

DANIELA BELTRAMI, LAURA CALZÀ, GLORIA GAGLIARDI, ENRICO GHIDONI,
NORINA MARCELLO, REMA ROSSINI FAVRETTI, FABIO TAMBURINI

Strumenti di *screening* linguistico per l'identificazione precoce della fragilità cognitiva

This paper presents the preliminary results of the OPLON project, co-funded by the Ministry of Education as part of the contract "Smart Cities and Communities and Social Innovation". It aimed at identifying early linguistic symptoms of cognitive decline in the elderly. This pilot study was conducted on a corpus composed of spontaneous speech sample collected from 20 subjects, who underwent a neuropsychological screening for visuo-spatial abilities, memory, language, executive functions and attention. A rich set of linguistic features was extracted from the digitalized utterances (at phonetic, suprasegmental, lexical, morphological and syntactic levels) and the statistical significance in pinpointing the pathological process was measured. Our results, far from being complete, show remarkable trends for what concerns both the linguistic traits selection and the automatic classifiers building.

1. *Introduzione*

Il progetto OPLON ["OPportunities for active and healthy LONGevity", Smart Cities and Communities – DD 391/RIC] si pone l'obiettivo di sviluppare strumenti di diagnosi precoce della fragilità finalizzati alla prevenzione del declino e alla promozione della salute dei soggetti anziani.

Uno dei fenomeni epidemiologici più rilevanti dal punto di vista sanitario e socio-economico che ha interessato la società occidentale negli ultimi decenni è l'aumento significativo della popolazione anziana: in Italia vi sono oltre un milione di pazienti fragili e tale numero è destinato a raddoppiare entro i prossimi vent'anni (ISTAT, 2006).

Tra gli effetti importanti della maggior aspettativa e qualità della vita vi è l'aumento delle patologie neurodegenerative a carattere progressivo, il cui sintomo più rilevante è il decadimento cognitivo. È oramai noto che il processo che conduce alla demenza inizia molti anni prima dell'esordio dei sintomi, quindi molto prima dell'inizio dell'effettiva fase "clinica".

La possibilità di identificare precocemente la malattia si pone come un imprescindibile obiettivo scientifico. Una diagnosi precoce, infatti, non solo permetterebbe di ottimizzare gli interventi assistenziali nei confronti dei pazienti e dei loro familiari (miglior definizione del quadro clinico; accesso precoce ad alcuni trattamenti; programmi di prevenzione sempre più efficaci; maggior *compliance* al percorso terapeutico; miglior risposta ai cambiamenti dello stile di vita; riduzione dei rischi secondari della patologia), ma consentirebbe anche al paziente di pianificare

la propria vita in una fase in cui le capacità sono ancora sostanzialmente mantenute (presa di decisioni riguardanti il futuro; riduzione dello stigma sociale; assunzione di consapevolezza graduale) e agevolerebbe la ricerca e la sperimentazione di terapie in grado di modificare il decorso clinico, fornendo ai pazienti la possibilità di accedere precocemente ai trattamenti farmacologici (i farmaci oggi disponibili agiscono efficacemente soltanto se somministrati in fase altamente precoce).

Vari studi (Kempler, 1987; Caramelli, 1998) mostrano che oltre alla memoria (specialmente quella episodica), anche il linguaggio (soprattutto quello spontaneo) può subire cambiamenti significativi già diversi anni prima dell'esordio clinico della malattia. Tali cambiamenti, tuttavia, non sono identificati facilmente dai test di *screening* tradizionali che, purtroppo, prestano insufficiente attenzione all'eloquio spontaneo (Lezak, 1995) e dovrebbero essere valutati con strumenti alternativi costruiti *ad hoc*.

2. Prospettive linguistico-computazionali

Nella letteratura specifica degli ultimi anni si rilevano lavori che si occupano dell'identificazione automatica dei vari stadi e tipi di demenze a partire dall'analisi statistica della produzione linguistica spontanea dei pazienti. La maggior parte di questi studi, quasi totalmente rivolti all'analisi della lingua inglese, assume come base d'indagine la rilevazione di opportuni tratti linguistici negli enunciati dei pazienti, identificati di solito con strumenti automatici, tentando di correlare tali tratti con i risultati dei test cognitivi utilizzati in ambito neuropsicologico per l'identificazione precoce di disturbi cognitivi (Lehr, 2012; Roark et al., 2011; Jarrold et al., 2010; Satt et al., 2013; Chapman et al., 2002).

Si è giunti all'elaborazione di metodologie automatiche e di procedure di analisi che aprono nuove prospettive di progettazione metodologica e nuovi obiettivi. In certa misura, tuttavia, può restare aperta una domanda sulla legittimità della correlazione fra piano linguistico e piano cognitivo. Questa trova ragione, a nostro avviso, in una pluralità di ambiti della linguistica, della psicologia, della neurologia e, in particolare, della neurolinguistica. Ambiti e ricerche che affondano le loro radici in tempi che appaiono lontani. Doveroso, in questa sede, appare far precedere alla presentazione del contributo un riferimento a Jakobson e agli studi da lui condotti sul linguaggio patologico, in cui l'afasia è indagata non di per sé, ma in relazione al farsi e al disfarsi del linguaggio, come campo produttivo per lo studio linguistico nella sua globalità (Jakobson, 2006; Jakobson, Halle, 1956).

