REVISTA SIGNOS. ESTUDIOS DE LINGÜÍSTICA ISSN 0718-0934
© 2022 PUCV, Chile • 55(110) 1029-1044

Distributional Semantics in Aphasia: An Exploratory Analysis Using Word Embeddings

Semántica distributiva en afasia: Un análisis exploratorio usando incrustación de palabras

Sebastián Bello Lepe
UNIVERSIDAD DE VALPARAÍSO
CHILE
sebastian.bello@uv.cl

Recibido: 01-IV-2022 / **Aceptado:** 07-XI-2022
DOI: 10.4067/S0718-09342022000301029

Abstract

Distributional semantics theory brings a cognitive/linguistic theoretical background to natural language processing. Thus, the analysis of the semantic distance between words could involve the study of language impairments, such as in individuals with aphasia. The present study explored the "Cinderella" storytelling of 80 neurotypical individuals, 108 individuals with non-fluent aphasia, and 138 individuals with fluent aphasia. Data was recollected from the AphasiaBank database. Semantic vectors were obtained using the word2vec model, a word embedding model to explore the vectorial projections of words and their meaning construction based on the context. Individuals with aphasia showed larger word vectorial distances in comparison with neurotypical controls. Furthermore, individuals with fluent aphasia had the larger word vectorial distance within the aphasia group. This is in concordance with findings of semantic feature impairments in individuals with fluent aphasia. The influence of using a computational approach with theoretical linguistic background in clinical linguistics is further discussed.

Keywords: Distributional semantics, word embedding, aphasia, semantic vector.

Resumen

La teoría de la semántica distributiva aporta una base teórica cognitiva/lingüística al procesamiento del lenguaje natural. Así, el análisis de la distancia semántica entre palabras podría implicar el estudio de las alteraciones del lenguaje, como en los individuos con afasia. El presente estudio exploró el discurso de "Cenicienta" de 80 individuos neurotípicos, 108 individuos con afasia no fluida y 138 individuos con afasia fluida. Los datos se recogieron de la base de datos AphasiaBank. Los vectores semánticos se obtuvieron utilizando el modelo *word2vec,* un modelo de incrustación de palabras para explorar las proyecciones vectoriales de las palabras y su construcción de significado en función del contexto. Los individuos con afasia mostraron mayores distancias vectoriales de palabras en comparación con los controles neurotípicos. Además, los individuos con afasia fluida tenían la mayor distancia vectorial de palabras dentro del grupo de afasia. Esto está en concordancia con los hallazgos alteraciones de ciertos rasgos semánticos en individuos con afasia fluida. Se discute además la influencia de utilizar un enfoque computacional con una base lingüística teórica en la lingüística clínica.

Palabras Clave: Semántica distributiva, incrustación de palabras, afasia, vector semántico.

INTRODUCTION

The development of data analysis based on natural language processing has been increasing exponentially in the last few years. Methodological approaches such as topic modelling or word embedding propose interesting approaches to explore linguistic features from a computational perspective. Word embeddings research comprises from sentimental analysis (Deho, Agangiba, Aryeh & Ansah, 2018) to behavioural analysis of social media (Sekar, Chandrakala & Prakash, 2021). Nevertheless, criticism of this method is related to the lack or misuse of a cognitive/linguistic theoretical background (Günter, Rinaldi & Marelli, 2019).

Distributional hypothesis and distributional semantics have emerged as theories behind word embedding and natural language processing. The main contribution of psychological/linguistic support to this analysis is the opportunity to explore semantic features beyond spared language. A robust computational analysis combined with theoretical background allows for observing and developing sensible explanations of language breakdown phenomena, such as linguistic performance in individuals with aphasia.

The present study uses a method from distributional semantics, the word embedding analysis, to explore semantic features in an individual with aphasia. The link between the distributional semantics, word embedding models, and linguistics characteristics in individuals with language breakdown seeks to enable comprehensive explanations for language research in clinical linguistics.

1. Theoretical Framework

1.1. Distributional Semantics and Word Embeddings

The distributional hypothesis posits similarities in meaning due to correlations of contexts where words occur and their distribution (Firth, 1957; Miller & Charles, 1991). Consequently, the differences in meanings are related to differences in their distribution (Harris,1954). Thus, distributional semantics (DS) could be defined as a usage-based model of meaning, taking into account the statistical distribution of lexical items in their context (Lenci, 2018), with similar words in meaning occurring in a similar context (Rubenstein & Goodenough, 1965). Linguistic and cognitive theories behind DS point out how the experience of a learner or speaker over a word or group of words determines how semantic representations, intraverbal connections, and word meanings are acquired (Jenkins, 1954; Landauer & Dumais, 1997).

