ALZUMERIC: A decision support system for diagnosis and monitoring of cognitive impairment

Unai Martinez de Lizarduy[1], Pilar Calvo Salomón[1], Pedro Gómez Vilda[2],
Mirian Ecay Torres[3,4], Karmele López de Ipiña[1]

[1] UPV/EHU (Basque Country University) Donostia-San Sebastián, Spain
[2] Centro de Tecnología Biomédica. UPM, Madrid, Spain
[3] Fundación CITA-Alzheimer Fundazioa, Donostia-San Sebastián, Spain
[4] Department of Basic Psychology, Psychology Faculty. UNED, Madrid, Spain
unai.martinezdelizarduy@ehu.eus ORCID: http://orcid.org/0000-0003-0969-4857
pilarmaria.calvo@ehu.eus ORCID: http://orcid.org/0000-0002-5886-5475
pedro@fi.upm.es ORCID: http://orcid.org/0000-0003-3283-378X
mecay@cita-alzheimer.org ORCID: http://orcid.org/0000-0001-7170-6180
karmele.ipina@ehu.eus ORCID: http://orcid.org/0000-0002-2135-2393

ABSTRACT: Internet of things and smart cities are becoming a reality. Nowadays, more and more devices are interconnected and in order to deal with this new situation, data processing speeds are increasing to keep the pace. Smart devices like tablets and smartphones are accessible to a wide part of society in developed countries, and Internet connections for data exchange make it possible to handle large volumes of information in less time. This new reality has opened up the possibility of developing client-server architectures focused on clinical diagnosis in real time and at a very low cost. This paper illustrates the design and implementation of the *ALZUMERIC* system that is oriented to the diagnosis of Alzheimer's disease (AD). It is a platform where the medical specialist can gather voice samples through non-invasive methods from patients with and without mild cognitive impairment (MCI), and the system automatically parameterizes the input signal to make a diagnose proposal. Although this type of impairment produces a cognitive loss, it is not severe enough to interfere with daily life. The present approach is based on the description of speech pathologies with regard to the following profiles: phonation, articulation, speech quality, analysis of the emotional response, language perception, and complex dynamics of the system. Privacy, confidentiality and information security have also been taken into consideration, as well as possible threats that the system could suffer, so this first prototype of services offered by *ALZUMERIC* has been targeted to a predetermined number of medical specialists.

Keywords: Alzheimer's Disease; Mild Cognitive Impairment; Automatic Speech Processing; Decision Support Systems; Convolutional Neural Networks.

RESUMEN: *ALZUMERIC: Un sistema de apoyo a la decisión para el diagnóstico y la monitorización del deterioro cognitivo.* – La Internet de las cosas o de las ciudades inteligentes se está convirtiendo en una realidad. Cada vez más dispositivos están interconectados y, para hacer frente a esta nueva situación, las velocidas de procesamiento de datos se incrementan. Los dispositivos inteligentes, como las tabletas y los teléfonos, son accesibles para una gran parte de la sociedad en los países desarrollados, y las mejoras en las conexiones a Internet para el intercambio de datos hacen posible manejar grandes volúmenes de información en menos tiempo. Esta nueva realidad ha abierto la posibilidad de desarrollar arquitecturas cliente-servidor centradas en el diagnóstico clínico en tiempo real y a un coste muy bajo. Este trabajo ilustra la concepción del sistema *ALZUMERIC* orientado al diagnóstico de la enfermedad de Alzheimer. Es una plataforma a partir de la cual el médico especialista puede tomar muestras de voz a través de métodos no invasivos a pacientes con y sin deterioro cognitivo leve (MCI), y en la cual se parametriza la señal de entrada automáticamente para posteriormente avanzar una propuesta de diagnóstico. El MCI es un tipo de deterioro neurológico que produce una pérdida cognitiva no lo suficientemente grave como para interferir en la vida cotidiana.

El presente estudio está enfocado hacia la descripción de las patologías del habla con respecto a los siguientes perfiles: fonación, articulación, calidad del habla, análisis de la respuesta emocional, percepción del lenguaje y dinámica de sistemas complejos. También se consideran aspectos relativos a la privacidad, la confidencialidad y la seguridad de la información frente a las posibles amenazas que pudiera sufrir el sistema, por lo que este primer prototipo de servicios ofrecidos por *ALZUMERIC* se ha dirigido a un número predeterminado de médicos especialistas.