L'opera di Jakobson, letta alla luce delle procedure sviluppate nell'ambito delle tecnologie informatiche si pone, oggi, come fondante nelle intuizioni e nei contributi. Che la patologia del linguaggio obbedisca a un insieme di regole è asserito fin dalle prime fasi della sua ricerca, così come che nessuna regola sottesa alla regressione del linguaggio possa essere individuata prescindendo dai prin-

cipi, dalle metodologie e dalle procedure proprie della linguistica. Come ben si chiarisce nelle seguenti parole:

Se l'afasia è un disturbo del linguaggio, come dice il termine stesso, ne consegue che ogni descrizione e classificazione dei disturbi afasici deve preliminarmente chiarire quali aspetti del linguaggio sono alterati nei diversi tipi di questo disturbo. (Jakobson, 1956: 22)

Costante, nella ricerca jakobsoniana, è l'asserzione dell'importanza di un'analisi dei deficit linguistici che sorgono in concomitanza con danni cerebrali e della correlazione che fra questi si instaura. Correlazione evidenziata già negli studi ancora precedenti di Hughlings Jackson, ricordato da Jakobson come "primo vero interprete dell'afasia", ed oggetto di analisi nelle ricerche di Kurt Goldstein, André Ombredane e Alexander R. Luria, indicati come autori dei migliori trattati apparsi negli ultimi anni '40 sull'afasia, nei cui libri è possibile trovare i primi sforzi dei neurologi verso l'utilizzazione sistematica dei principi linguistici moderni per l'analisi delle sindromi afasiche. Così come è ripetutamente asserita l'opportunità, se non la necessità, di un approccio interdisciplinare in cui in cui sia presente l'apporto della teoria matematica della comunicazione e della teoria dell'informazione, e in cui possa costantemente accrescersi la cooperazione di linguisti, psicologi, psichiatri, neurologi. Nella nuova dimensione di analisi un ruolo primario spetterebbe alla linguistica quale punto di convergenza di ambiti disciplinari diversi¹.

Con le opportune modifiche queste affermazioni, formulate intorno alla metà del secolo scorso, potrebbero essere accettate ancora oggi. Certo, gli ambiti di indagine nei gruppi di lavoro si sono modificati. Oggi facciamo riferimento al Natural Language Processing, alle biotecnologie, alle neuroscienze, alle scienze cognitive. Al generico concetto di afasia si sostituisce in letteratura una classificazione più specifica dei disturbi del linguaggio. In particolare appaiono superati i limiti del modello afasiologico: già da alcuni decenni, con gli sviluppi delle neuroscienze e con il diffondersi delle tecniche di neuroimmagine, si è rilevata la necessità di operare modifiche alla rigidità posta dal modello classico nella localizzazione e nell'organizzazione cerebrale delle funzioni linguistiche (Bambini, 2009). Nuovi aspetti sono presi in esame, come la modalità o i livelli di strutturazione; sono indicati nuovi oggetti di ricerca, ma non è confutata la validità che i risultati ottenuti nello studio delle afasie mantengono sia a livello linguistico sia, in una dimensione multidisciplinare, clinico. Non è negata la rilevanza che gli studi linguistici possono assumere nell'analisi del linguaggio patologico ma appare dimenticata una lunga tradizione risalente al secolo scorso, che può essere

¹ "Mi sembra che siano ormai pochi i ricercatori nel campo dei disturbi del linguaggio che credono ancora alla non rilevanza del ruolo della linguistica nello studio dell'afasia. [...] Vi sono parecchi centri in cui neurologi, psicologi, linguisti, e altri specialisti, lavorano insieme per descrivere, esaminare, analizzare l'afasia, e ottenere le diagnosi e prognosi più esatte." (Jakobson, 2006: 171).

riduttivo fare risalire a Jakobson ma che in Jakobson certamente trova uno studioso particolarmente attento agli aspetti teorici e metodologici.

Negli ultimi decenni si è sviluppato un nuovo interesse verso la possibilità di giungere all'individuazione delle caratteristiche distintive del linguaggio dei soggetti con diverso grado di fragilità cognitiva utilizzando anche approcci di analisi delle produzioni verbali spontanee (March et al., 2003). Vari ed importanti sono stati gli studi sull'analisi del discorso e del testo². Sono state oggetto di indagine le abilità conversazionali e comunicative e aspetti più specifici come la ricchezza lessicale, la complessità sintattica, il livello di competenza pragmatica. Si è configurato un panorama alquanto articolato che mostra, ad esempio, come la struttura sintattica sia meglio preservata della lessico-semantica (Kempner et al., 1987) e come alcune alterazioni linguistiche siano apprezzabili a livello di discorso (Caramelli et al., 1998) nei soggetti affetti da demenza di Alzheimer. L'eloquio di tali pazienti, descritto in letteratura come tangenziale e incoerente, è caratterizzato in particolare dall'abuso di termini deittici e pronomi senza antecedente e da un aumento di revisioni e ripetizioni, a cui si accompagnano difficoltà nel *retrieval* lessicale e deficit delle abilità pragmatiche (Taler, Phillips, 2008).

Tuttavia, occorre ammettere che, pur nella varietà e molteplicità delle ricerche, i contributi che sono stati dati nell'ambito della linguistica appaiono insufficienti nella definizione di un quadro concettuale di riferimento.

Nuove prospettive sembrano oggi configurarsi nell'attuale alto interesse verso tecniche di analisi computerizzate: questo ha portato alla creazione di "corpora linguistici elettronici", intendendo con ciò collezioni di testi in forma elettronica scelti in virtù della loro rappresentatività nel caratterizzare una particolare lingua o varietà linguistica, e allo sviluppo di numerosi algoritmi per l'analisi e la classificazione del "parlato".

Applicato al linguaggio patologico, questo approccio e le tecnologie relate hanno il notevole vantaggio di rappresentare una registrazione del linguaggio naturale e spontaneo, al di fuori quindi degli schemi dei test neuropsicologici del linguaggio, potenzialmente applicabili ad ampie fasce di popolazione con strumenti a basso costo (Lehr, 2012).

L'analisi computazionale di dati linguistici e paralinguistici può dare un contributo rilevante allo *screening* delle patologie neurodegenerative di tipo demenziale: tecniche proprie del Natural Language Processing sono state recentemente utilizzate con successo nell'analisi automatica di testi scritti, testi orali elicitati in contesto clinico controllato e produzioni spontanee di soggetti affetti da demenza, in particolare Alzheimer e Demenza Fronto-Temporale (Afasia Primaria progressiva e Demenza Semantica) (Peintner et al., 2008; Jarrold et al., 2010; Roark et al., 2011; Fraser et al., 2014).

² "In clinical groups, discourse research can provide important insights into the pragmatic aspects of language changes and the role of non-linguistic factors in communication (e.g., communicative context, conversational partner) as associated to brain damage." (March et al., 2006: 311-312).

