

CLUB

Working Papers in Linguistics

A cura di Elisabetta Magni e Yahis Martari

Volume 4, 2020



CLUB – CIRCOLO LINGUISTICO DELL'UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

CLUB

Working Papers in Linguistics

A cura di Elisabetta Magni e Yahis Martari

Volume 4, 2020

CLUB – CIRCOLO LINGUISTICO DELL'UNIVERSITÀ DI BOLOGNA

ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

Collana
CLUB WORKING PAPERS IN LINGUISTICS
(CLUB-WPL)

Comitato di Direzione

Cristiana De Santis
Nicola Grandi
Francesca Masini
Fabio Tamburini

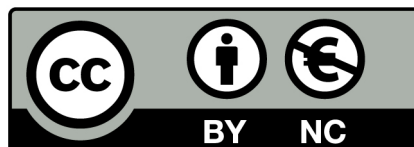
Comitato Scientifico

Marianna Bolognesi
Claudia Borghetti
Chiara Gianollo
Elisabetta Magni
Yahis Martari
Caterina Mauri
Marco Mazzoleni
Emanuele Miola
Rosa Pugliese
Mario Vayra
Matteo Viale

*Il CLUB – Circolo Linguistico dell'Università di Bologna nasce nel 2015
con l'obiettivo di riunire coloro che, all'interno dell'Alma Mater,
svolgono attività di ricerca in ambito linguistico.*

*Il CLUB organizza ogni anno un ciclo di seminari e pubblica una selezione
degli interventi nella collana CLUB WORKING PAPERS IN LINGUISTICS.*

*I volumi, sottoposti a una procedura di peer-review, sono pubblicati online sulla
piattaforma AMS Acta dell'Università di Bologna e sono liberamente accessibili.*



CC BY-NC

CLUB Working Papers in Linguistics, Volume 4, 2020
ISBN: 9788854970182

CLUB – CIRCOLO LINGUISTICO DELL'UNIVERSITÀ DI BOLOGNA
<http://corpora.ficlit.unibo.it/CLUB>

CLUB Working Papers in Linguistics

Volume 4

a cura di *Elisabetta Magni e Yahis Martari*

Indice

Presentazione <i>Elisabetta Magni e Yahis Martari</i>	5
The areal typology of grammaticalization: the case of northern China <i>Giorgio Francesco Arcodia</i>	7
Conventionality, deliberateness, and creativity in metaphors: toward a typology of figurative expressions in Latin semantics <i>Chiara Fedriani</i>	33
Reduplicazione e ripetizione in tipologia: due strategie separate o espressione dello stesso fenomeno? <i>Alessandra Barotto e Simone Mattiola</i>	47
Biomarker linguistici per la diagnosi di condizioni cliniche e precliniche <i>Gloria Gagliardi</i>	67
<i>Parlo come mangio!</i> Il lessico populista su Twitter <i>Claudia Roberta Combei</i>	103
Analyser linguistiquement l'écriture à l'école: EcriScol, un corpus génétique <i>Claire Doquet</i>	123
Uno studio acustico per l'interpretazione del mutamento delle labiovelari indoeuropee in occlusive bilabiali <i>Dennis Fucci</i>	137

Presentazione

Elisabetta Magni

Università di Bologna

elisabetta.magni@unibo.it

Yahis Martari

Università di Bologna

yahis.martari@unibo.it

Questo volume, il quarto della collana *CLUB Working Papers in Linguistics*, raccoglie una selezione dei contributi presentati durante l'anno accademico 2018/2019 nell'ambito delle attività del CLUB – Circolo Linguistico dell'Università di Bologna. Inoltre, come nelle precedenti edizioni, viene qui pubblicato anche un saggio tratto dalla tesi di laurea magistrale risultata vincitrice del premio CLUB Day 'Una tesi in linguistica' (<http://corpora.fielit.unibo.it/CLUB/>): quest'anno, il lavoro scelto è di Dennis Fucci.

I sette articoli, redatti in italiano, inglese o francese, sono (in ordine alfabetico) di Giorgio Francesco Arcodia, Alessandra Barotto, Claudia Roberta Combei, Claire Doquet, Chiara Fedriani, Dennis Fucci, Gloria Gagliardi e Simone Mattiola. Come di prassi, tutti i contributi sono stati sottoposti a un processo di revisione anonima da parte di due *reviewer* esperti, scelti all'interno o all'esterno del comitato scientifico della collana.