Palabras clave: enfermedad de Alzheimer; deterioro cognitivo leve; procesamiento automático del habla; sistemas de apoyo a la decisión; Redes Neuronales convolucionales.

1. INTRODUCTION

In a world in continuing evolution, where digital interactions are part of daily life and Internet of things or smart cities are becoming a reality in many fields, health services start demanding new interactive and digital solutions where the final user will play an active role. Besides, classic clinical practice relies increasingly on technological support for making clinical decisions, which serves as a help for specialists, reduces the time to perform a reliable diagnosis, and substantially improves health resources by decreasing medical testing time and costs (Alzheimer's Disease International, 2015).

In this scenario, the dramatic aging of the world population is causing cognitive impairment to reach epidemic level. In this sense, its early detection by premature symptoms would be crucial to optimize management, and could open a new scenario for potential interventions. Undoubtedly, an early and accurate diagnosis of this kind of diseases could help to decrease its social impact and consequences. Thus, early clinical tests should be reinforced and consistent with complementary evaluations, rehabilitation and monitoring. Over the last decades there have been useful advances not only in classic assessment techniques but also in novel screening strategies for preliminary assessment or monitoring during the evolution of the disease. Non-invasive intelligent techniques of diagnosis may become valuable tools for clinical and domestic environments that can complement diagnosis and/or monitor its progress. These methodologies are easy to apply, and can be introduced in the standard medical protocols without altering daily life. Moreover, they are capable of yielding information easily, quickly, and inexpensively (Faundez-Zanuy et al., 2013; Klimova et al., 2015; Laske et al., 2015; López-de-Ipiña et al., 2013). Thus, even non-technical staff in the usual environment of the patient could use these methodologies without altering or blocking the patient abilities, but a medical supervision is required (Laske et al., 2015). Among non-invasive methods, automatic analysis of speech, which involves verbal communication (one of the first damaged skills in MCI/AD patients), provides a natural, friendly and powerful tool for early diagnosis and low cost screening.

This work presents the conception and preliminary evaluation of the *ALZUMERIC* system, a low-cost technological platform oriented to the diagnosis of Alzheimer's disease (AD) by speech analysis. It is a platform where the medical specialist can gather voice samples through non-invasive methods from patients with and without mild cognitive impairment (MCI) and AD. The server parameterizes the input signal for medical assessment and it makes an evaluation so that the patient is assigned to the group with MCI or excluded from it. The score of diagnosis has substantially improved with deep learning algorithms. This diagnosis is crucial because it may allow to treat the disease at an early stage. The entire process is monitored in real time and is based on protocols included under TCP/IP. Privacy and information security have also been taken into consideration, as well as possible threats that the system could suffer, so this first prototype of services offered by *ALZUMERIC* has been targeted to a predetermined number of medical specialists. In this framework, HTTP/HTTPS clients, on web protocols, are responsible for the coordination and exchanges of information between the final user and the servers, and are integrated into all terminals with Internet access. They are the most important part of the tool. It is likely that the expansion of desktop and mobile devices for clinical use will keep increasing, and since their use is widespread, the platform has a really short learning curve. The analysis of the information matrix to obtain a reliable diagnosis is based on a set of algorithms of deep learning that achieves substantial improvements on results.

This paper is organized as follows. Materials are described in Section 2. Section 3 summarizes the used methods. The protocols and algorithms are discussed in Section 4. The results and discussion are summarized in Section 5, and finally concluding remarks are drawn in Section 6.

2. MATERIALS

In the development and clinical evaluation protocol of the *ALZUMERIC* system, three tasks with different levels of language complexity have been used: Categorical Verbal Fluency (CVF, based on Animal Naming, AN), Picture Description (PD), and Spontaneous Speech (SS). AN and PD have been recorded in a clinical environment, and SS has been carried out in a domestic environment. The participants are different in the three tasks. PD and SS analyze AD patients vs. individuals from a control group (CR), and AN analyzes MCI patients vs. CR. All the work was performed strictly following the ethical guidelines of the organizations involved in the project.