Despite the theoretical background, DS models are algorithms representing word meanings as high dimensional numerical vectors (Günter et al., 2019). A DS model implies co-occurrences between lexical and linguistic contexts, distribution of lexical items, contexts weights, and semantic similarities between lexemes and row vectors in their lexical distribution of items (Lenci, 2018). Vector space models of semantics are grounded on usage-based and statistical properties of language, where lexical meaning is configured from word usage (Turney & Pantel, 2010). As in morphemes, differences in word meanings relates to different contexts, non-related to their semantical properties (Harris, 1968).

Several methods have been developed to explore these semantic representations, such as Latent Semantic Analysis (Landauer & Dumais, 1997) or Latent Dirichlet Analysis (Griffiths, Steyvers & Tenenbaum, 2007).

Recently, word2vec has become a prominent model to explore distributional word vectors in a context (Mikolov, Yih & Zweig, 2013). The model trains word vectors, building a hidden layer in a neural network, predicting words from their word neighbours and backwards. This process, also called word embedding, allows converting a word into a number (Sivakumar, Videla, Rajesh Kumar, Nagaraj, Itnal & Haritha, 2020). Two approaches are proposed to word embeddings: 1) Continuous Bag of Words (CBOW) and 2) Skip Gram. CBOW uses context, learning the embedding from all nearby words. In CBOW, the input is a context of words, whereas the output is a single word (Sivakumar et al., 2020). The architecture of a CBOW model (Fig. 1) implies a window size (C), a weight matrix (W), a number of words in vocabulary (V), a number of nodes in the concealed layer, and it is the training parameter (N) and the output (Y) (Sivakumar et al., 2020).

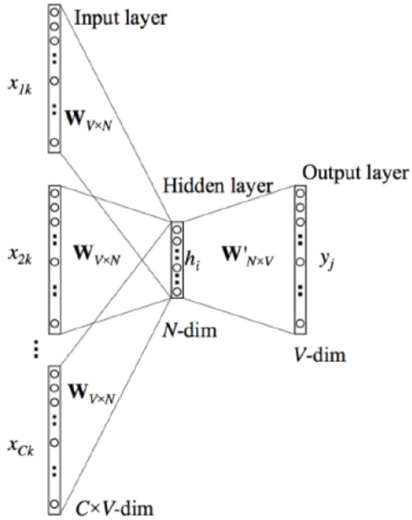

Figure 1. Architecture of CBOW model (From Sivakumar et al., 2020).

In Skip-Gram models, the context words are predicted by the main word. Unlike CBOW, the input in Skip-Gram models is single words, whereas the output is in the context of these words. The Skip-Gram architecture (Figure 2) compresses similar parameters as CBOW architecture plus a concealed layer with its weighted sum of the input vector given (Hi) and output of the concealed layer (k^{th}) (Sivakumar et al., 2020).

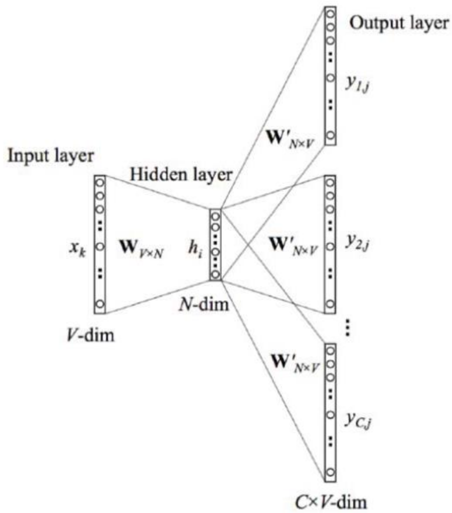

Figure 2. Architecture of Skip-Gram model (From Sivakumar et al., 2020).

Independent of the approach, word embeddings could capture some statistical patterns of word retrieval across discourse. Landauer and Dumais (1997) point out how the dimensional distribution of vectors could reflect latent semantic dimensions.