Le prospettive di analisi computazionali coprono vari ambiti linguistici, spaziando dall'esame approfondito della produzione orale spontanea alla ricerca di disfluenze, incertezze e/o modificazioni nella sequenza fonica e nei parametri soprasegmentali fino a un'approfondita analisi lessicale, morfosintattica e sintattica con l'obiettivo di individuare regolarità che evidenzino una maggiore semplificazione dell'eloquio nei pazienti rispetto ai controlli. La letteratura ha prodotto, negli ultimi anni, una moltitudine di indicatori, misurabili spesso con strumenti automatici, che, da studio a studio, si sono rivelati significativi nel discriminare tra popolazioni diverse di soggetti. Tali metodologie si sono quindi dimostrate in grado di rilevare *pattern* latenti e regolarità nel linguaggio di queste popolazioni di pazienti, potenzialmente utili all'identificazione precoce, classificazione e descrizione del decadimento cognitivo (Elvevåg, Garrard, 2014).

Date queste premesse, l'obiettivo a lungo termine del progetto OPLON riguarda lo sviluppo di tecniche facilmente somministrabili di raccolta e analisi di campioni di linguaggio spontaneo per lo screening sistematico della popolazione potenzialmente a rischio. Tali strumenti devono essere economici, semplici e utilizzabili anche in ambiti non specializzati (Medicina di Base, ad esempio).

A medio termine, questo studio si propone di dimostrarne la fattibilità in ambiente controllato, visto che in letteratura non risultano al momento studi simili sulla lingua italiana.

3. Costruzione, annotazione e analisi del corpus

3.1 Raccolta dati

La ricerca prevede il reclutamento di novanta soggetti bilanciati per sesso, età (*range* 50-75) e scolarità (licenza media o licenza elementare con buona stimolazione intellettuale nel corso della vita), di cui quarantacinque soggetti di controllo (neurologicamente indenni) e altrettanti affetti da declino cognitivo. Tale declino fa riferimento a due diverse categorie:

1. Mild Cognitive Impairment (MCI): deterioramento cognitivo lieve non in grado di interferire in modo significativo con le attività della vita quotidiana, riguardante una singola funzione cognitiva (ad es. memoria) o più funzioni (ad es. memoria, attenzione e linguaggio). I soggetti con MCI sono stati divisi in due sottogruppi:
 - a. MCI con isolato deficit di memoria (a-MCI single domain; n = 15);
 - b. MCI con deficit di funzioni cognitive ma non di memoria (non a-MCI *single or multiple domain*; n = 15).

In entrambi i casi la compromissione cognitiva non interferisce con il funzionamento quotidiano. Tale differenziazione (a-MCI vs. non-a-MCI) è stata fatta poiché sembra che diversi tipi di MCI possano essere prodromici di differenti tipi di demenza (*Alzheimer Disease* – AD, demenza vascolare, demenza fronto-temporale, ecc.). La variante amnesica, ad esempio, sembrerebbe

evolvere più frequentemente verso una demenza (AD in particolare) rispetto ad altri tipi di MCI.

2. Demenza in fase iniziale (*early-D*; $n = 15$): i soggetti appartenenti a questo gruppo hanno ricevuto la diagnosi al termine di un articolato iter diagnostico. In questo caso i deficit cognitivi influenzano (sebbene ancora solo parzialmente) le autonomie nelle attività della vita quotidiana (ad es. assunzione autonoma della terapia farmacologica, cura personale, cura dell'ambiente domestico, ecc.). Un ulteriore requisito per fare parte di questo gruppo sperimentale era quello di ottenere un punteggio maggiore o uguale a 18 al MMSE (Mini Mental State Examination).

Ai soggetti sono stati somministrati i test neuropsicologici più efficaci nella discriminazione tra soggetti normali, con MCI e con Demenza. Tali prove, infatti, valutano le funzioni che precocemente subiscono variazioni in seguito ad un processo neurodegenerativo (abilità verbali, esecutive, visuo-spaziali, attentive, mnesiche e di orientamento (Chen et al., 2001; Backman et al., 2005; Grober et al., 2008; Ismail et al., 2010)), e risultano essere le più efficaci nella discriminazione tra soggetti normali, con demenza e Mild Cognitive Impairment. In particolare sono stati somministrati:

- MMSE - Mini-Mental State Examination (Measso et al., 1993; Magni et al., 1996);
- MoCA - Montreal Cognitive Assessment (Nasreddine et al., 2005; Pirani et al., 2006);
- GPCog - General Practitioners assessment of Cognition (Brodaty et al., 2002; Pirani et al., 2010);
- CDT - Clock Drawing Test (Freedman et al., 1994; Mondini et al., 2003);
- Fluenze Verbali (fluency fonemica e semantica (Carlesimo et al., 1996; Novelli et al., 1986));
- Paired Associate Learning (PAL, subtest del Cambridge Neuropsychological Test Automated Battery – CANTAB (Junkkila et al., 2012)).

È stato presentato anche un questionario di valutazione dell'indice di Riserva Cognitiva (Cognitive Reserve Index Questionnaire, CRI-q (Nucci et al., 2011)).

Dopo la valutazione tradizionale, i soggetti sono stati sottoposti ad una registrazione dell'eloquio spontaneo durante alcuni compiti:

1. “Descrivi questa immagine”. L'immagine (tratta da Esame del Linguaggio – II (Ciurli et al., 1996)), in bianco e nero, ritrae un salotto nel quale vi sono una donna, un uomo, due bambini ed un gatto che compiono alcune semplici azioni (ad es. guardare la televisione, giocare con dei cubetti, ecc.).
2. “Descrivi una giornata lavorativa”;
3. “Descrivi l'ultimo sogno che ricordi”.

