

The Dutch translation of the Linguistic Inquiry and Word Count (LIWC) 2007 dictionary

Peter Boot, Hanna Zijlstra, Rinie Geenen

Appeared *Dutch Journal of Applied Linguistics* 6 (2018) 1, pp. 65-76

<http://www.ibe-platform.com/content/journals/10.1075/dujal.6.1.04boo>

Abstract

The words we use in everyday language reveal our thoughts, feelings, personality, and motivations. Linguistic Inquiry and Word Count (LIWC) is a software program to analyse text by counting words in certain psychologically meaningful categories that are catalogued in a dictionary of words. Several versions of the program and dictionary have been released and translated into a number of languages. The 2001 version has been translated into Dutch. This article presents the Dutch translation of LIWC 2007. It describes the procedure of translation and it compares the dictionaries on a corpus of parallel (Dutch-English) texts. The Dutch and English dictionaries give similar results on a collection of translated texts in both languages, except for a small number of categories. Correlations are high to very high, effect sizes of the differences between word counts are low to medium. The LIWC 2007 categories can now be used to analyse Dutch language texts.

Keywords: Linguistic Inquiry and Word Count; LIWC; text analysis; psychology; tools; translation

1. Introduction

Language, whether spoken or written, is an important window into people's thoughts, feelings, personality, and motivations. Text analysis is a means to understand these psychological processes. Linguistic Inquiry and Word Count (LIWC) is a software program for counting words in certain psychologically meaningful categories. It was initially developed for the analysis of traumatized patients' narratives (Pennebaker & Graybeal, 2001), but has been applied in many contexts and for many purposes. Tausczik & Pennebaker (2010) reviewed investigations into addiction, suicidal tendencies, online behaviour, deception, bereavement, and other psychological phenomena. Meaningful results were obtained in many contexts. It was for instance shown that the use of personal pronouns referring to self is high in poets that commit suicide, and that the phases of traumatic recovery after the September 11th disaster are reflected in word use on the internet. Moreover, LIWC is an important tool in researching the effectiveness of communication (Scielzo, Patel, & Smith-Jentsch, 2011) and is used in a commercial setting to analyse mail from staff members, customers or job applicants (Receptiviti, 2016).

LIWC comes with a dictionary of words in 66 categories. This dictionary has been translated in many languages. The 2001 version has been translated into Dutch (Zijlstra, Van Meerveld, Van Middendorp, Pennebaker, & Geenen, 2004; Zijlstra, Van Middendorp, Van Meerveld, & Geenen, 2005). The current research note presents and evaluates the Dutch translation of the LIWC 2007 dictionary (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007).

The LIWC program counts, in one or more texts, the number of words in each of the categories. These word categories include syntactic categories (e.g. pronouns), psychological processes

(affective, cognitive, perceptual and biological) and personal concerns, such as work and death. The word categories are shown in Table 1 and a full list of categories and sample words can be found in Pennebaker et al. (2007). Texts can be transcriptions of conversations, naturally occurring texts (school assignments, email), internet texts such as forum posts, or any other type of text. The LIWC output consists of a file that gives, for each input document, the percentage of words in each category in the document. The output file can be manually inspected or read using Excel, SPSS, R and many other programs. The LIWC word counts are necessarily crude (as they are independent of context), and they do not aim at the subtlety of e.g. a (manually coded) conversation analysis. LIWC is however remarkably fast and efficient in a quantitative research setting, it can analyse huge amounts of text, it is independent of human judgment and it gives replicable results.

Changes in the 2007 dictionary include the removal of little-used categories, a reorganisation of the category Body and a merging of the Job and School categories. Many words were added to, or sometimes removed from, the individual categories. However, the most important innovation of the 2007 LIWC dictionary is the increased coverage of function words such as conjunctions, frequent adverbs and auxiliary verb. As Chung and Pennebaker (2007) argue, function words (pronouns, articles, prepositions, auxiliary verbs, etc.) have important social functions and they tell us something about linguistic style. Function words usage reflects gender, political views and mental states (Pennebaker, 2011).