2.1. Task of *Categorical Verbal Fluency* (CVF), based on animal naming

The task of *Categorical Verbal Fluency* (CVF, or Animal naming, AN), is a test that measures and quantifies the progression of cognitive impairment in neurodegenerative diseases (López-de-Ipiña, 2015). It is widely used to assess language skills, semantic memory and executive functions. During the CVF task the interviewer asks patients to list in one minute all the names they can remember from a category, animal names in our case. The sample consists of 187 healthy controls and 38 patients with MCI, from the cohort of Gipuzkoa-Alzheimer Project (PGA, CITA-PGA, 2017) of the CITA-Alzheimer Foundation.

2.2. Picture description task

This pilot study, which is part of the Gipuzkoa-Alzheimer Project (PGA) and the Memory Clinic of the CITA-Alzheimer Foundation, includes six subjects with diagnosis of AD and 12 healthy controls (sub-database MINI-PGA), that is, the sample included data from 18 subjects. The task consists in the verbal description of a picture (CITA-PGA, 2017).

2.3. Spontaneous speech task

In order to develop a new methodology applicable to a wide range of individuals of different sex, age, language, and cultural and social background, a multicultural and multilingual spontaneous speech database has been built in English, French, Spanish, Catalan, Basque, Chinese, Arabic, and Portuguese, comprising video recordings of 50 healthy people (12 hours), and 20 patients with a prior diagnosis of AD (8 hours). The age span of the individuals in the full database is 20-98 years. This database is called AZTIAHO (López-de-Ipiña et al., 2015). The complete speech database consists of about 60 minutes for the AD group, and about 9 hours for the control group. Hereafter, the speech was divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers, yielding a database of approximately 600 segments of spontaneous speech. Finally, in order to perform our experiments a subset was selected, balanced with regard to the age of the participants and the emotional response level. The subset consists of 20 subjects (9 women and 11 men) from the control group, and 20 AD patients (12 women and 8 men). This subset of the database is called *AZTIAHORE*.

3. METHODS

Recent works point out the relevance of dysfluency in speech as a hallmark of MCI and AD. In López-de-Ipiña et al. (2015) it is suggested that shorter speech segments reflect that AD patients require a greater effort to produce speech than healthy individuals.

AD patients speak more slowly with longer pauses, and they spend more time looking for the correct word, which leads to speech dysfluencies or broken messages. Speech dysfluencies are any break, irregularity or non-lexical element that occurs within the period of fluent speech, and that could start or interrupt it.

These include among others, false starts, repeated or re-started phrases, repeated or extended syllables, grunts or non-lexical utterances such as fillers and repaired speech, and instances of speakers correcting their own lisps or mispronunciations (Dingemanse et al., 2013; López-de-Ipiña et al., 2017). In AD patients sometimes these become a verbal utterance of the internal cognitive process or an inner dialogue, and they produce sentences such as: "What is that?", "What was the name?", "Uhm, I can't remember". The increase in the number of dysfluencies and silences may be a sign of deterioration of the disease, which could lead to a deficit in clear communication. As a conclusion, dysfluencies are a direct reflection of the cognitive process during communication, and become an unquestionable hallmark for the detection of this kind of disorders. Although AD is mainly a cognitive disease, it may also present neuromechanical alterations in phonation and articulation. The task of Categorical Verbal Fluency (CVF; animal naming, AN), or animal fluency task, is a test that measures and quantifies the progression of cognitive impairment in neurodegenerative diseases (Lezak et al., 2012; Ruff et al., 1997). The CVF protocol is widely used to assess language skills, semantic memory and executive functions (Lezak et al., 2012). In Figure 1 two examples collected using this protocol may be seen. In Figure 1 (left) the production of a control subject is presented. The subject can utter as many as 40 names in 62 s, whereas in Figure 1 (right) the production of an MCI subject is of only 6 names in 37 s.

The data sample produced using CVF consists of 187 healthy controls and 38 MCI patients being a subset of the cohort of Gipuzkoa-Alzheimer Project (CITA-Alzheimer Foundation, 2017; López-de-Ipiña et al., 2015). Age distributions are given in Table 1. A subset from PGA-OREKA was selected for experimentation.