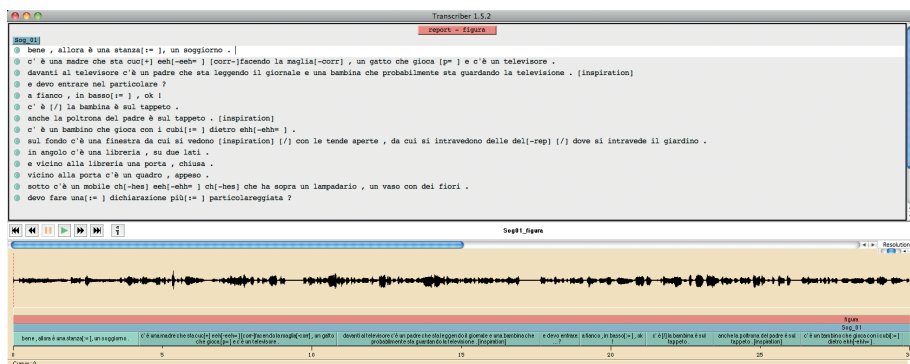
Oltre all'analisi linguistica, questi compiti permettono di valutare l'eventuale compromissione delle funzioni mnesiche; infatti, in tutte e tre le prove è importante ricordare ciò che si sta dicendo (memoria di lavoro), che si è già riferito (memoria episodica) o si vorrà raccontare (memoria prospettica); inoltre, è importante conoscere il significato dei termini utilizzati (memoria semantica) ed avere ricordi rispetto al passato più o meno remoto (memoria autobiografica recente e remota).

L'intera valutazione viene ripetuta a distanza di circa un anno.

3.2 Trascrizione e annotazione dei testi orali del corpus

I campioni di linguaggio parlato raccolti nel corso delle sedute di test nella forma di file .WAV (segnale mono-canale digitalizzato a 44.1KHz di frequenza, *range* dinamico di 16 bit) vengono inizialmente sottoposti a trascrizione ortografica manuale. Le trascrizioni, conformi al "Protocollo di trascrizione" definito nell'ambito del progetto e ispirato alle *guidelines* del progetto CLIPS (Albano Leoni, 2003), sono realizzate con il software *opensource* Transcriber; il programma genera un output in formato .trs (.xml con allineamento temporale degli enunciati alla waveform) e ne consente l'esportazione in formato plain-text (si veda la fig. 1).

Figura 1 - Uno screenshot del programma di trascrizione "Transcriber"

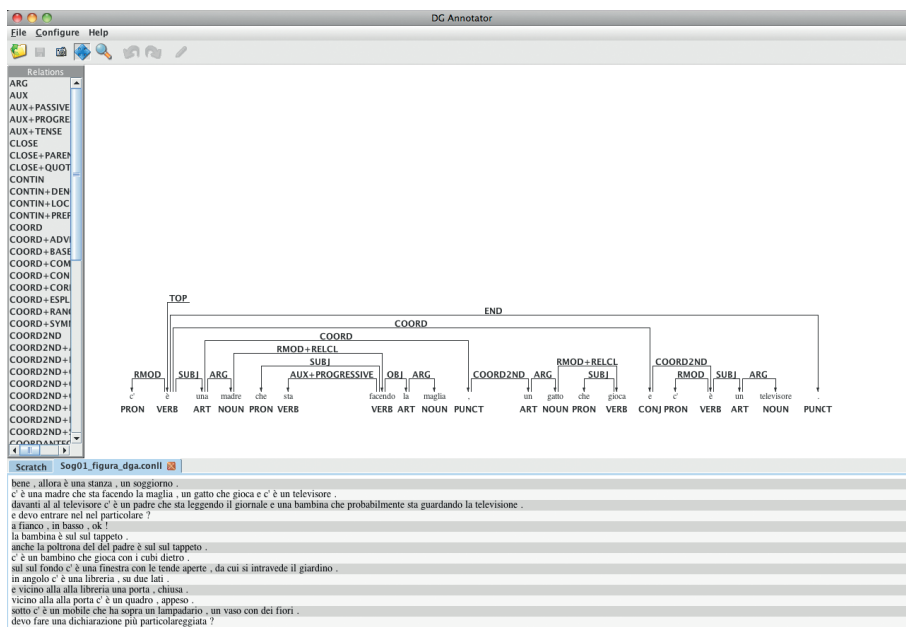


L'unità di allineamento scelta è l'enunciato, definito e identificato sulla base di criteri intonativi. Contestualmente alla trascrizione vengono annotati anche i principali fenomeni paralinguistici (pause, disfluenze, lapsus, etc.).

Vengono quindi isolati i turni di parola del soggetto sottoposto a test: gli enunciati selezionati, emendati manualmente dalle disfluenze, sono sottoposti ad annotazione morfosintattica semi-automatica, ovvero subiscono Part of Speech (PoS) Tagging e Parsing automatico mediante il *parser* a dipendenze TULE (Lesmo, 2007), che assicura prestazioni confrontabili con lo stato dell'arte per la lingua

italiana³. L'annotazione viene revisionata da un linguista esperto utilizzando il software *open-source* DGA – Dependency Grammar Annotator, strumento grafico per l'annotazione sintattica di testi secondo il quadro teorico delle Grammatiche a dipendenze (fig. 2).

Figura 2 - Uno screenshot del programma “DGA” utilizzato per la correzione manuale dell'annotazione sintattica



Il *tagset* adottato è quello del TUT - Turin University TreeBank (Bosco et al., 2000), compatibile con il *parser* TULE. L'annotazione morfosintattica è prodotta nel formato tabulare standard CoNLL (Buchholz, Marsi, 2006).

3.3 Estrazione dei tratti linguistici

Per ciascun livello di analisi vengono calcolati indici linguistici e stilometrici descritti in letteratura (e risultati significativi negli studi condotti su lingue diverse dall'italiano) o creati ex novo. Le misure acustiche vengono calcolate direttamente sul segnale vocale, utilizzando un Voice Activity Detector⁴. I tratti lessicali e sintattici vengono invece derivati dall'annotazione automatica revisionata manualmente dal linguista (si veda la tabella 1).

³ Si rinvia ai dati delle campagne di valutazione EVALITA, Subtask “Dependency Parsing” (<http://www.evalita.it>).

⁴ Il VAD utilizzato nello studio è contenuto nel pacchetto software di Julius ASR (http://julius.osdn.jp/en_index.php), v. 4.3.1, denominato “adintool”.