An important reason for these changes in the LIWC dictionary was the interest in 'linguistic style matching' (LSM), linguistic accommodation between participants in a conversation. For instance, during speed dating, language style matching of function words such as 'the', 'not', 'all', 'none' and 'me' predicted the initiation of romantic relationships and relationship stability (Ireland et al., 2011). LSM also predicted cohesiveness and task performance in groups (Gonzales, Hancock, & Pennebaker, 2010). The amount of accommodation has been shown to depend on power differences between participants (Danescu-Niculescu-Mizil, Lee, Pang, & Kleinberg, 2012). LSM plays an important role in situations as different as police interrogations and therapeutic conversations (Lord, Sheng, Imel, Baer, & Atkins, 2015; Richardson, Taylor, Snook, Conchie, & Bennell, 2014). The pervasiveness of linguistic accommodation is an important reason why a translation of LIWC 2007 into Dutch is desirable.

The existing translation of LIWC 2001 has been used in a psychological context for research into e.g. alexithymia (Meganck, Vanheule, Inslegers, & Desmet, 2009), depression (Balsters, 2013; Van der Zanden et al., 2014) and to evaluate an emotional disclosure intervention (Van Middendorp & Geenen, 2008). LIWC's range of application goes much beyond the realm of psychology: the translation has been used in age prediction on Twitter (Nguyen, Gravel, Trieschnigg, & Meder, 2013), to study online book discussions (Boot, 2013) and to examine customer attrition in a business context (Coussement & Van den Poel, 2009). The width of application shows that LIWC is an essential tool for the Dutch research context.

In this article, we describe the procedure of translation of LIWC 2007 in Section 2. Section 3 describes how we tested the equivalence of the Dutch and English dictionaries. Section 4 is the conclusion.

2. The procedure of translation

For most categories of the LIWC 2007 dictionary, we translated just those words that were added to the 2001 dictionary. For some heavily reorganised categories (such as Body and Health) we opted for a completely new translation. For those categories of function words that can be limitatively enumerated (articles, pronouns, auxiliary verbs) we chose to build the list based on the Dutch reference grammar E-ANS (E-ANS, 2012). The translations were done by one person and reviewed by two. Unlike the 2001 translation, the present translation frequently uses the '*' wildcard: *deur** will catch *deur*, *deuren* and *deurtje* (door, doors, small door).

The most common problem in translating the dictionary was homonymy (words with the same spelling having a different meaning). An ankle is a body part, but its Dutch translation *enkel* also means 'just', 'only', which is an adverb. Any of the possibilities (Body, Adverb, both or none) implies that some words will be counted wrongly or not at all. To choose the most appropriate category (the one that covers most instances), one needs information about the relative frequency of these word senses. For these frequency estimates we used a large story corpus (Boot, 2011), the OpenSoNar corpus (Oostdijk, Reynaert, Hoste, & Schuurman, 2013) and frequency information from e-Lex (Nederlandse Taalunie, 2014).

The resulting Dutch LIWC 2007 dictionary contains 11091 entries/words in 66 categories. The categories and their hierarchical structure correspond to those of the English dictionary, except for two additional categories: one for singular and plural third person pronouns (Shehethey) and one for pronominal adverbs, such as *waarmee* (wherewith) or *daarin* (therein). For a list of categories, including examples, see Pennebaker et al. (2007).

3. Equivalence of dictionary and translation

3.1 Method

To test the equivalence of the English dictionary and its Dutch translation, we collected 564 texts that were available in Dutch and English. The texts included a collection of letters by Vincent van Gogh (source: <http://vangoghletters.org>), the books of the Bible (source: <http://iw.org>), and three collections described in Tiedeman (2012): TED-talks, transcripts of sessions of the European Parliament and a collection of EU official texts (the so-called 'acquis communautaire'). The selection was based on the desire for a wide range of text types (personal, expository, formal) as well as pragmatic reasons (availability). We removed texts that contained less than a thousand words or where the dictionary coverage was below 60% (either in Dutch or in English).

We applied the same procedure as was applied for the translation of LIWC 2001 (Zijlstra et al., 2004). The English and Dutch texts were analysed using the English and Dutch LIWC 2007 dictionaries. For each of the word categories, the correlation was computed between the word counts in the Dutch and English texts (so n=564). If the skewness of the score distribution was between -1.5 and 1.5, we computed the Pearson correlation, if otherwise we used Spearman's rank correlation coefficient. We also computed the median, minimum and maximum frequencies in all categories, as well as the effect sizes. For the effect size, we took the difference of the means of the percentages of the word counts for the English and Dutch texts and divided this by the square root of the mean of their variances. Conventionally, an effect size of .2 is considered small, one of .5 is considered medium, and an effect size of .8 is considered large (Cohen, 1988).