4. PROTOCOLS AND ALGORITHMS

The speech analysis approach to evaluate the CVF protocol is based on the integration of several kinds of features in order to model speech and dysfluencies: linear and non-linear. Besides, this approach is based on the description of speech pathologies with regard to phonation, articulation, speech quality, human perception, and the complex dynamics of the system. In this work, we will use some of the most efficient speech features for differentiation between healthy and pathological speech in the state of the art (Gómez-Vilda et al., 2015; López-de-Ipiña et al., 2016; López-de-Ipiña et al., 2017). Most of them are well known in the field of speech signal processing,

Figure 1: Two examples from the Categorical Verbal Fluency (CVF) task by a normative individual (left) and by an individual of the MCI group (right). Time domain (top) and spectral contents (bottom) are shown.
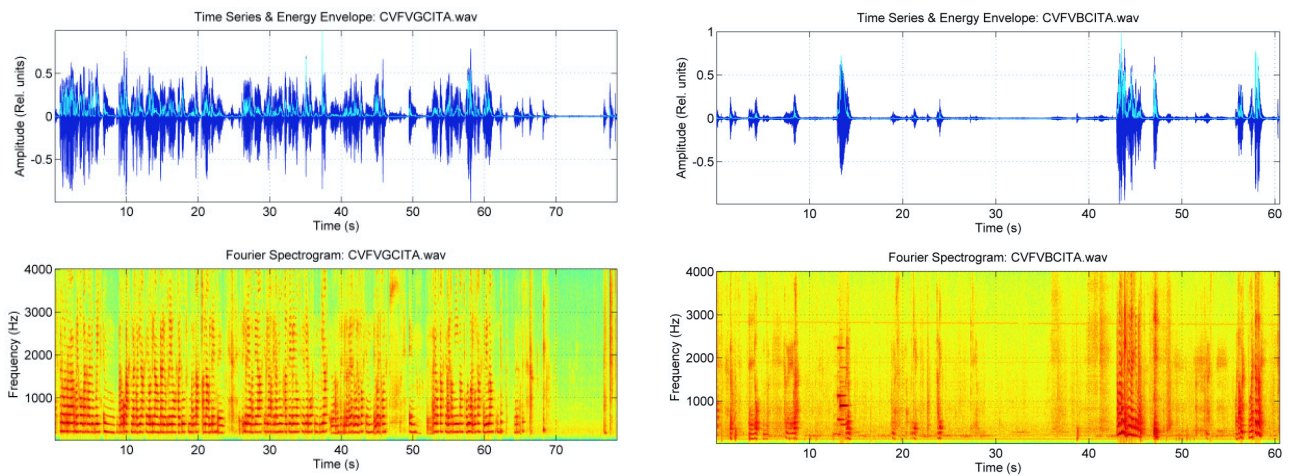
Table 1: Age distribution of the subsets selected for the experimentation (AN: Animal Naming).

	PGA-OREKA AN task-subset (CR/MCI)
Females	36/21
Males	26/17
Age range	39-74/42-79
Age mean	56.73/57.15
Age std. dev.	7.8/8.9

and thus for each parameter a reference is provided where a deeper description and further information can be found. All features are calculated using a software tool developed in our research group using Praat (Boersma and Weenink, 2017) and MATLAB (MATLAB, 2017) (López-de-Ipiña et al., 2017; Mekyska et al., 2015).

4.1. Automatic dysfluency segmentation

The recordings have been automatically segmented in speech signal and dysfluencies by means of a Voice Activity Detection (VAD) algorithm (Solé-Casals & Zaiats, 2010).