Le diverse discipline del linguaggio che sono rappresentate in questa miscellanea offrono al lettore uno sguardo d'insieme sulle ultime tendenze della ricerca in vari ambiti, tra cui la tipologia linguistica, la linguistica storica, la *corpus linguistics*, l'analisi del discorso e la linguistica clinica.

Il saggio di Francesco Arcodia analizza le proprietà tipologiche del cinese nell'ampio contesto genetico e areale delle lingue dell'est e del sud-est asiatico. Dal punto di vista diacronico, l'appartenenza al tipo isolante è tradizionalmente associata a fenomeni di grammaticalizzazione senza evoluzione formale. Tuttavia, alcune lingue settentrionali del diasistema in oggetto presentano possibili eccezioni a questa tendenza generale, dando luogo a processi di 'grammaticalizzazione secondaria' che l'autore discute nel dettaglio e in rapporto al quadro della ricerca tipologica.

Chiara Fedriani propone uno studio sulle metafore nella lingua latina. La studiosa discute in primo luogo i criteri utili a distinguere quelle ricorrenti e di uso convenzionale, quelle impiegate strategicamente per suscitare l'attenzione e la riflessione del lettore, e quelle prodotte creativamente dall'inventiva degli scrittori. Inoltre, i dati tratti da un ampio corpus di testi, permettono di ricostruire lo sviluppo delle metafore relative all'espressione dell'ira e della rabbia, nonché i fattori testuali e culturali che favoriscono l'emergere delle immagini associate alle interpretazioni metaforiche.

Alessandra Barotto e Simone Mattioli analizzano i fenomeni di reduplicazione e ripetizione, discutendo alcuni casi particolari che mettono in crisi i precedenti tentativi di demarcazione dei due processi sul piano formale e funzionale. Lo studio propone un nuovo modo di concettualizzare queste strategie seguendo, e ricalibrando a livello interlinguistico, l'approccio che considera reduplicazione e ripetizione come i poli prototipici di un *continuum* in cui forme e funzioni risultano distinte e inequivocabili solamente a livello di singole lingue.

Il lavoro di Gloria Gagliardi offre una panoramica sul contributo che la linguistica, e in particolare la linguistica computazionale, può fornire in ambito medico per lo *screening* e la diagnosi di condizioni cliniche e precliniche. A partire dalla definizione di “(bio)marker linguistico”, la ricerca presenta i risultati di alcuni studi multidisciplinari finalizzati allo sviluppo di strumenti per la diagnosi precoce delle condizioni di deterioramento cognitivo in età presenile e senile, e per l'individuazione di indizi linguistici utili nell'indagine sull'anoressia nervosa.

Claudia Roberta Combei conduce una ricerca sulle scelte lessicali che caratterizzano il discorso dei politici afferenti all'area cosiddetta ‘populista’. Nello specifico, l'indagine si basa su un corpus di oltre 6000 messaggi pubblicati su Twitter dagli esponenti del Movimento 5 Stelle e della Lega nel primo anno del loro governo. I risultati mettono in luce l'adozione di strategie differenti nella comunicazione dei due partiti, le cui politiche fanno perno su temi rispettivamente più e meno vari, ma anche su atteggiamenti diversi nei confronti del proprio elettorato.

L'articolo di Claire Doquet esplora le produzioni linguistiche non standard negli scritti di studenti sulla base del corpus EcriScol, sviluppato presso i laboratori della Sorbonne Nouvelle. Il database in questione offre un quadro diacronico sullo sviluppo delle capacità produttive nel periodo che va dall'inizio della scuola primaria fino al passaggio all'università. I materiali del progetto permettono di ricavare indicazioni non solo sulla qualità e la tipologia dei testi prodotti, ma anche sui differenti stadi evolutivi delle competenze di scrittura.

Infine, Dennis Fucci indaga un mutamento fonologico ampiamente noto nelle lingue indoeuropee antiche: il passaggio delle labiovelari a occlusive bilabiali. La presenza di processi analoghi anche in lingue di altre famiglie ha spinto a formulare diverse ipotesi sui fattori fonetici che assecondano il cambiamento. Nel solco delle spiegazioni di tipo acustico-percettivo, l'autore propone uno studio pilota sui suoni prodotti da un parlante igbo, ed elabora un algoritmo che individua automaticamente le transizioni all'interno di una formante, calcolando le variazioni di pendenza.