We aimed for high ($> .50$) correlations between word counts within the corresponding Dutch and English categories, because they imply that effects that have been found in research on English texts can probably be transferred to the Dutch situation, and vice-versa. This form of equivalence testing was also applied for many other LIWC translations (Bjekić, Lazarević, Živanović, & Knežević, 2014; Ramirez-Esparza, Pennebaker, García, & Suriá, 2007; Wolf et al., 2008). Additionally, we computed the effect size. Here we aim for low to medium effect sizes. A high effect size implies a large difference between the means of the computed frequencies of word counts, which in the case of a perfect translation would be unlikely.

3.2 Results

Table 1 reports the results of the comparison between the two dictionaries. For 60 out of 64 categories, the correlations are high (above $.50$) or very high (above $.70$). The correlations are lower for four categories: Future, Swear words, Nonfluencies and Fillers. With regards to the effect sizes, 55 categories have an effect size below $.65$ (i.e. half way between the conventional medium and large effect sizes). The effect sizes are higher for Articles, Prepositions, Conjunctions, Tentative, Exclusive, Relative, Space, Nonfluencies and Fillers. For the psychologically very relevant categories Pronouns, Social, Affect and Cognitive Mechanisms the effect size is small or medium, as desired.

For some of the categories where the results could have been better (e.g. Future), the cause is probably a difference in word use between the languages. Other cases (e.g. Conjunctions) may be caused by homonymies in Dutch or in English. For the Nonfluencies and Fillers our test corpus may have been less suitable.

Table 1. Results of equivalence test.

Category	Word counts				Equivalence statistics	
	English		Dutch		<i>r</i>	<i>d</i>
	<i>Median</i>	<i>Min - Max</i>	<i>Median</i>	<i>Min - Max</i>		
General information						
Word count	3,128.50	1,021 - 44,512	3,053.00	1,021 - 48,010	1.00	0.00
Words per sentence	24.45	9.32 - 117.52	24.36	9.21 - 288	0.78 ^a	0.19
Words \geq 6 letters	20.46	8.03 - 36.83	24.56	11.81 - 39.99	0.96	0.46
Dictionary coverage	82.41	60.69 - 94.07	81.16	62.49 - 92.42	0.85	0.33
Linguistic dimensions						
Function words	57.09	42.23 - 71.21	57.40	44.08 - 67.53	0.91	0.01
Pronouns	14.63	2.41 - 26.53	12.15	2.35 - 25.05	0.97	0.39
Personal pronouns	7.66	0.28 - 16.97	7.78	0.41 - 20.86	0.98	0.02
I	1.67	0.02 - 10.68	1.67	0 - 11.26	1.00	0.02
We	0.92	0 - 5.96	0.87	0 - 5.87	0.99	0.02
You	1.34	0 - 7.91	1.33	0 - 15.27	0.98 ^a	0.02
Shehe	0.38	0 - 6.18	0.70	0.08 - 7.65	0.82 ^a	0.19
They	0.67	0 - 5.27	0.86	0.12 - 5.59	0.81 ^a	0.39

