	0	0.02	73	0	0
2	40	42.13	14	0	0.61	26	0	0.16	38	0	0.07	50	0	0.04	62	0	0.02	74	0	0
3	11	17.41	15	2	0.52	27	1	0.14	39	0	0.07	51	0	0.04	63	0	0.02	75	1	0
4	9	9.3	16	2	0.45	28	0	0.13	40	0	0.06	52	0	0.03	64	0	0.02			
5	4	5.72	17	0	0.4	29	0	0.12	41	0	0.06	53	0	0.03	65	0	0.02			
6	4	3.84	18	1	0.35	30	0	0.12	42	0	0.06	54	0	0.03	66	0	0.02			
7	3	2.75	19	0	0.31	31	0	0.11	43	0	0.05	55	0	0.03	67	0	0.02			
8	4	2.05	20	0	0.28	32	0	0.1	44	0	0.05	56	0	0.03	68	0	0.02			
9	0	1.59	21	0	0.25	33	0	0.09	45	0	0.05	57	0	0.03	69	0	0.02			
10	0	1.26	22	1	0.23	34	1	0.09	46	0	0.05	58	0	0.03	70	0	0.02			
11	0	1.03	23	0	0.21	35	0	0.08	47	0	0.04	59	0	0.03	71	0	0.02			
12	2	0.85	24	0	0.19	36	0	0.08	48	0	0.04	60	0	0.03	72	0	0.02			
	a = 2.18			**R = 75**			**X^2 = 14.54**			**DF = 14**			**P(X^2) = 0.41**							

Table 15: Herculano. Primary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	188	183	13	1	1	25	0	0.3	37	0	0.1	49	0	0.1	61	0	0.04	73	0	0
2	37	44.2	14	0	0.8	26	0	0.2	38	0	0.1	50	0	0.1	62	0	0.04	74	0	0
3	23	19.3	15	1	0.7	27	0	0.2	39	0	0.1	51	0	0.1	63	0	0.04	75	0	0
4	11	10.7	16	0	0.6	28	0	0.2	40	0	0.1	52	0	0.1	64	1	0.04	76	0	0
5	9	6.77	17	0	0.6	29	0	0.2	41	0	0.1	53	0	0.1	65	0	0.04	77	0	0
6	3	4.66	18	0	0.5	30	0	0.2	42	0	0.1	54	0	0.1	66	0	0.03	78	0	0
7	3	3.4	19	0	0.4	31	0	0.2	43	0	0.1	55	0	0.1	67	0	0.03	79	0	0
8	1	2.59	20	1	0.4	32	0	0.2	44	0	0.1	56	0	0.1	68	0	0.03	80	1	0
9	6	2.03	21	1	0.4	33	0	0.1	45	0	0.1	57	0	0.1	69	0	0.03			
10	0	1.64	22	0	0.3	34	0	0.1	46	0	0.1	58	0	0	70	0	0.03			
11	3	1.35	23	0	0.3	35	0	0.1	47	0	0.1	59	0	0	71	0	0.03			
12	1	1.13	24	0	0.3	36	0	0.1	48	0	0.1	60	0	0	72	0	0.03			
	a = 2.05			**R = 80**			**X^2 = 21.49**				**DF = 17**			**P(X^2) = 0.21**						

Table 16: Herculano. Secondary Process

x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx	x	g(x)	NPx
1	81	75.2	6	1	1.9	11	2	0.5	16	0	0.3	21	2	0.1	26	0	0.09	31	0	0.1
2	13	18	7	2	1.4	12	1	0.5	17	0	0.2	22	0	0.1	27	0	0.08	32	0	0.1
3	6	7.81	8	0	1	13	0	0.4	18	0	0.2	23	1	0.1	28	0	0.08	33	0	0.1
4	2	4.32	9	2	0.8	14	1	0.3	19	0	0.2	24	0	0.1	29	0	0.07	34	0	0.1
5	2	2.73	10	0	0.7	15	0	0.3	20	1	0.2	25	0	0.1	30	0	0.07	35	1	0.1
	a = 2.06			**R = 35**			**X^2 = 8.93**				**DF = 9**			**P(X^2) = 0.44**						

4 DISCUSSION

The analysis has shown that the right-truncated Zeta distribution can be fitted satisfactorily to word-frequency spectra within the categories of primary and secondary process, as identified in texts using the Portuguese translation of the Regressive Imagery Dictionary. Acceptable fits were found for all six biblical books examined in the study, as well as for the romance by Herculano, with the majority showing a very good quality of fit. This would seem, therefore, to provide further support for the validity of the main categories within the RID. As well as showing good construct validity in previous studies, the categories have also now demonstrated formal quantitative behaviour that shows them to be correctly delimited entities in texts.