Although the kind of semantic features that word embeddings could encompass is still not precise (Günter et al., 2019), exploratory approaches have found in semantic spaces some essential components of semantic dimensions, such as concreteness or affective features (Osgood, Suci & Tannenbaum, 1957; Hollis & Westbury, 2016). Thus, the application of word embedding methods for language analysis could offer explanations beyond the methodological characteristics, given possible frameworks for the studies on language breakdown.

1.2. Language in Individuals with Aphasia

Aphasia could be defined as an impairment of language, affecting verbal comprehension, production of speech, the ability to write or read (National Aphasia Association, 2021). Frequent symptomatology includes difficulty initiating speech (Crosson, Ford & Raymer, 2018), non-fluent speech reducing the phrase length agrammatism, comprehension impairment of single words or sentences, problems in repetition and difficulties in lexical retrieval (Sheppard & Sebastian, 2021).

From a linguistic component of language perspective, individuals with aphasia (IWA) have impairment in phonology, morphology, syntax and semantics structures (Bryant, Ferguson & Spencer, 2016). The linguistic impairments are related to several types and or aphasia classifications. Nevertheless, a familiar dichotomy in clinical aphasia includes non-fluent and fluent individuals, where in the former, the impairment in linguistic components affects speech connectivity, and in the latter, there is no effect on speech fluency.

A common discussion about semantic deficits arises among individuals with fluent aphasia (fIWA). Observations on Wernicke's aphasia, a type of fIWA, were constantly focused on impairments in perceptual attributions of language (Luria, 1970). However, other researchers have found multimodal semantic impairments (Robson Robson, Grube, Lambon Ralph, Griffiths & Sage, 2012; Thompson, Robson, Lambon Ralph & Jefferies, 2015) and deficits in the executive process involved in the semantic activation control (Jefferies & Lambon Ralph, 2006). The semantic deficits in individuals with non-fluent aphasia (nfIWA) are less common, finding some semantic errors in auditory comprehension and picture naming (Butterworth, Howard & Mcloughilin, 1984).

Regardless of the nature of possible semantic deficits in IWA, word embedding approaches offer an opportunity to explore, from a DS perspective, how the linguistic impairments could manifest in vectorial semantic spaces. Moreover, whether semantic deficits arise from IWA, they might be detected by models involving measuring semantic distances of words in discourse. Thus, DS models could be useful to explore some lexical patterns of the IWA speech.

1.3. Aim of the Study

Considering the evidence about semantic deficits in IWA, the study aims to explore whether a DS-based approach, a word embedding analysis, is sensitive to semantic deficits. The research questions are as follows:

1. Does IWA show higher values for vector distances compared to neurotypical individuals?
2. Does fIWA show higher values for vector distances compared to nfIWA?

The first hypothesis establishes that IWA has greater vector distances values than neurotypical controls. The second hypothesis poses that the vector distances values from the fIWA group are more significant than the nfIWA group. The present study is an exploratory analysis using the word embeddings approach, precisely the word2vec technique (Mikolov, Sutskever, Chen, Corrado & Dean, 2013).

2. Methods

2.1. Sample and Demographic Description

For the present study, 83 neurotypical, 108 nfIWA and 138 fIWA English speakers' samples narrations of the Cinderella storytelling (Grimes, 2005) were obtained from the AphasiaBank database (MacWhinney, 2000; MacWhinney, Fromm, Forbes & Holland, 2011). Table 1 shows demographic and narration features for each group.

Table 1. Individuals with nfIWA are slightly younger in comparison with the other groups.

	Age mean and SD	Years of education Mean and SD	Number of words
Neurotypical	73.94 (10.31)	15.48 (2.71)	30.718
Non-Fluent Aphasia	59.78 (12.21)	14.82 (3.08)	15.678
Fluent Aphasia	63.92 (12.59)	15.29 (3)	33.899

A one-way ANOVA shows statistically significant age differences ($p < .01$) and the number of words ($p <.01$) across the groups. No statistical differences in years of education. From each sample, all the individuals declared to know the Cinderella storytelling. Dillow (2013) founded 19 core nouns lexicon in Cinderella (Table 2). The core nouns were used as reference for word embedding analysis.

Table 2. Principal nouns are related with the successful of Cinderella narration (Dillow, 2013).