Impersonal pronouns	6.35	1.99 - 10.83	5.34	2.35 - 9.37	0.85	0.56
Article	7.93	3.77 - 13.44	9.61	5.02 - 16.45	0.93	0.86
Adverb	4.32	0.26 - 9.00	3.51	0.26 - 7.96	0.88	0.45
Prepositions	14.87	7.96 - 20.08	11.00	4.75 - 18.47	0.89	1.09
Conjunctions	6.91	3.59 - 12.1	5.38	3.03 - 10.79	0.78	0.88
Negations	1.21	0.16 - 4.44	1.27	0.14 - 4.97	0.95	0.03
Quantities	2.52	0.92 - 4.83	3.09	0.90 - 6.45	0.80	0.63
Numbers	1.48	0.28 - 11.08	1.38	0.05 - 11.28	0.61 ^a	0.07
Verbs	12.62	4.93 - 18.38	13.12	5.18 - 18.92	0.91	0.15
Auxiliary verbs	8.40	3.99 - 13.02	7.55	2.63 - 11.66	0.88	0.62
Past	2.11	0.28 - 7.24	1.97	0.60 - 7.94	0.89 ^a	0.02
Present	7.26	1.25 - 14.01	8.28	3.51 - 13.53	0.89	0.46
Future	1.56	0.08 - 5.18	1.12	0.11 - 4.69	0.47	0.61
Psychological processes						
Swear words	0	0 - 0.29	0	0 - 0.29	0.32 ^a	0.20
Social words	7.96	2.12 - 21.39	7.71	2.90 - 19.39	0.94	0.05
Family	0.06	0 - 4.76	0.11	0 - 4.77	0.87 ^a	0.11
Friend	0.06	0 - 0.54	0.05	0 - 0.47	0.68 ^a	0.15
Humans	0.76	0 - 3.9	0.80	0.06 - 3.46	0.83	0.06
Affect	4.19	1.01 - 9.15	3.68	0.86 - 8.87	0.84	0.39
Positive emotions	2.87	0.65 - 7.94	2.35	0.51 - 6.73	0.73	0.59
Negative emotions	1.15	0.04 - 4.38	1.08	0.16 - 3.95	0.88	0.05
Anxiety	0.2	0 - 2.99	0.16	0 - 2.67	0.79 ^a	0.10
Anger	0.26	0 - 2.26	0.25	0 - 1.93	0.79 ^a	0.00
Sadness	0.22	0 - 1.37	0.38	0 - 1.48	0.70	0.59
Cognitive mechanisms	16.69	9.56 - 23.17	17.91	11.42 - 25.51	0.88	0.44
Insight	2.2	0.26 - 4.81	2.43	0.33 - 4.84	0.72	0.15
Cause	1.77	0.55 - 5.27	1.49	0.28 - 3.97	0.76	0.53
Discrepancy	1.67	0.08 - 4.24	1.91	0.31 - 4.18	0.78	0.32
Tentative	2.20	0.24 - 5.63	2.90	0.60 - 6.89	0.86	0.69
Certain	1.53	0.42 - 3.47	1.56	0.43 - 3.28	0.57	0.04
Inhibition	0.54	0 - 2.06	0.35	0 - 2.01	0.86	0.60
Inclusion	5.22	2.25 - 10.80	5.10	2.64 - 12.19	0.67 ^a	0.09
Exclusion	2.24	0.51 - 6.00	2.99	0.80 - 7.60	0.90	0.79
Perception	1.64	0.04 - 11.44	1.95	0.05 - 9.80	0.94 ^a	0.17
See	0.56	0 - 9.78	0.55	0 - 8.27	0.91 ^a	0.03
Hear	0.44	0 - 4.79	0.53	0 - 4.89	0.93 ^a	0.14
Feel	0.25	0 - 3.53	0.30	0 - 1.26	0.74 ^a	0.06
Biological processes	0.88	0.10 - 6.20	0.78	0.10 - 6.38	0.90 ^a	0.17
Body	0.23	0 - 3.04	0.19	0 - 2.94	0.86 ^a	0.14
Health	0.32	0 - 5.09	0.24	0 - 3.92	0.72 ^a	0.19
Sexual	0.05	0 - 1.56	0.03	0 - 2.21	0.68 ^a	0.20
Ingest (food/drink/drugs)	0.14	0 - 4.04	0.13	0 - 3.62	0.84 ^a	0.00
Relative words	12.00	8.01 - 19.45	13.91	7.91 - 19.41	0.75	1.05
Motion	1.46	0.57 - 5.03	1.79	0.78 - 5.92	0.86	0.49

Space	6.40	3.72 - 10.10	8.43	3.63 - 11.99	0.72	1.65
Time	4.00	1.59 - 7.82	4.12	1.9 - 7.92	0.87	0.16
Personal Concerns						
Work	1.95	0.07 - 7.35	1.62	0.22 - 7.02	0.95	0.25
Achieve	1.93	0.30 - 4.82	1.66	0.36 - 4.54	0.86	0.42
Leisure	0.43	0 - 5.04	0.30	0 - 3.88	0.51 ^a	0.36
Home	0.17	0 - 2.07	0.08	0 - 2.25	0.69 ^a	0.37
Money	0.67	0 - 5.32	0.57	0 - 4.19	0.91 ^a	0.16
Religion	0.12	0 - 6.64	0.05	0 - 4.62	0.79 ^a	0.19
Death	0.06	0 - 1.44	0.05	0 - 0.94	0.77 ^a	0.19
Spoken language						
Assent	0.06	0 - 2.35	0.07	0 - 2.63	0.58 ^a	0.01
Nonfluencies	0.09	0 - 1.11	0	0 - 0.12	0.06 ^a	1.38
Fillers	0.10	0 - 1.42	0	0 - 0.16	0.12 ^a	1.30
Dutch only						
Shehethy			0.68	0.05 - 2.08		
Pronominal adverbs			1.27	0.27 - 9.17		

r: correlation, *d*: effect size (Cohen's *d*).

a: Correlations with superscript 'a' were computed using Spearman's rank correlation coefficient.