In all cases, the parameter a of the right-truncated Zeta distribution approximated towards a value of 2 (mean = 2.00, min = 1.81, max = 2.21, sd = 0.13). However, in five out of the seven texts, the value of a for the secondary process category was slightly larger than its value for the primary process category.[2] On the basis of these few preliminary data alone, it is difficult to hazard any very strong claims as to why the size of a should tend to be slightly higher for the secondary process category. It certainly does not seem to be related to the dominance of one or the other category within a text, since the Epistle to the Romans, which has a significant dominance of secondary process words, shows exactly the same pattern of a-values as texts such as the Apocalypse, which has a primary process dominance (cf. Table 2). It seems more likely that the relative size of a may be related to the *a priori* probability of occurrence of certain words that are contained within the pre-defined dictionary categories, since, in all cases, the value of R – i.e. the frequency of the most common word in the category – was higher for primary process than for secondary process: in other words, the most frequent primary process word always occurred more often than the most frequent secondary process word, even in a text with an overall dominance of secondary process words. However, this relationship requires further explicit investigation on a larger number of texts.

---

[2] In the case of two other texts – the Gospel according to John and the romance by Herculano – the values of a for primary and secondary process were almost equal.

Further research should also aim to extend and consolidate these findings in a number of other ways. First, the Portuguese RID should be applied to a larger sample of texts, including other genres, in order to confirm the model suggested here. In such a study, any further systematic patterns of variation in the parameter *a* should be examined, as the parameters of frequency distributions have previously shown clear evidence of links to text typology and authorship. Second, the experiments should be replicated using other versions of the RID, in order to check that the findings of the present study are generally valid and not an artefact of the Portuguese version alone. For example, Hogenraad (2005) has investigated the behaviour of several versions of the RID on parallel translation texts and has found that the size of the correlations between category counts produced by different versions does vary. Third, the research should be extended to cover other kinds of psychoanalytic categories in texts – for example, the categories of body boundary definiteness developed in Wilson (2006) and the categories of anality and orality constructed by Vanheule, Desmet & Meganck (2008). Finally, the most challenging task would be to integrate these empirical findings into a synergetic theory of text production that incorporates cognitive elements such as consciousness states and concept-word mappings. Researchers such as Roy (2004) and Spivak (2004) have already made interesting contributions in this direction, but there is still much work that remains to be done in developing a comprehensive model.

REFERENCES

Altmann, G. (2002): Zipfian linguistics, in: *Glottometrics,* 3, 19-26.

Aulagnier, P. (2001): *The violence of interpretation: From pictogram to statement* (A. Sheridan, Trans.). London: Brunner-Routledge. (Original work published 1975.)

Baayen, R.H. (2001): *Word frequency distributions*. Dordrecht: Kluwer.

Bucci, W. (1997): *Psychoanalysis and cognitive science: A multiple code theory*. New York: Guilford Press.

Elter-Nodvin, E. (2000): Computerized content analysis: a comparison of the verbal productions of high hypnotizable, low hypnotizable and simulating subjects. Ph.D. dissertation, University of Tennessee, Knoxville.

Goldstein, K. (1939): *The organism*. Boston: Beacon.

Freud, S. (1900): *Die Traumdeutung*. Leipzig & Vienna: Franz Deuticke.

Hogenraad, R. (2005): The Regressive Imagery Dictionary: a test of five versions (English, French, German, Portuguese, and Swedish). Paper presented at the International Congress on Aesthetics, Creativity, and Psychology of the Arts, Perm, Russia, June 2005.