Tabella 1 - *Elenco dei tratti linguistici considerati nello studio*

<i>Acoustic Features</i>
Silence segments duration: mean, median and Std. Deviation (Satt et al., 2012; 2013)
Speech segments duration: mean, median and Std.Deviation (Satt et al., 2012; 2013)
Temporal regularity of voiced segment durations (Satt et al., 2012; 2013)
Verbal Rate (Singh et al., 2001; Roark et al., 2007; 2011)
Transformed Phonation Rate (Singh et al., 2001; Roark et al., 2011)
Standardized Phonation Time (Singh et al., 2001; Roark et al., 2011)
Standardized Pause Rate (Singh et al., 2001; Roark et al., 2007; 2011)
Root Mean Square energy: mean and Std.Deviation (López-de-Ipiña, 2013)
Pitch: mean and Std. Deviation (López-de-Ipiña, 2013)
Spectral Centroid, mean and Std. Deviation (López-de-Ipiña, 2013)
Higuchi Fractal Dimension, mean and Std. Deviation (López-de-Ipiña, 2013)
<i>Lexical Features</i>
Lexical density: Open-class ratio
Content Density (Roark et al., 2011)
PoS rate (Holmes, Singh, 1996; Bucks et al., 2000; Vigorelli, 2004; Garrard et al., 2005; Thomas et al., 2005; Peintner et al., 2008; Cantos-Gomez et al., 2009; Jarrold et al., 2010; 2014)
Reference Rate to Reality (Vigorelli, 2004)
Personal, Spatial and Temporal Deixis rate (March et al., 2006; Cantos-Gomez et al., 2009)
Relatives pronouns and negative adverbs rate
Lexical Richness: Type-Token Ratio, W - Brunét's Index and R - Honoré's Statistic (Brunét, 1978; Honoré, 1979; Bucks et al., 2000; Holmes, 1992; Holmes, Singh, 1996; Thomas et al., 2005)
Action Verbs rate (Gagliardi, 2014)
Frequency-of-use tagging (De Mauro/Paravia dictionary) (De Mauro, 1980; 2000; Barbagli et al., 2014)
Propositional Idea Density (Snowdon et al., 1996; Brown et al., 2008; Jarrold et al., 2010; Roark et al., 2011)
<i>Syntactic Features</i>
Number of dependent elements linked to the noun, mean
Mean Global Dependency Distance (Roark et al., 2011)
Syntactic complexity (Szmrecsanyi, 2004)
Syntactic embeddedness: maximum depth of the structure, mean and Std. Deviation
Utterance length, mean and Std. Deviation

3.4 Test statistici e classificazione automatica

L'effettiva significatività statistica per la lingua italiana degli indici e delle possibili correlazioni tra variabili è stata valutata attraverso test e metriche statistiche standard (parametriche e non parametriche) e i parametri risultati significativi sono stati ulteriormente testati mediante algoritmi di *feature selection*. Tramite metodologie di *machine learning* gli indici selezionati vengono integrati in un classificatore automatico, le cui performance sono valutate in termini di *accuracy*, *precision*, *recall* e *F-measure*.

Nella fase conclusiva del progetto i risultati del test linguistico verranno unificati in un punteggio di *scoring* sintetico, in grado di esprimere un indicatore di impoverimento del linguaggio dell'individuo sottoposto a test, da integrare nel modello OPLON di predizione del rischio di insorgenza di fragilità nell'anziano.

4. Risultati preliminari e discussione

L'effettiva fattibilità della metodologia è stata verificata mediante un *Proof of Concept* su un campione di 20 soggetti, 12 MCI e 8 controlli.

Tabella 2 - I risultati di significatività statistica e le prestazioni di classificazione per i tratti linguistici maggiormente significativi

Task "Figura"		
		KS test
Tratti selezionati	Silence segment duration (mean)	p = 0.047035
	Spectral Centroid (mean)	p = 0.009033
	Relatives pronouns and negative adverbs rate	p = 0.047035
	Utterance length (SD)	p = 0.047035
Prestazioni del Classificatore	Accuratezza = 0.9000	
	Precision = 1.0000	
	Recall = 0.8000	
	F1 = 0.8889	
Task "Lavoro"		
		KS test
Tratti selezionati	Higuchi Fractal Dimension (SD)	p = 0.047035
	Spectral Centroid (SD)	p = 0.028057
	De Mauro Frequency-of-use labels	p = 0.047035
Prestazioni del Classificatore	Accuratezza = 0.7000	
	Precision = 0.6429	
	Recall = 0.9000	
	F1 = 0.7500	
Task "Sogno"		
		KS test
Tratti selezionati	Utterance length (SD)	p = 0.009033
	Silence segment duration (mean)	p = 0.047035
	Personal Deixis rate	p = 0.076262
Prestazioni del Classificatore	Accuratezza = 0.8500	
	Precision = 0.8182	
	Recall = 0.9000	
	F1 = 0.8571	

Per ciascun *task* del test linguistico sono state identificate le *feature* statisticamente significative nel discriminare le due popolazioni di soggetti (*p-value* < 0.05) applicando il test non parametrico di Kolmogorov-Smirnov, in ragione della ridotta

dimensione del campione, e l'algoritmo di *feature selection* mRMR (minimum-Redundancy-Maximum Relevance) (Peng et al., 2005).

Gli indici selezionati sono stati integrati in un classificatore automatico kNN 3-neighbours. Il classificatore esegue un campionamento *random* dei testi e 5 iterazioni delle fasi di *training* e *testing*; la dimensione relativa del training set è 80%, ovvero 16 soggetti sono stati utilizzati per l'addestramento del sistema e 4 per il test.

È evidente che la ridotta quantità di dati non permette l'applicazione produttiva di queste metodologie, se non per ricercare tendenze e informazioni al fine di predisporre metodologie più adeguate in vista del campione completo.

Le prestazioni del classificatore in relazione a ciascun *task* sono sintetizzate nella tabella 2:

4. Conclusioni e sviluppi futuri

I risultati presentati, ancorché preliminari, appaiono incoraggianti. Pur utilizzando un set di dati estremamente ridotto, l'esperimento ha permesso di saggiare la validità della metodologia proposta e ha fornito indicazioni preliminari sulla significatività che alcuni degli indici linguistici presentano nei testi in esame. Ad oggi non risultano presenti studi simili sull'italiano.