Elisabetta Magni e Yahis Martari

Biomarker linguistici per la diagnosi di condizioni cliniche e precliniche

Gloria Gagliardi

Università degli Studi di Napoli “L’Orientale”

ggagliardi@unior.it

Abstract

L’articolo si propone di fornire una panoramica del contributo che la linguistica, e in particolare la linguistica computazionale, può fornire in ambito clinico per lo *screening* e la diagnosi di condizioni cliniche e precliniche.

In apertura verrà fornita, sulla scia della letteratura scientifica biomedica, una definizione di “(bio)marker linguistico”, e verranno illustrati alcuni dei domini che nell’ultimo decennio hanno accolto con crescente favore l’avvento e l’applicazione delle tecniche di Natural Language Processing all’analisi delle condizioni patologiche; verranno poi presentati, a testimonianza dell’importanza strategica che le collaborazioni multidisciplinari possono rivestire anche in ambito umanistico, i risultati di alcuni studi clinici a cui ho collaborato nell’ultimo quinquennio: OPLON (*OPportunities for active and healthy LONGevity*), finalizzato allo sviluppo di strumenti per la diagnosi precoce delle condizioni di deterioramento cognitivo attraverso l’analisi computazionale delle produzioni verbali spontanee di soggetti in età presenile e senile, e PMLAN, studio prospettico osservazionale finalizzato all’individuazione di biomarker linguistici dell’Anoressia Nervosa. La possibilità di estendere questa metodologia di analisi ai disturbi del neurosviluppo verrà infine dimostrata presentando alcuni casi di studio sul Disturbo Primario del Linguaggio e sul Disturbo dello Spettro Autistico.

1. Introduzione¹

Nella letteratura medica e biomedica, così come nel linguaggio comune, si sono fatte strada negli ultimi anni le parole “marker” e “biomarker”, ad indicare parametri fisici quantificabili che, se opportunamente misurati, consentono ai clinici di stabilire una diagnosi, pianificare un trattamento e monitorarne l’efficacia (WHO 1993a; 2001; Biomarkers Definitions Working Group 2001; Strimbu & Tavel 2010). Vengono di solito annoverate in questa categoria le proprietà fisico-chimiche di liquidi biologici e tessuti, quantificate mediante i comuni esami ematochimici e citologici, oppure le alterazioni funzionali o morfologiche rilevate mediante gli esami strumentali (es. frequenza cardiaca e pressione sanguigna, *neuroimaging*).

¹ Il contributo riassume risultati già pubblicati negli anni passati e qui ricompresi in una riflessione più vasta sul ruolo del linguista (e della linguistica) negli studi clinici. Ai numerosi colleghi coinvolti va la mia gratitudine, non soltanto per il lavoro condotto insieme “sul campo”, ma anche per i preziosi momenti di confronto e i numerosi spunti di riflessione.

Crescente impiego ha registrato anche, nell'ambito delle scienze del linguaggio, la locuzione “(bio)marker linguistico”, con cui generalmente si fa riferimento a tratti linguistici evidenziabili nelle produzioni orali o scritte dei parlanti, che correlino con la presenza di un disturbo (neuro)psicologico oppure di una patologia organica.

Il linguaggio è la più evoluta e complessa tra le funzioni cognitive: non stupisce dunque che sia soggetto ad erosione, in modo più o meno vistoso, in una vasta gamma di condizioni mediche, in età evolutiva, adulta o involutiva. Talvolta questi deficit assumono la forma di veri e propri *breakdown* delle abilità verbali, a livello fonetico/fonologico, morfo-sintattico, semantico o pragmatico; più spesso si manifestano come insidiose microalterazioni, difficilmente rilevabili (e quantificabili) anche per i clinici più esperti. Se correttamente isolati, tuttavia, possono essere efficacemente utilizzati in funzione diagnostica o, meglio ancora, per lo screening nelle fasi precliniche e subcliniche delle patologie, quando altri sintomi organici sono poco o per nulla percettibili, perfino dallo stesso paziente.