Ball
Cinderella
Daughter
Dress
Fairy
Foot
Girl
Glass
Home
Horse
House
Midnight
Mother
Mouse
Prince
Pumpkin
Sister
Slipper
Time

2.2. Word Embedding Analysis

For word embedding, the word2vec model (Mikolov et al., 2013) was applied using an open-source library for unsupervised natural language processing in Python V 3.0 (van Rossum & Drake, 2009) called Gensim (Rehurek & Sojka, 2011). The code lines are mainly based on Pierremegret (2018) work. A model was created and trained for each group independently. The reference model for word embedding was based on the Skip-Gram approach. Each model was trained based on 'en_core_web_sm' v.3.2.0, a trained pipeline from spaCy v.3.2 (Honnibal, Montani, van Landeghem & Boyd, 2020) that includes vocabulary, syntax and entities. Stopwords removed included special characters (e.g., "") and uppercases. Each model had a two-window size, 300 vector size with an alpha = 0.03 and 30 epochs. The reference word for a similar vector was "cinderella". A T-distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten & Hinton, 2008) tool was used to visualize the semantic distance between similar words. A cosine distance was applied between the word "cinderella" and each storytelling core noun to measure the vectorial semantic space in each group.

3. Results

3.1. Similar Words

Table 3 shows the highest similar words to Cinderella for each group.

Table 3. No words matches within the three groups.

Neurotypical group		nfIWA		fIWA	
Word	**Semantic distance**	**Word**	**Semantic distance**	**Word**	**Semantic distance**
time	0.9997892379760742	fairy	0.9997804760932922	let	0.9997683763504028
hadta	0.999787449836731	fit	0.9997733235359192	clothe	0.9997678995132446
father	0.99977707862854	thing	0.9997717142105103	man	0.9997645616531372
clean	0.9997737407684326	go	0.9997702836990356	like	0.9997575283050537
show	0.9997674226760864	slipper	0.9997639656066895	place	0.9997544884681702
ready	0.9997652769088745	get	0.9997628927230835	bad	0.9997528791427612
tell	0.9997652769088745	o'clock	0.9997514486312866	mouse	0.9997522234916687
pretty	0.999764084815979	maid	0.9997509121894836	woman	0.9997453689575195
lose	0.9997621774673462	girl	0.9997508525848389	find	0.9997447729110718
decide	0.9997615814208984	time	0.999748945236206	god	0.9997444152832031

A linear regression showed no influence of age over semantic distances ($\beta = 2.49e\text{-}07$, $t = .82$, $p = .41$). Figure 3 shows the dimensional space of similar words in the neurotypical group.

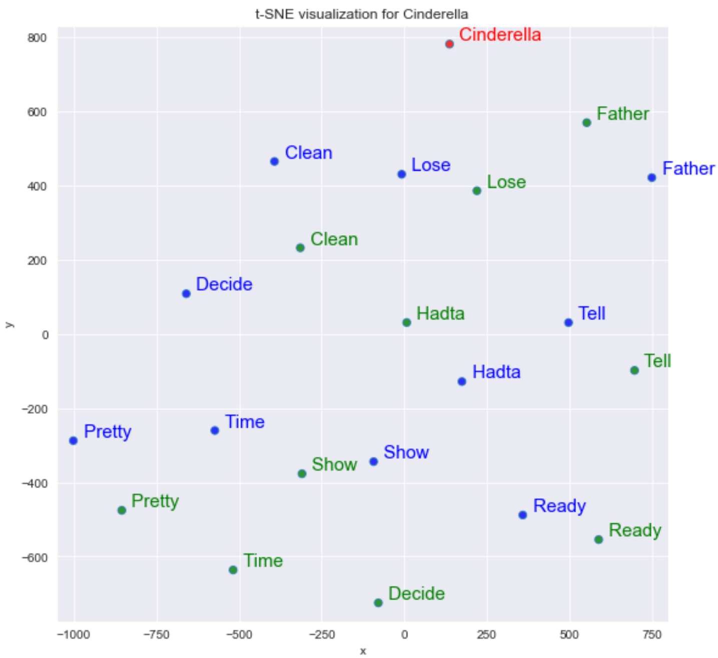

Figure 3. Red word represents the query word (Cinderella). Blue words represent the most similar words with "Cinderella". Green words represent other words from the vocabulary (duplicated with similar words).

Figure 4 shows the dimensional space of similar words in the nfIWA group.