4.2. Feature extraction

In this approach the extracted features are the same for both genders; moreover, the approach is independent of the language, so that it is appropriate for a multilingual environment and it is also flexible with regard to the task. The final selected feature test will be different according to the used task. This is the feature set used in the evaluation of speech quality under the aforementioned premises:

- *Classical features (CF):*
  1. Spectral domain features: harmonic contents, harmonic-to-noise ratio (HNR), noise-to-harmonic ratio (NHR), pitch, jitter, shimmer, Amplitude Perturbation Quotient (APQ), spectrum centroids and formants and their variants (mean, median, min, max, mode, standard deviation).
  2. Time domain features: voiced/unvoiced segments, breaks, Zero-Crossing Rate (ZCR) and their variants.
  3. Energy, intensity, short time energy, and spectrum centroids.
- *Perceptual features (PF):*
  1. Mel-Frequency Cepstral Coefficients (MFCC): as the human ear behaves as a filter concentrating only on certain frequency components, the logarithm of the power spectrum of the signal is processed by a bank of filters simulating this behavior. The cosine transform encodes the frequency contents in a vector that constitutes a compact representation of the signal. These filters are non-uniformly spaced on the frequency axis. There are more filters at low frequencies and fewer filters at high frequencies. Mel-Frequency analysis and specifically MFCC are oriented to simulate auditory perception.
  2. Coefficients that provide information related to voice quality, perception, adaptation or amplitude modulation (Mekyska et al., 2015): Modulation Spectra Coefficients (MSC) provide complementary information to MFCC; Perceptual Linear Predictive coefficients (PLP) take into account an adjustment to the equal loudness curve and intensity-loudness power law; Linear Predictive Cepstral coefficients (LPCC); Linear Predictive Cosine Transform coefficients (LPCT) and Adaptive Component Weighted coefficients (ACW), which are less sensitive to channel distortion, and Inferior Colliculus Coef-

ficients (ICC), which analyze amplitude modulations in voice using a biologically inspired model of the inferior colliculus. These features are sometimes extended by their first and second order time derivatives (Δ and $\Delta\Delta$, respectively).
- *Non-linear features (NLF):*
  1. Fractal Dimension, Shannon Entropy, and Multiscale Permutation Entropy have been calculated (López-de-Ipiña et al., 2017).

4.3. Automatic selection of features by Mann-Whitney U-test

In this step the best features are automatically selected in regard to a common significance level. Thus, automatic feature selection is performed by Mann-Whitney U-test with a p value < 0.1 in order to get a larger set for the second phase of feature selection (MATLAB, 2017).

4.4. Automatic selection of features by WEKA

Then a new selection phase is carried out using WEKA, which is a collection of machine learning algorithms for data mining tasks (WEKA, 2017). The functional used is the *SVMAttributeEval* algorithm. This provides a selection by analyzing the feature set.

4.5. Feature normalization by WEKA

During the preprocessing of data, all the features are normalized by WEKA, which is a standard software package.

4.6. Automatic classification

Four classifiers have been used: k-Nearest Neighbors (k-NN), Support Vector Machines (SVM's), Multilayer Perceptrons (MLP's) with L layers and N neurons, and a Convolutional Neural Network (CNN) with L layers of N neurons, a convolution of cxc and a pool of pxp (Eibe et al., 2016).

The WEKA software suite (WEKA, 2016) was used to carry out the experiments. The results were evaluated using the merit figure known as Accuracy (%). For training and validation steps we used k-fold cross-validation with k = 10 (Picard & Cook, 1984). The results of the validation process are summarized in Figure 2.

5. RESULTS AND DISCUSSION

5.1. Animal Names

Around 40 speech samples were used in the experimentation for the MCI group and 60 for the control group

Figure 2: Accuracy (%) for classes and selected classifiers: k-Nearest Neighbors (KNN), Support Vector Machines (SVM), Multilayer Perceptron (MLP) with L layers and N neurons and a Convolutional Neural Network (CNN).
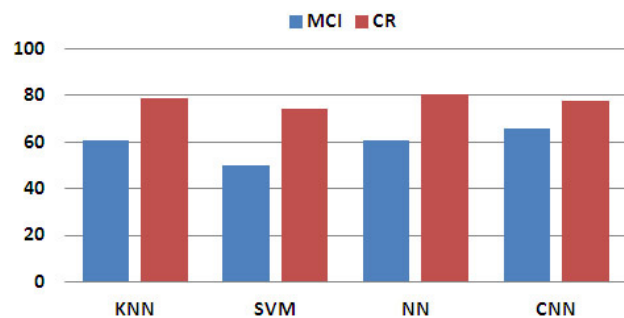
(CR), both from PGA-OREKA (see Table 1). Initially, the number of features included in the study was about 920 (473 for speech and 447 for dysfluencies), obtained at a sampling frequency of 22.05 kHz. After a normalization step, an automatic feature selection was performed based on a non-parametric Mann-Whitney U-test with a p value < 0.1, after which the feature set was reduced to 150. In the second optimization step, the attribute selection algorithm *SVMAttributeEval* of WEKA yielded about 80 features that are finally used in the experiments. The proposed feature set includes features from all the feature types described in Subsection 4.2 for speech and dysfluencies.