In questo paradigma, dunque, indici quantitativi estraibili dalle verbalizzazioni dei parlanti, al pari di altri parametri (neuro)biologici, possono essere utilizzati come “marker”, in quanto “caratteristiche oggettivamente misurabili, che covariano e quindi possono essere considerate indicatori di un processo biologico normale, di un processo patogenico o le risposte farmacologiche ad un intervento terapeutico”.

La comunità di linguisti che si occupa di Natural Language Processing (NLP) sta apportando un contributo rilevante in questo ambito di studi: l'applicazione in campo medico di tecniche statistiche e informatiche ha consentito ad esempio, nell'ultimo decennio, l'individuazione di *pattern* latenti nel linguaggio dei pazienti colpiti da demenza, potenzialmente utili per l'identificazione precoce e la fenotipizzazione della patologia (Vincze et al. 2016; Asgari et al. 2017; König et al. 2018; Tóth et al. 2018; Themistocleous et al. 2018; Gosztolya et al. 2019; Fraser et al. 2019a; 2019b).

La metodologia è stata però applicata anche alla descrizione dell'eloquio e al *profiling* linguistico di numerose condizioni patologiche con esordio in età pediatrica oppure adulta. A titolo di esempio, una mole crescente di studi è stata dedicata al parlato disartrico, in particolare nella forma ipocinetica legata al Parkinson, di cui sono stati estesamente indagati i pattern ritmici ed acustici associati alle modificazioni fonarticolatorie conseguenti ai disordini del movimento tipici della patologia (Bayestehtashk et al. 2015; Pettorino et al. 2018; Benba et al. 2017; Arias-Vergara et al. 2018); in letteratura sono ampiamente attestati studi sui correlati acustici delle disfonie di natura organica oppure funzionale (Godino-Llorente & Gómez-Vilda 2004; Kazinczi et al. 2015) e dell'apnea ostruttiva del sonno (Montero-Benavides et al. 2014). Sono state altresì descritte alterazioni di natura morfo-sintattica e semantico-lessicale associate a vari disturbi mentali, quali schizofrenia, disordine bipolare e depressione (Ramirez-Esparza et al. 2008; Nguyen et al. 2017; Smirnova et al. 2018; Chandran et al. 2019; Stasak et al. 2019; Williamson et al. 2019). Per quanto riguarda invece le patologie dell'età evolutiva, hanno ricevuto crescente attenzione negli ultimi anni la sindrome dello spettro autistico (MacFarlane et al. 2017; Boorse et al. 2019; Ming et al. 2019), il Disturbo Primario del Linguaggio (Uwer et al. 2002; Ringeval et al. 2011; Clahsen et al. 2014; Dudy et al. 2018; Cacheda et al. 2019), la disprassia verbale evolutiva (Childhood Apraxia of Speech – CAS; Hosom, Shriberg, Green 2004) e le balbuzie (Świetlicka et al. 2013).

Si tratta dunque di un campo di ricerca in forte espansione, in cui la consulenza del linguista per l'individuazione, l'estrazione e l'analisi degli indici linguistici dovrebbe

essere imprescindibile; purtroppo, ancora troppo spesso, l'utilità di tali competenze è non soltanto sottovalutata, ma addirittura neanche percepita (Favilla & Ferroni 2009).

Questo articolo si propone dunque di fornire un esempio di come i linguisti possano dare il loro contributo nei gruppi di ricerca multidisciplinari in campo medico, illustrando la metodologia e i principali risultati di alcuni studi clinici a cui ho collaborato nell'ultimo quinquennio: OPLON (OPportunities for active and healthy LONGevity), finalizzato allo sviluppo di strumenti per la diagnosi precoce delle condizioni di deterioramento cognitivo attraverso l'analisi computazionale delle produzioni verbali spontanee di soggetti in età presenile e senile, e PMLAN, studio prospettico osservazionale che ha come scopo l'individuazione di biomarker linguistici dell'Anoressia Nervosa. La possibilità di estendere con profitto questa metodologia di analisi ai disturbi del neurosviluppo verrà inoltre dimostrata attraverso alcuni casi di studio sul Disturbo Primario del Linguaggio e sul Disturbo dello Spettro Autistico, svolti in collaborazione con i servizi di riabilitazione logopedica dell'Azienda USL Toscana Centro. Per ciascuna delle condizioni indagate verrà fornito un breve quadro clinico ed epidemiologico, verranno illustrate le specifiche motivazioni che determinano la necessità di sviluppare strumenti per la diagnosi precoce, e illustrate le strategie di ricerca adottate.