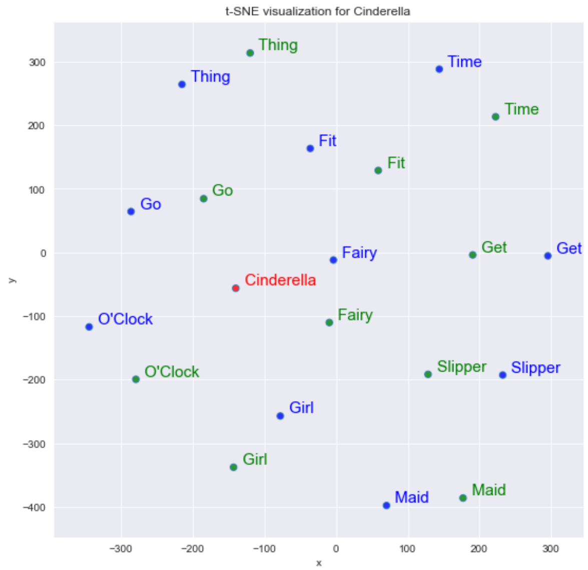

Figure 4. Red word represents the query word (Cinderella). Blue words represent the most similar words with "Cinderella". Green words represent other words from the vocabulary (duplicated with similar words).

Figure 5 shows the dimensional space of similar words in the fIWA.

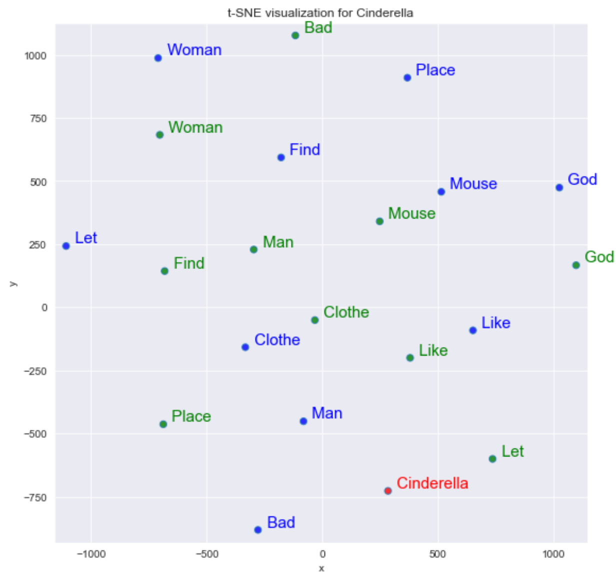

Figure 5. Red word represents the query word (Cinderella). Blue words represent the most similar words with "Cinderella". Green words represent other words from the vocabulary (duplicated with similar words).

3.1. Semantic Vectorial Distance between Cinderella and the Core Nouns

Table 4 shows by group the cosine distance between the word "cinderella" with each of the core nouns described in Dillow (2013) study.

Table 4. Neurotypical and nfIWA groups did not show semantic vectorial distance for two words. There were semantic vectors for all the core nouns in the fIWA group.

Word	Neurotypical semantic vector values	nfIWA semantic vector values	fIWA semantic vector values
Ball	0.00027555227279663086	0.0002605915069580078	0.000268399715423584
Daughter	0.0002846717834472656	0.0003024935722351074	0.00028330087661743164
Dress	0.0002728104591369629	0.00027817487716674805	0.0003019571304321289
Fairy	not present	0.00021952390670776367	0.00028502941131591797
Foot	0.0002722740173339844	not present	0.0002714395523071289
Girl	0.000263214111328125	0.0002490878105163574	0.00031059980392456055
Glass	not present	0.000319063663482666	0.0002981424331665039
Home	0.0002783536911010742	0.0003037452697753906	0.000268399715423584
Horse	0.00025159120559692383	0.00026851892471313477	0.0002752542495727539
House	0.00025594234466552734	0.0002574920654296875	0.0002778768539428711
Midnight	0.0002620816230773926	0.0003204941749572754	0.0002689361572265625
Mother	0.00027489662170410156	0.0002745389938354492	0.00033777952194213867
Mouse	0.00025081634521484375	not present	0.00024771690368652344
Prince	0.00024968385696411133	0.000261843204498291	0.0002694129943847656
Pumpkin	0.0002570152282714844	0.00026553869247436523	0.00027996301651000977
Sister	0.00028902292251586914	0.0003007054328918457	0.00028449296951293945
Slipper	0.0002722740173339844	0.00023609399795532227	0.00030106306076049805
Time	0.00021082162857055664	0.0002509951591491699	0.00026220083236694336

A Shapiro-Wilk test showed a normal distribution of semantic vector values ($W = .96$, $p = .209$). A one-way ANOVA was applied to determine semantic vector values differences within the group. There was a statistically significant difference in the semantic vectors values within groups (($f(1, 40) = 6.54$, $p = .01$; $Eta^2 = .14$, 90% CI [0.02, 0.31]). A pairwise comparison was applied to analyse the differences between groups (Table 5). No significant effects of age over cosine values ($\beta = 9.91e\text{-}07$, $t = 1.81$, $p = .07$)

Table 5. Pairwise comparison shows statistically significant differences between fIWA group and neurotypical group. No differences between neurotypical group with nfIWA group and nfIWA with fIWA group.