Figure 2 shows the results of the automatic separation in two classes, MCI and CR. The merit figure Accuracy (%) is evaluated for all the classifiers detailed in Subsection 4.6. The integration of dysfluency analysis outperforms previous results (López-de-Ipiña et al., 2015) for most classifiers. The results are acceptable, robust, and balanced for all the classifiers (average of 75 %). The deep learning option with CNN yields the best results for a configuration of 2 layers of 20 neurons, a convolution of 3×3 and a pool of 2×2. This option outperforms MLP for 2 layers of 100 neurons.

5.2. Comparing the performance of AN, PD and SS classifiers

The evaluation of the models has been carried out with the balanced subsets: PGA-OREKA, MINI-PGA and AZTIAHORE.

In the first stage all the 473 features described in Section 3 were extracted: (a) 70 classic, (b) 60 perceptual, (c) 180 advanced perceptual and (d) 30 non-linear. Then, an automatic feature selection was carried out and the number of features was reduced to around 60 % for Animal Naming (AN), 50 % for Picture Description (PD), and 40 % for Spontaneous Speech (SS). Finally, an automatic classification by the classifiers was performed by means of a cross-validation (for the tasks SS, AN and PD) as described in Section 3. The attained global classification accuracy rates for the three tasks (%Accuracy) are shown in Table 2.

Table 2: Recognition rate accuracy (%Acc) for the three tasks with the best classifiers.

Task	%Acc
SS	95
AN	80
PD	94

Summarizing the results in the table:

- The best global results are obtained for the Spontaneous Speech (SS) task (Accuracy=95%). This is mainly due to the relevant emotional level of the recording obtained in a relaxing atmosphere, the presence of subtle cognitive changes in the signal due to a more open language, and the inclusion of AD patients instead of MCI subjects.
- The Animal Naming task (AN) shows the lowest results (82%). This is mainly due to the restricted task that could be an easier exercise for people with MCI and the inclusion of MCI subjects instead of AD patients.
- In the case of the Picture Description task (PD), despite the small size of the set, the results are significantly good: 94%.
- The results for the tasks AN and PD are promising and require validation on larger sample sets.

6. CONCLUSIONS

The automatic integration of the most relevant features by Convolutional Neural Networks (CNN) provides useful information not available from statistical tests. This work presents a novel proposal based on a platform *(ALZUMERIC)* used to collect and automatically analyze speech and dysfluencies in order to support MCI and AD diagnosis. Regarding the platform the following conclusions may be drawn:

- The system is friendly and easy to use.
- It is a valuable platform to objectify the patient's state and evolution for the clinician.
- The novel specific features (linear and non-linear) are very useful for a multilingual and multicultural environment, because they are independent from the language.
- The configurable options are very useful, especially for health professionals.
- The results are easy to interpret and manage.
- In this approach the extracted features are the same for both genders.
- It is also flexible with regard to the task.
- The final selected feature test will be different with regard to the used task.

Regarding the analysis methods, it may be said that a non-linear multi-feature modeling is presented based on

the selection of the most relevant features by statistical tests (under medical criteria) and automatic attribute selection: Mann-Whitney U-test and Support Vector Machine Attribute (SVM) evaluation. The approach includes deep learning by means of CNN. The results are encouraging and open a new research perspective.

REFERENCES

Alzheimer's Disease International (2015). World Alzheimer Report 2015. Retrieved from https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf

Boersma, P., & Weenink, D. (2017). *Praat: doing phonetics by computer* [Computer program]. Retrieved from http://www.praat.org

CITA-Alzheimer Foundation, PGA Project (2017). Retrieved from http://www.cita-alzheimer.org/investigacion/proyectos

Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is 'Huh?' a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PLOS ONE, 8*(11): e78273. doi.org/10.1371/journal.pone.0078273

Eibe, F., Mark, A., Hall, I., & Witten, H. (2016) (4th ed). The WEKA Workbench. Online Appendix for I. Witten, E. Frank, M. Hall, & C. Pal (Eds.), *Data mining: Practical machine learning tools and techniques.* New York: Elsevier, Morgan Kaufmann.