L'articolo è organizzato come segue: nel paragrafo 2 verrà presentato il progetto OPLON; nel § 3 verrà descritto il progetto PMLAN; il § 4 illustrerà invece i *case study* dedicati alla costruzione di profili linguistici qualitativi e quantitativi dei due disturbi con esordio in età evolutiva. Nel § 5 verranno infine tratte alcune brevi conclusioni.

2. Strumenti per l'identificazione degli stati di deterioramento cognitivo mediante l'analisi computazionale delle produzioni verbali: OPLON

2.1 Quadro epidemiologico

Uno dei fenomeni demografici che, dal punto di vista sanitario e socioeconomico, ha maggiormente interessato la società occidentale negli ultimi decenni è l'aumento significativo della popolazione anziana. I dati e le stime probabilistiche pubblicati dal Dipartimento per gli affari economici e sociali del Segretariato delle Nazioni Unite (DESA)² sono chiari: in Europa Occidentale la popolazione di età superiore a 65 anni sta crescendo più velocemente delle altre fasce demografiche (figure 1 e 2); entro il 2050, un europeo su quattro sarà ultrasessantenne.³

² <https://population.un.org/wpp/>

³ Ovviamente questi dati non tengono conto delle modificazioni che la pandemia di Covid-19, tutt'ora in corso, produrrà sulle caratteristiche demografiche della popolazione anziana. I dati di mortalità sono talmente variabili nei diversi paesi da impedire al momento qualsiasi stima.

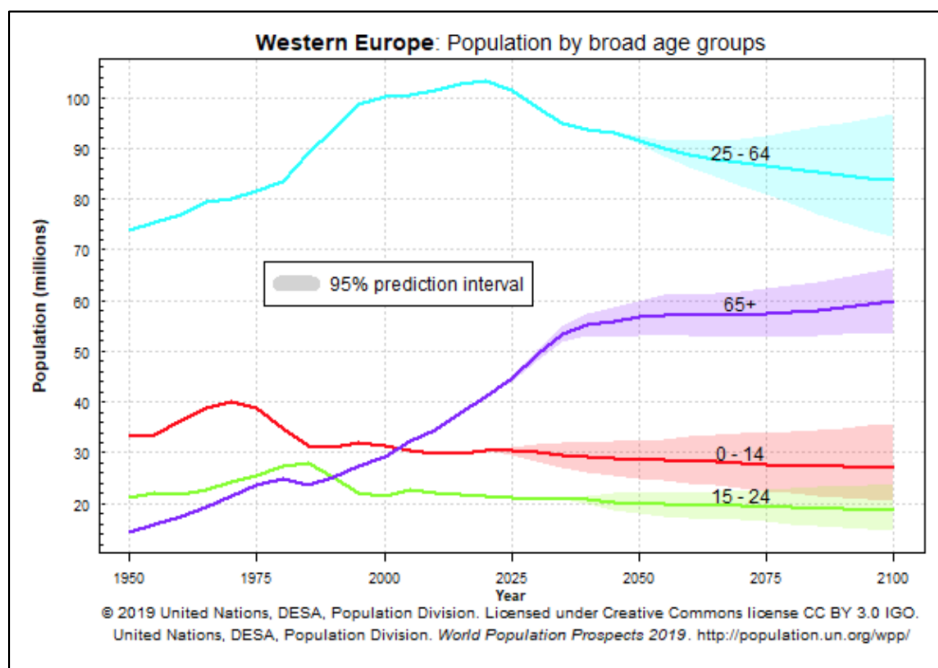


Figura 1. Andamento demografico per fasce di età (0-14; 15-24; 25-64; 65+), Europa Occ. (fonte: United Nations, DESA 2019).