	Neurotypical group	nfIWA group
nfIWA group	.25	-
fIWA group	.04	.25

4. Discussion

The present study explored the sensitivity of word embedding analysis to detect semantic vector distances differences within neurotypical individuals, nfIWA and fIWA, from a storytelling discourse. In concordance with hypothesis one, there were significant differences in word semantic spaces within groups. Moreover, the fIWA group had statistically significant greater values of semantic vector distances than the neurotypical and nfIWA groups.

As word embedding is a model based on DS, these results support the claim that fIWA has semantic impairments, reflected in an increase of the word vectorial distance in discourse compared with nfIWA and neurotypical individuals. Considering the theoretical premises of DS, in fIWA, some patterns of the distributional properties of lexemes are impaired concerning their linguistic context. However, neither lexical properties or linguistic context could be suggested as the main impairment factor. Interestingly, nfIWA did not show differences with both groups, even when the group had the lowest number of words. This is in concordance with Alyahya, Conroy, Halai and Lambon-Ralph (2021), where IWA discourse shows a decrease in word quantity but not in lexical-semantic complexity. The researchers found this pattern in both fluent and non-fluent individuals with aphasia. Nevertheless, considering the number of words and vectorial semantic distance results in the present study might be possible to posit that this effect is more significant for nfIWA than fIWA.

Lenci (2018) remarks that conceptual differences exist between words with similar semantic features and strong associations between words. As part of DS methods, word embedding is related to the construction of word-association networks (Lenci, 2018). This implies that, for the present study, the fIWA results could represent a weakly word association instead of impairments on the semantic features that compose the nodes between words. This weakly association could rely on the contextual characteristics of the task than degradation of semantic processing. For instance, storytelling as "Cinderella" is a structured narration with a closed vocabulary. Word embeddings explore semantic relations based on contexts. Thus, lexical-semantic impairments in fIWA could depends on contextual features of the task than disruptions on lexical-semantic meanings. Some contexts could boost the residual capacity, whereas others may reduce it. Further studies could compare word embedding analysis with other context-driven theories, other types of narrations and statistical properties of language, such as lexical frequency.

Since the current study is an exploratory approach to word embedding for clinical linguistics analysis, there are significant limitations. First, it is necessary to increase the number of words in each group to improve the model accuracy. Typically, a significant word embedding model involves millions to billions of words. The AphasiaBank database is continually increasing the number of samples so that further analysis could include more Cinderella discourses. The second is the relationship with other variables. As mentioned before, lexical frequency could be a factor that influences the IWA word pool. Some studies have shown that IWA benefits from statistical properties of language, easing the retrieval of words with high lexical frequency (Kittredge, Dell, Verkuilen & Schwartz, 2008; DeDe, 2014).

Further studies could consider how and what is the relationship between vectorial semantic spaces and lexical frequency. Third, it was not possible to acquire the severity of aphasia. Butterworth et al. (1984) mentioned that the aphasia severity is a stronger factor than the type of aphasia for semantic impairments. Thus, nfIWA could have severe linguistic problems that might have a greater effect on semantic processing when is compared to the fIWA group. A future study may consider groups by level of severity instead of the fluent/non-fluent dichotomy. Finally, the word embedding model used in the present study could be replaced for models with larger data trained. For example, Glove (Pennington, Socher & Manning, 2014) is an unsupervised learning algorithm from Stanford University for obtaining vector representations for words. One of its pre-trained word vector models contains almost six billion tokens and more than 400.000 vocabularies. A stronger pre-trained model helps to improve the accuracy of the input's semantic vectors.