4. Conclusion

The analyses show high associations between texts processed with the original English and translated Dutch dictionaries of LIWC 2007. The findings indicate that except for a few word categories the English and Dutch LIWC give equivalent results. The analyses imply that research findings based on LIWC scores on English-language texts are relevant for Dutch-language subjects, and vice versa.

Automated text analysis is a research technique that is of growing importance in many research fields. LIWC has spearheaded development in the field of psychology, but is increasingly being used outside of its original scope, for instance in the blossoming field of digital humanities (Meder, Nguyen, & Gravel, 2015; Yu, 2014). A Dutch translation of the LIWC 2007 dictionary is now available. The Dutch and English dictionaries give similar results on a collection of translated texts in both languages, except for a small number of categories. Correlations are high to very high, effect sizes are low to medium. Research that uses the LIWC 2007 categories can now be done on Dutch language texts. A promising field of study is the research into linguistic style matching.

Since the development of the first LIWC dictionaries, a number of electronic resources have become available (dictionaries, corpora, grammars) that have helped us in building the translation. We wrote some Python scripts to support the translation process and these scripts are now available for translators of LIWC into other languages and for other LIWC users (Boot 2016). The next step is the Dutch translation of the 2015 version of the LIWC dictionary. Changes in that dictionary include new categories among others for references to men and women, for interrogative words, for drive, power and for netspeak. There are also a number of computed categories, such as analytic thinking and

authenticity. The development of this text analysis tool is an ongoing process that is fed by new theories and empirical findings.

Literature

- Balsters, M. (2013). *Expression and perception of emotions*. Tilburg: Tilburg University.
- Bjekić, J., Lazarević, L. B., Živanović, M., & Knežević, G. (2014). Psychometric evaluation of the Serbian dictionary for automatic text analysis LIWCser. *Psihologija*, 47(1).
- Boot, P. (2011). *Predicting long-term activity in online writing communities: A Quantitative Analysis of Amateur Writing*. Paper presented at the Supporting Digital Humanities 2011. Retrieved 2015-12-27 from http://crdo.up.univ-aix.fr/SLDRdata/doc/show/copenhagen/SDH-2011/submissions/sdh2011_submission_1.pdf
- Boot, P. (2013). *The desirability of a corpus of online book responses*. Paper presented at the Second Workshop on Computational Linguistics for Literature. Retrieved 2015-12-27 from <http://www.aclweb.org/anthology-new/W/W13/W13-1405.pdf>
- Boot, P. (2016). LIWCtools. Tools for working with LIWC dictionaries (Version 0.0.1). Retrieved 2016-09-02 from <https://github.com/pboot/LIWCtools>.
- Chung, C. K., & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedeler (Ed.), *Social communication* (pp. 343-359). New York: Psychology Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second ed.). Mahwah, New Jersey: Lawrence Erlbaum.
- Coussement, K., & Van den Poel, D. (2009). Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers. *Expert Systems with Applications*, 36, 6127-6134.
- Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). *Echoes of power: Language effects and power differences in social interaction*. Paper presented at the Proceedings of the 21st international conference on World Wide Web.
- E-ANS. (2012). Elektronische ANS. Version 1.3 Available from <http://ans.ruhosting.nl/e-ans/index.html>. 2015-12-11.
- Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language Style Matching as a Predictor of Social Dynamics in Small Groups. *Communication Research*, 37(1), 3.
- Ireland, M. E., & Pennebaker, J. W. (2010). Language Style Matching in Writing: Synchrony in Essays, Correspondence, and Poetry. *Journal of Personality and Social Psychology*, 99(3), 549–571.
- Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, 22(1), 39-44.
- Lord, S. P., Sheng, E., Imel, Z. E., Baer, J., & Atkins, D. C. (2015). More than reflections: Empathy in motivational interviewing includes language style synchrony between therapist and client. *Behavior therapy*, 46(3), 296-303.
- Meder, T., Nguyen, D., & Gravel, R. (2015). The Apocalypse on Twitter. *Digital Scholarship in the Humanities*, First published online: 31 January 2015. Retrieved from <http://dx.doi.org/10.1093/llc/fqu073>. doi:<http://dx.doi.org/10.1093/llc/fqu073>
- Meganck, R., Vanheule, S., Inslegers, R., & Desmet, M. (2009). Alexithymia and interpersonal problems: A study of natural language use. *Personality and Individual Differences*, 47(8), 990-995.
- Nederlandse Taalunie. (2014). e-Lex 1.1.1 Available from <http://tst-centrale.org/nl/producten/lexica/e-lex/7-25>. 21-11-2015.
- Nguyen, D., Gravel, R., Trieschnigg, D., & Meder, T. (2013). "How Old Do You Think I Am?"; A Study of Language and Age in Twitter. Paper presented at the Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. Retrieved 2015-12-31 from <http://www.dongnguyen.nl/publications/nguyen-icwsm2013.pdf>