Nelle fasi prodromiche delle malattie dementigene è già presente un deficit linguistico, seppur subliminale; tale deficit, identificabile con strumenti automatici, non è circoscrivibile ad un singolo livello ma, come si può notare dai tratti risultati statisticamente significativi nei diversi *task*, interessa il piano acustico, lessicale e sintattico.

Per confermare tali tendenze è però assolutamente necessaria una maggior quantità di dati, organizzati in un corpus adeguatamente bilanciato. Il reclutamento dei pazienti è, al momento, ancora in corso.

Si prevede inoltre di identificare ed implementare nuovi indici linguistici, e testare nuovi algoritmi di *feature selection* e classificatori ML (ad esempio *neural network*, *Support Vector Machine*, *Quantum Classifiers*).

Bibliografia

- ALEGRIA, R., GALLO, C., BOLSO, M., DOS SANTOS, B., PRISCO, C.R., BOTTINO, C. & NOGUEIRA, M.I. (2013). Comparative study of the uses of grammatical categories: adjectives, adverbs, pronouns, interjections, conjunctions, and prepositions in patients with Alzheimer's disease. In *Alzheimer's & Dementia: The journal of the Alzheimer's Association*, 9 (4), 882.
- ALBANO LEONI, F. (2003). Tre progetti per l'italiano parlato. In MARASCHIO, N., POGGI SALANI, T. (Eds.), *Italia linguistica anno Mille. Italia linguistica anno Duemila. Atti del XXXIV Congresso Internazionale di studi della Società di linguistica italiana (SLI)*. Roma: Bulzoni, 675-683.

- ALEXOPOULOS, P., GRIMMER, T., PERNECZKY, R., DOMES, G. & KURZ, A. (2006). Progression to Dementia in Clinical Subtypes of Mild Cognitive Impairment. In *Dementia and geriatric cognitive disorders*, 22, 27-34.
- BÄCKMAN, L., JONES, S., BERGER, A.K., LAUKKA, E.J. & SMALL, B.J. (2005). Cognitive Impairment in Preclinical Alzheimer's Disease: A Meta-Analysis. In *Neuropsychology*, 19 (4), 520-531.
- BAMBINI, V. (2009). Reti metaforiche: il contributo delle neuroimmagini alla modellizzazione della pragmatica. In FERRARI, G., BENATTI, R. & MOSCA, M. (Eds.), *"Linguistica e modelli tecnologici di ricerca": Atti del XXXIV Congresso Internazionale di studi della Società di Linguistica Italiana (SLI)*. Roma: Bulzoni, 33-45.
- BAMBINI, V., RICCI, I. & BERTINETTO, P.M. (Eds.) (2012). *"Linguaggio e cervello - Semantica / Language and the brain - Semantics": Atti del XLII Convegno della Società di Linguistica Italiana (SLI)*. Roma: Bulzoni.
- BARBAGLI, A., LUCISANO, P., DELL'ORLETTA, F., MONTEMAGNI, S. & VENTURI, G. (2014). Tecnologie del linguaggio e monitoraggio dell'evoluzione delle abilità di scrittura nella scuola secondaria di primo grado. In BASILI, R., LENCI, A. & MAGNINI, B. (Eds.), *Proceeding of the First Italian Conference on Computational Linguistics (CLiC-IT 2014)*, 23-27.
- BOSCO, C., LOMBARDO, V., VASSALLO, D. & LESMO, L. (2000). Building a Treebank for Italian: a data-driven annotation schema. In *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC-2000)*, ELRA - European Language Resources Association.
- BRODATY, H., POND, D., KEMP, N.M., LUSCOMBE, G., HARDING, L., BERMAN, K. & HUPPERT, F.A. (2002). The GPCOG: A New Screening Test for Dementia Designed for General Practice. In *Journal of the American Geriatrics Society*, 50 (3), 530-534.
- BROWN, C., SNODGRASS, T., KEMPER, S.J., HERMAN R. & COVINGTON, M.A. (2008). Automatic measurement of Propositional Idea Density from Part-of-Speech Tagging. In *Behavior Research Methods*, 40 (2), 540-545.
- BRUNET, E. (1978). *Le Vocabulaire de Jean Giraudoux. Structure et Evolution*. Geneve: Slatkine.
- BUCKS, R.S., SINGH S., CUERDEN J.M. & WILCOCK, G.K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. In *Aphasiology*, 14 (1), 71-91.
- BUCHHOLZ, S., MARSI, E. (2006). CoNLL-X shared task on Multilingual Dependency Parsing. In *Proceedings of CoNLL-X, ACL - Association for Computational Linguistics*, 149-164.
- CANTOS-GÓMEZ, P. (2009). Featuring linguistic decline in Alzheimer's disease: A corpus-based approach. In MAHLBERG, M., GONZÁLEZ DÍAZ, V. & SMITH, C. (Eds.), *Proceedings of the Corpus Linguistics Conference 2009 (CL2009)*.
- CARAMELLI, P., MANSUR, L.L. & NITRINI, R. (1998). Language and communication disorders in dementia of the Alzheimer type. In STEMMER, B., WHITAKER, H.A. (Eds.), *Handbook of neurolinguistics*, 463-473.
- CARLESIMO, G.A., CALTAGIRONE, C., GAINOTTI, G. & THE GROUP FOR THE STANDARDIZATION OF THE MENTAL DETERIORATION BATTERY (1996). The Mental

Deterioration Battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. In *European Neurology*, 36, 379-384.

CHAPMAN, S.B., ZIENTZ, J., WEINER, M.F., ROSENBERG, R.N., FRAWLEY, W.H. & BURNS, M.H. (2002). Discourse changes in early Alzheimer disease, Mild Cognitive Impairment, and normal aging. In *Alzheimer Disease & Associated Disorders*, 16 (3), 177-186.

CHEN, P., RATCLIFF, G., BELLE, S.H., CAULEY, J.A., DEKOSHY, S.T. & GANGULI, M. (2001). Patterns of cognitive decline in presymptomatic Alzheimer disease. A prospective community study. In *Archives of general psychiatry*, 58, 853-858.

CIURLI, P., MARANGOLO, P. & BASSO, A. (1996). *Esame del Linguaggio – II*. Firenze: Organizzazioni Speciali.