CONCLUSIONS

Word embedding could be sensitive to semantic differences between neurotypical individuals, nfIWA and fIWA. Individuals with fluent aphasia have been described to have linguistic impairments at the semantic level. This group's higher semantic vectorial distances could reflect these difficulties, expressing weaker semantic association networks during storytelling like "Cinderella". Despite further analysis, the present study shows that it is possible to apply natural language processing approaches to clinical linguistics. When models such as word embeddings of topic modelling have been already applied in clinical language-impaired individuals, the lack of relationship to a linguistic theoretical framework denies plausible explanations for the data analysed. DS could be a strong linguistic theory to support word embedding analysis in individuals with aphasia.

REFERENCES

Alyahya, R., Conroy, P., Halai, A. & Lambon Ralph, M. (2021). An Efficient, Accurate and Clinically-Applicable Index of Content Word Fluency in Aphasia. *Aphasiology*, 1-19.

Bryant, Ferguson, A. & Spencer, E. (2016). Linguistic Analysis of Discourse in Aphasia: A Review of the Literature. *Clinical Linguistics & Phonetics, 30*(7), 489-518. DOI: https://doi.org/10.3109/02699206.2016.1145740

Butterworth, B., Howard, D. & Mcloughlin, P. (1984). The Semantic Deficit in Aphasia: The Relationship between Semantic Errors in Auditory Comprehension and Picture Naming. *Neuropsychologia*, *22*(4), 409-426. DOI: https://doi.org/10.1016/0028-3932(84)90036-8

Crosson, B. A., Ford, A. & Raymer, A. M. (2018). Transcortical Motor Aphasia. In A. M. Raymer & L. J. Gonzalez Rothi (Eds.), The Oxford Handbook of Aphasia and Language Disorders (pp. 171-186). Oxford: Oxford Library of Psychology. DOI: https://doi.org/10.1093/oxfordhb/9780199772391.013.11

DeDe G. (2012). Effects of Word Frequency and Modality on Sentence Comprehension Impairments in People with Aphasia. *American Journal of Speech-Language Pathology,* *21*(2), S103-S114. DOI: https://doi.org/10.1044/1058-0360(2012/11-0082)

Deho, B., Agangiba A., Aryeh, L. & Ansah, A. (2018) Sentiment Analysis with Word Embedding. *IEEE 7th International Conference on Adaptive Science & Technology (ICAST)*, 1-4. DOI: 10.1109/ICASTECH.2018.8506717

Dillow, E. (2013). *Narrative Discourse in Aphasia: Main Concept and Core Lexicon Analyses of the Cinderella Story*. Doctoral dissertation, University of South Carolina, United State.

Firth J. R. (1957). (Ed.). A Synopsis of Linguistic Theory, 19301955. In Studies in Linguistic Analysis (pp. 1-32). Oxford, UK: Blackwell.

Griffiths, T., Steyvers, M. & Tenenbaum, J. B. (2007). Topics in Semantic Representation. *Psychological Review,* *114*(2), 211-244. DOI: https://doi.org/10.1037/0033-295X.114.2.211

Grimes N. (2005). *Walt Disney's Cinderella.* New York: Random House.

Günther, F., Rinaldi, L. & Marelli, M. (2019). Vector-Space Models of Semantic Representation from a Cognitive Perspective: A Discussion of Common Misconceptions. *Perspectives on Psychological Science,* *14*(6), 1006-1033. DOI: https://doi.org/10.1177/1745691619861372

Harris Z. S. (1954). Distributional structure. *Word,* 10, 146-62.

Harris, Z. (1968). *Mathematical Structures of Language*. New York, USA: Wiley.

Hollis, G. & Westbury, C. (2016). The Principals of Meaning: Extracting Semantic Dimensions from Co-Occurrence Models of Semantics. *Psychonomic Bulletin & Review, 23*(6), 1744-1756. DOI: https://doi.org/10.3758/s13423-016-1053-2

Honnibal, M., Montani, I., van Landeghem, S. & Boyd, A. (2020). spaCy: Industrial-strength Natural Language Processing in Python. DOI: 10.5281/zenodo.1212303

Jefferies, E. & Lambon Ralph, M. A. (2006). Semantic Impairment in Stroke Aphasia Versus Semantic Dementia: A Case-Series Comparison. *Brain: A Journal of Neurology*, *129*(Pt 8), 2132-2147. DOI: https://doi.org/10.1093/brain/awl153

Jenkins, J. J. (1954). Transitional Organization: Association Techniques. In C. E. Osgood & T. A. Sebeok (Eds.), *Psycholinguistics. A Survey of Theory and Research Problems* (pp. 112-118). Bloomington: Indiana University Press.