Hogenraad, R.; Daubies, C.; Bestgen, Y. & Mahau, P. (2003): Une théorie et une méthode générale d'analyse textuelle assistée par ordinateur. Le système PROTAN (PROTocol ANalyzer). 32-bits version of November 10, 2003 by Pierre Mahau. Psychology Department, Catholic University of Louvain, Louvain-la-Neuve.

Kristeva, J. (1996/1997): Freudian models of language: A conversation. In: *Psychomedia: Journal of European Psychoanalysis, 3/4.* [Retrieved December 8, 2006 from [http://www.psychomedia.it/jep/number3-4/kristeng.htm].

Lévy-Bruhl, L. (1966): *How natives think.* New York: Washington Square Press. (Original work published 1910.)

Martindale, C. (1975): *Romantic progression: the psychology of literary history.* Washington, DC: Hemisphere.

Martindale, C. (1976): Primitive mentality and the relationship between art and society, in: *Scientific Aesthetics,* 1, 5-18.

Martindale, C. & Fischer, R. (1977): The effects of psilocybin on primary process content in language. In: *Confinia Psychiatrica, 20,* 195-202.

Piaget, J. (1954): *The construction of reality in the child.* New York: Basic Books.

Popescu, I.-I.; Best, K.-H. & Altmann, G. (2007): On the dynamics of word classes in text, in: *Glottometrics,* 14, 58-71.

Prün, C. (1999): G.K. Zipf's conception of language as an early prototype of synergetic linguistics, in: *Journal of Quantitative Linguistics,* 6, 78-84.

Roy, P.K. (2004): Stochastic resonance as an emerging technique for neuron-modulation and pharmacolinguistics: using nonlinear dynamics to analyze drug-induced language transition and EEG, in: *Journal of Quantitative Linguistics,* 11, 49-77.

Sorokin, P. (1957): *Social and cultural dynamics: A study of change in major systems of art, truth, ethics, law and social relationships.* Boston, MA: Porter Sargent.

Spivak, D. (2004): Linguistics of altered states of consciousness: problems and prospects, in: *Journal of Quantitative Linguistics,* 11, 27-32.

Strauss, U.; Fan, F. & Altmann, G. (2008): *Problems in quantitative linguistics 1* (2nd ed.). Lüdenscheid: RAM-Verlag.

Uhlířová, L. (1995): On the generality of statistical laws and individuality of texts. A case of syllables, word forms, their length and frequencies, in: *Journal of Quantitative Linguistics,* 2, 238-247.

Vanheule, S., Desmet, M. & Meganck, R. (2008): Anal and oral word use in relation to dependency and self-criticism. Poster presentation at the Winter Meeting of the American Psychoanalytic Association, New York, 2008.

Werner, H. (1948): *Comparative psychology of mental development.* New York: International Universities Press.

West, A. (1991): Primary process content in the King James Bible: the five stages of Christian mysticism, in: *Computers and the Humanities,* 25, 227-238.

West, A. & Martindale, C. (1988): Primary process content in paranoid schizophrenic speech, in: *Journal of Genetic Psychology,* 149, 547-553.

West, A.; Martindale, C.; Hines, D. & Roth, W. (1983): Marijuana-induced primary process content in the TAT, in: *Journal of Personality Assessment,* 47, 466-467.

West, A.; Martindale, C. & Sutton-Smith, B. (1985): Age trends in the content of children's spontaneous fantasy narratives, in: *Genetic, Social, and General Psychology Monographs,* 111, 391-405.

Wilson, A. (2002): The application of computer content analysis in sexology: a case study of primary process content in fictional fetishistic narratives, in: *Electronic Journal of Human Sexuality,* 5. [Retrieved October 31, 2006, from: http://www.ejhs.org/volume5/wilson.html].

Wilson, A. (2006): Development and application of a content analysis dictionary for body boundary research, in: *Literary and Linguistic Computing,* 21, 105-110.

Wilson, A. (2007): Barrier and penetration imagery as a supplementary measure of altered states of consciousness discourse: replicating the five-stage model of Christian mysticism in the Bible. Forthcoming.

Ziegler, A., Best, K.-H. & Altmann, G. (2002): Nominalstil, in: *Empirical Text and Culture Research,* 2, 72-85.

Zipf, G.K. (1935): *The psycho-biology of language: an introduction to dynamic philology.* Boston: Houghton Mifflin.