DE MAURO, T. (1980). *Guida all'uso delle parole*. Editori Riuniti.

DE MAURO, T. (2000). *Il dizionario della lingua italiana*. Paravia.

ELVEVÅG, B., GARRARD, P. (2014). *Language, Computer and Cognitive Neuroscience*. In *Cortex Special Issue*, 55.

FRASER, K.C., MELTZER, J.A., GRAHAM, N.L., LEONARD, C., HIRST, G., BLACK, S.E. & ROCHON, E. (2014). Automated classification of Primary Progressive Aphasia subtypes from narrative speech transcripts. In *Cortex*, 55, 43-60.

FREEDMAN, M., LEACH, L., KAPLAN, E., WINOCUR, G., SHULMAN, K.I. & DELIS, D.C. (1994). *Clock Drawing: A neuropsychological analysis*. Oxford: Oxford University Press.

GAGLIARDI, G. (2014). *Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnosi del Mild Cognitive Impairment*, Tesi di dottorato, Università degli Studi di Firenze.

GARRARD, P., MALONEY, L.M., HODGES, J.R. & PATTERSON, K. (2005). The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author. In *Brain*, 128, 250-260.

GROBER, E., HALL, C.H., LIPTON, R.B., ZONDERMAN, A.B., RESNICK, S.M. & KAWAS, C. (2008). Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. In *Journal of the International Neuropsychological Society*, 14 (2), 266-278.

HOLMES, D.I. (1992). A stylometric analysis of mormon scripture and related texts. In *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 155 (1), 91-120.

HOLMES, D.I., SINGH, S. (1996). A stylometric analysis of conversational speech of aphasic patients. In *Literary and Linguistic Computing*, 11 (3), 133-140.

HONORÉ, A. (1979). Some simple measures of richness of vocabulary. In *Association of Literary and Linguistic Computing Bulletin*, 7, 172-177.

ISMAIL, Z., RAJJI, T.K. & SHULMAN, K.I. (2010). Brief cognitive screening instruments: an update. In *International Journal of Geriatric Psychiatry*, 25 (2), 111-120.

JAKOBSON, R. (2006). *Linguaggio infantile e afasia*. Torino: Einaudi.

JAKOBSON, R., HALLE, M. (1956). *Fundamentals of Language*. Den Haag: Mouton & co.

JARROLD, W.L., PEINTNER, B., YEH, E., KRASNOW, R., JAVITZ, H.S. & SWAN, G.E. (2010). Language Analytics for Assessing Brain Health: Cognitive Impairment, Depression and Pre-symptomatic Alzheimer's Disease. In YAO, Y., SUN, R., POGGIO, T., LIU, J., ZHONG, N. & HUANG, J. (Eds.), *Brain Informatics: International Conference (BI 2010)*. Berlin - Heidelberg: Springer-Verlag, 299-307.

- JARROLD, W.L., PEINTNER, B., WILKINS, D., VERGRY, D., RICHEY, C., GORNO-TEMPINI, M.L. & OGAR, J. (2014). Aided Diagnosis of Dementia Type through Computer-Based Analysis of Spontaneous Speech. In RESNIK, P., RESNIK, R. & MITCHELL, M. (Eds.), *Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality*. ACL - Association for Computational Linguistics, 27-37.
- JUNKKILA, J., OJA, S., LAINE, M. & KARRASCH, M. (2012). Applicability of the CANTAB-PAL computerized memory test in identifying amnesic Mild Cognitive Impairment and Alzheimer's disease. In *Dementia and Geriatric Cognitive Disorder*, 34, 83-89.
- KEMPLER, D., CURTISS, S. & JACKSON, C. (1987). Syntactic preservation in Alzheimer's disease. In *Journal of Speech and Hearing Research*, 30 (3), 343-350.
- LARRIEU, S., LETENNEUR, L., ORGOGOZO, J., FABRIGOULE, C., AMIEVA, H., LE CARRET, N., BARBERGER-GATEAU, P. & DARTIGUES, J. (2002). Incidence and outcome of mild cognitive impairment in a population-based prospective cohort. In *Neurology*, 59 (10), 1594-1599.
- LEHR, M., PRUD'HOMMEAUX, E., SHAFRAN, I. & ROARK, B. (2012). Fully Automated Neuropsychological Assessment for Detecting Mild Cognitive Impairment. In *Proceedings of InterSpeech 2012*.
- LESMO, L. (2007). Il parser basato su regole del Gruppo NLP dell'Università di Torino. In *Intelligenza Artificiale*, IV (2), 46-47.
- LEZAK, M.D. (1995). *Neuropsychological Assessment*. Oxford: Oxford University Press.
- LÓPEZ-DE IPIÑA, K., ALONSO, J.B., TRAVIESO, C.M., SOLÉ-CASALS, J., EGIRAUN, H., FAUNDEZ-ZANUY, M., EZEIZA, A., BARROSO, N., ECAY-TORRES, M., MARTINEZ-LAGE, P. & MARTINEZ DE LIZARDUI, U. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. In *Sensors*, 13, 6730-6745.
- MAGNI, E., BINETTI, G., BIANCHETTI, A., ROZZINI, R. & TRABUCCHI, M. (1996). Mini-Mental State Examination: a normative study in italian elderly population. In *European Journal of Neurology*, 3, 198-202.
- MARCH, E.G., WALES, R. & PATTISON, P. (2003). Language use in normal ageing and dementia of the Alzheimer type. In *Clinical Psychologist*, 7 (1), 44-49.
- MARCH, E.G., WALES, R. & PATTISON, P. (2006). The uses of nouns and deixis in discourse production in Alzheimer's disease. In *Journal of Neurolinguistics*, 19, 311-340.
- MEASSO, G., CAVARZERAN, F., ZAPPALÀ, G., LEBOWITZ, B.D., CROCK, T., PIROZZOLO, F.J., AMADUCCI, L.A., MASSARI, D. & GRIGOLETTO, F. (1993). The Mini-Mental State Examination: Normative study of an italian random sample. In *Developmental Neuropsychology*, 9 (2), 77-95.
- MONDINI, S., MAPELLI, D., VESTRI, A. & BISIACCHI, P.S. (2003). *Esame Neuropsicologico Breve (ENB). Una batteria di test per lo screening neuropsicologico*. Milano: Raffaello Cortina Editore.
- NASREDDINE, Z.S., PHILLIPS, N.A., B'EDIRIAN, V., CHARBONNEAU, S., WHITEHEAD, V., COLLIN, I., CUMMINGS, J.L. & CHERTKOW, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. In *Journal of the American Geriatrics Society*, 53 (4), 695-699.