Kittredge, A. K., Dell, G. S., Verkuilen, J. & Schwartz, M. F. (2008). Where is the Effect of Frequency in Word Production? Insights from Aphasic Picture-Naming Errors. *Cognitive Neuropsychology, 25*(4), 463-492. DOI: https://doi.org/10.1080/02643290701674851

Landauer, T. & Dumais, S. (1997). A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction and Representation of Knowledge. *Psychological Review, 104*(2), 211-240.

Lenci. (2018). Distributional Models of Word Meaning. *Annual Review of Linguistics, 4(*1), 151-171. DOI: https://doi.org/10.1146/annurev-linguistics-030514-125254

Luria A.R. (1970). *Traumatic Aphasia*. Mouton: The Hague.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk (3rd Edition). Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B., Fromm, D., Forbes, M. & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology,* 25,1286-1307.

Mikolov, T., Yih. W.-T. & Zweig, G. (2013). Linguistic Regularities in Continuous Space Word Representations. In *Proceedings of the 2013 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 746-51). Stroudsburg, PA: Assoc. Comput. Linguist.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. & Dean, J. (2013). Distributed Representations of Words and Phrases and their Compositionality. In *Proceedings of the 26th Conference on Advances in Neu- ral Information Processing Systems*, pp. 311119 [online]. Retrieved from: https://papers.nips.cc/book/advances-in-neural-information-processing-systems-26-2013

Miller, G. & Charles, W. (1991). Contextual Correlates of Semantic Similarity. *Lang. Cogn. Process,* 6,1-28.

National Aphasia Association (2021). [online]. Retrieved from: https://www.aphasia.org/aphasia-definitions/

Osgood, C., Suci, G. J. & Tannenbaum, P. H. (1957). *The Measurement of Meaning.* Urbana: University of Illinois Press.

Pennington, J., Socher, R. & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP-2014)*. Stroudsburg, PA: Association for Computational Linguistics.

Pierremegret. (2018). *Gensim word2vec Tutorial.* Kaggle. [online]. Retrieved from: https://www.kaggle.com/code/pierremegret/gensim-word2vec-tutorial/notebook

Rehurek, R. & Sojka, P. (2011). Gensim–Python Framework for Vector Space Modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, *3*(2).

Robson, H., Grube, M., Lambon Ralph, M. A., Griffiths, T. D. & Sage, K. (2012). Fundamental Deficits of Auditory perception in Wernicke's Aphasia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *49*(7), 1808-1822. DOI: https://doi.org/10.1016/j.cortex.2012.11.012

Rubenstein, H. & Goodenough, J. (1965). Contextual Correlates of Synonymy. *Communications of the ACM, 8*(10), 627-633.

Sekar, N. & Chandrakala, S. & Prakash, G. (2021). Analysis of Global Word Representations for Depression Detection. Conference: CEUR Workshop Proceedings (CEUR-WS.org); International Semantic Intelligence Conference 2021, Volume 2786, New Delhi, India,

Sheppard, S. M. & Sebastian, R. (2021). Diagnosing and Managing Post-Stroke Aphasia. *Expert Review of Neurotherapeutics, 21*(2), 221-234. DOI: https://doi.org/10.1080/14737175.2020.1855976

Sivakumar, S., Videla, L. S., Rajesh Kumar, T., Nagaraj, J., Itnal, S. & Haritha, D. (2020). Review on word2vec Word Embedding Neural net. *2020 International Conference on Smart Electronics and Communication (ICOSEC)*. DOI: https://doi.org/10.1109/icosec49089.2020.9215319

Thompson, H. E., Robson, H., Lambon Ralph, M. A. & Jefferies, E. (2015). Varieties of Semantic 'Access' Deficit in Wernicke's Aphasia and Semantic Aphasia. *Brain: A Journal of Neurology*, *138*(Pt 12), 3776-3792. DOI: https://doi.org/10.1093/brain/awv281

Turney, P. & Pantel, P. (2010). From Frequency to Meaning: Vector Space Models of Semantics. *Journal of Artificial Intelligence Research*, 37, 141-188.

van der Maaten, L. & Hinton, G. (2008). Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research, 9*(86), 2579-2605.

van Rossum, G. & Drake, F. L. (2009). *Python 3 Reference Manual.* Scotts Valley, CA: CreateSpace.