Faundez-Zanuy, M., Hussain, A., Mekyska, J., Sesa-Nogueras, E., Monte-Moreno, E., Esposito, A., … Lopez-de-Ipiña, K. (2013). Biometric applications related to human beings: There is life beyond security. *Cognitive Computation, 5*(1), 136-151. doi.org/10.1007/s12559-012-9169-9

Gómez-Vilda P., Rodellar-Biarge, V., Nieto-Luis, V., López-de-Ipiña, K., Álvarez-Marquina, A., Martínez-Olalla, R., … Martínez-Lage, P. (2015). Phonation biomechanic analysis of Alzheimer's Disease cases. *Neurocomputing, 167,* 83-93. doi.org/10.1016/j.neucom.2015.03.087

Klimova, B., Maresova, P., Valis, M., Hort J., & Kuca, K. (2015). Alzheimer's disease and language impairments: Social intervention and medical treatment. *Clinical Interventions in Aging, 10,* 1401–1408. doi.org/10.2147/CIA.S89714

Laske, C., Sohrabi, H. R., Frost, S. M., López-de-Ipiña, K., Garrard, P., Buscema, M. … O'Bryant, S. E. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer & Dementia, 11*(5), 561-578. doi.org/10.1016/j.jalz.2014.06.004

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012) (5th ed.). *Neuropsychological assessment.* Oxford: Oxford University Press.

López-de-Ipiña K., Alonso, J.-B., Manuel Travieso, C., Solé-Casals, J., Egiraun, H., Faundez-Zanuy, M., … Martinez de Lizardui, U. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors, 13,* 6730–6745. doi.org/10.3390/s130506730

López-de-Ipiña, K., Martínez-de-Lizarduy, U., Barroso, N., Ecay-Torres, M., Martínez-Lage, P., Torres, F., & Faundez-Zanuy, M. (2015). Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impartment detection: A non-linear language independent approach. (2015). *4th International Work Conference on Bioinspired Intelligence (IWOBI)*, 1, 101-104. doi.org/10.1109/IWOBI.2015.7160151

López-de-Ipiña, K., Martinez de Lizarduy, U., Calvo, P., Beita, B., García-Melero, J., Ecay-Torres, M., … Faundez-Zanuy, M.

(2017). Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach. *4th International Work Conference on Bioinspired Intelligence (IWOBI)*, 1, 1-4.

López-de-Ipiña K., Satue-Villar, A., Faundez-Zanuy, M., Arreola, V., Ortega, O., Clavé, P. … Calvo, P. (2016). Advances in a multimodal approach for dysphagia analysis based on automatic voice analysis. In S. Bassis, A. Esposito, F. Morabito, & E. Pasero (Eds.), *Advances in Neural Networks*. *Smart Innovation, Systems and Technologies, Vol.* 54 (pp. 201-211). Cham: Springer. doi.org/10.1007/978-3-319-33747-0_20

MATLAB (2017). www.mathworks.com

Mekyska, J., Janousova, E., Gomez-Vilda, P., Smekal, Z., Rektorova, I., Eliasova, I., … López-de-Ipiña, K. (2015). Robust and complex approach of pathological speech signal analysis. *Neurocomputing, 167,* 94-111. doi.org/10.1016/j.neucom.2015.02.085

Picard, R. R., & Cook, D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association, 79*(387), 575–583. doi.org/10.1080/01621459.1984.10478083

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1997). The psychological construct of word fluency. *Brain and Language, 57,* 394-405. doi.org/10.1006/brln.1997.1755

Solé-Casals, J., & Zaiats, V. (2010). A non-linear VAD for noisy environments. *Cognitive Computation, 2*(3), 191-198. doi.org/10.1007/s12559-010-9037-4

WEKA (2017). Retrieved from http://www.cs.waikato.ac.nz/ml/weka