- NOVELLI, G., PAPAGNO, C., CAPITANI, E., LAIACONA, M., VALLAR, G. & CAPPA, S.F. (1986). Tre test clinici di ricerca e produzione lessicale. Taratura su soggetti normali. In *Archivio di Psicologia, Neurologia e Psichiatria*, 4, 477-506.
- NUCCI, M., MAPELLI, D. & MONDINI, S. (2011). The Cognitive Reserve Questionnaire (CRIq): a new instrument for measuring the cognitive reserve. In *Aging clinical and experimental research*, 24, 218-226.
- PEINTNER, B., JARROLD, W.L., VERGYRIY, D., RICHEY, C., GORNO-TEMPINI, M.L. & OGAR, J. (2008). Learning diagnostic models using speech and language measures. In *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 4648-4651.
- PENG, H., LONG, F. & DING, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 (8), 1226-1238.
- PETERSEN, R.C., STEVENS, J., GANGULI, M., TANGALOS, E.G., CUMMINGS, J.L. & DEKOSKY, S.T. (2001). Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). report of the quality standards subcommittee of the american academy of neurology. In *Neurology*, 56 (9), 1133-1142.
- PIRANI, A., NASREDDINE, Z.S., TULIPANI, C., ZACCHERINI, D., MARTINI, E. & NERI, M. (2007). Montreal Cognitive Assessment (MoCA): uno strumento rapido per lo screening del Mild Cognitive Impairment. Dati preliminari della versione italiana. In *Atti IV Congresso Regionale Associazione Italiana Psicogeriatrica*.
- PIRANI, A., BRODATY, H., MARTINI, E., ZACCHERINI, D., NEVIANI, F. & NERI, M. (2010). The validation of the italian version of the GPCOG (GPCOG-It): a contribution to cross-national implementation of a screening test for dementia in general practice. In *International Psychogeriatrics*, 22 (1), 82-90.
- ROARK, B., HOSOM, J.P., MITCHELL, M. & KAYE, J.A. (2007). Automatically derived spoken language markers for detecting Mild Cognitive Impairment. In *Proceedings of the 2nd International Conference on Technology and Aging (ICTA)*.
- ROARK, B., MITCHELL, M. & HOLLINGSHEAD, K. (2007). Syntactic complexity measures for detecting Mild Cognitive Impairment. In COHEN, K.B., DEMNER-FUSHMAN, D., FRIEMAN, C., HIRSCHMAN, L. & PESTIAN, J. (Eds.), *Proceedings of the Workshop BioNLP 2007: Biological, translational, and clinical language processing*. ACL - Association for Computational Linguistics, 1-8.
- ROARK, B., MITCHELL, M., HOSOM, J.P., HOLLINGSHEAD, K. & KAYE, J.A. (2011). Spoken language derived measures for detecting Mild Cognitive Impairment. In *IEEE Transactions on Audio Speech, and Language Processing*, 19 (7), 2081-2090.
- SATT, A., SORIN, A. & TOLEDO-RONEN, O. (2012). Vocal biomarkers for dementia patient monitoring. In *Proceedings of Interspeech 2012*, ISCA.
- SATT, A., SORIN, A., TOLEDO-RONEN, O., BARKAN, O., KOMPATSIARIS, I., KOKONOZI, A. & TSOLAKI, M. (2013). Evaluation of speech-based protocol for detection of early-stage dementia. In *Proceedings of Interspeech 2013, 14th annual conference of the International Speech Communication Association*, ISCA, 1692-1696.

- SINGH, S., BUCKS, R.S. & CUERDEN, J.M. (2001). An evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. In *Aphasiology*, 15 (6), 571-583.
- SNOWDON, D.A., KEMPER, S.J., MORTIMER, J.A., GREINER, L.H., WEKSTEIN, D.R. & MARKESBERY, W.R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the nun study. In *The journal of the American Medical Association*, 275, 528-532.
- SZMRECS'ANYI, B.M. (2004). On operationalizing syntactic complexity. In PURNELLE, G., FAIRON, C. & DISTER, A. (Eds.), *Proceedings of the 7th International Conference on Textual Data Statistical Analysis*. Louvain-la-Neuve: Presses Universitaires de Louvain, 1031-1038.
- TALER, V., PHILLIPS, N.A. (2004). Language performance in Alzheimer's disease and Mild Cognitive Impairment: A comparative review. In *Journal of Clinical and Experimental Neuropsychology*, 30 (5), 501-556.
- THOMAS, C., KESELJ, V., CERCONE, N., ROCKWOOD, K. & ASP, E. (2005). Automatic detection and rating of dementia of Alzheimer type through lexical analysis of spontaneous speech. In GU, J., LIU, P.X. (Eds.), *Proceedings of the IEEE International Conference on Mechatronics & Automation*, vol. 3, 1569-1574.
- VIGORELLI, P. (Ed.) (2004). *La conversazione possibile con il malato Alzheimer*. Franco Angeli.
- VISSER, P.J., VERHEY, F.R., PONDS, R. & JOLLES, J. (2001). Diagnosis of preclinical Alzheimer's disease in a clinical setting. In *International Psychogeriatrics*, 13 (4), 411-423.
- WILLIAMS, J.A., WEAKLEY, A., COOK, D.J. & SCHMITTER-EDGECOMBE, M. (2013). Machine learning techniques for diagnostic differentiation of Mild Cognitive Impairment and Dementia. In MICHALOWSKI, M., MICHALOWSKI, W., O'SULLIVAN, D. & S. WILK (Eds.), *Expanding the Boundaries of Health Informatics Using Artificial Intelligence: Papers from the AAAI 2013 Workshop*, 71-76.