- Oostdijk, N., Reynaert, M., Hoste, V., & Schuurman, I. (2013). The construction of a 500-million-word reference corpus of contemporary written Dutch *Essential Speech and Language Technology for Dutch: Results by the STEVIN-programme*: Berlin: Springer Verlag.
- Pennebaker, J. W. (2011). *The Secret Life of Pronouns*. New York: Bloomsbury Press.
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. Retrieved 2015-11-27, from <http://www.liwc.net>
- Pennebaker, J. W., Chung, C. K., Ireland, M. E., Gonzales, A., & Booth, R. J. (2007). The development and psychometric properties of LIWC2007.
- Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). Linguistic inquiry and word count: LIWC 2001. Mahwah, New Jersey: *Lawrence Erlbaum Associates*, 71, 2001.
- Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science*, 10(3), 90-93.
- Ramirez-Esparza, N., Pennebaker, J. W., García, F. A., & Suriá, R. (2007). La psicología del uso de las palabras: Un programa de computadora que analiza textos en español. *Revista Mexicana de Psicología*, 24(1), 85-99.
- Receptiviti. (2016). Enabling AI platforms with emotional intelligence. Retrieved 2016-07-04, 2016, from <http://www.receptiviti.ai/>
- Richardson, B. H., Taylor, P. J., Snook, B., Conchie, S. M., & Bennell, C. (2014). Language style matching and police interrogation outcomes. *Law and Human Behavior*, 38(4), 357.
- Scielzo, S. A., Patel, A., & Smith-Jentsch, K. A. (2011). Academic Mentoring Relationship Communication Processes and Participant-Reported Effectiveness. *Journal of Organizational Psychology*, 11(2), 81.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54.
- Tiedemann, J. (2012). *Parallel Data, Tools and Interfaces in OPUS*. Paper presented at the Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'2012). Retrieved
- Van der Zanden, R., Curie, K., Van Londen, M., Kramer, J., Steen, G., & Cuijpers, P. (2014). Web-based depression treatment: Associations of clients' word use with adherence and outcome. *Journal of Affective Disorders*, 160, 10-13.
- Van Middendorp, H., & Geenen, R. (2008). Poor cognitive-emotional processing may impede the outcome of emotional disclosure interventions. *British Journal of Health Psychology*, 13(1), 49-52.
- Wolf, M., Horn, A. B., Mehl, M. R., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative Textanalyse. Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. *Diagnostica*, 54(2), 85-98.
- Yu, B. (2014). Language and gender in Congressional speech. *Literary and Linguistic Computing*, 29(1), 118-132.
- Zijlstra, H., Van Meerveld, T., Van Middendorp, H., Pennebaker, J. W., & Geenen, R. (2004). De Nederlandse versie van de 'Linguistic Inquiry and Word Count' (LIWC) [The Dutch version of 'Linguistic Inquiry and Word Count' (LIWC)] *Gedrag & Gezondheid [Behaviour & Health]*, 32, 271-281.
- Zijlstra, H., Van Middendorp, H., Van Meerveld, T., & Geenen, R. (2005). Validiteit van de Nederlandse versie van de Linguistic Inquiry and Word Count (LIWC) [Validity of the Dutch version of Linguistic Inquiry and Word Count (LIWC)]. *Nederlands Tijdschrift voor Psychologie [Netherlands Journal of Psychology]*, 60(3), 50-58.