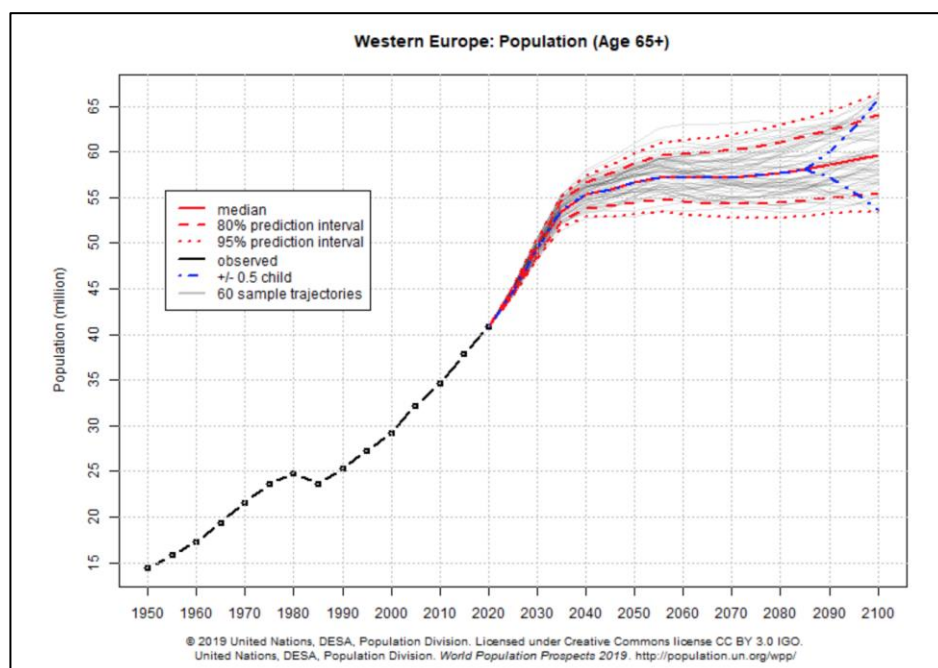


Figura 2. Proiezione demografica per la fascia di età 65+, Europa Occ. (fonte: United Nations, DESA 2019).

Il trend descritto è il risultato del sostanziale aumento dell'aspettativa di vita alla nascita che l'Europa Occidentale ha esperito a partire dal secondo dopoguerra (figura 3).

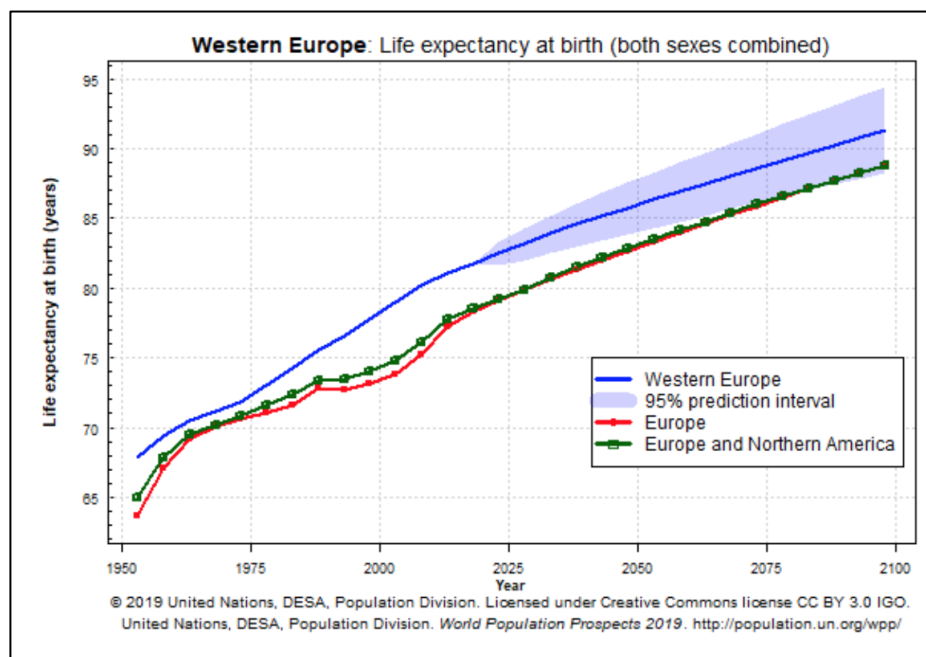


Figura 3. Aspettativa di vita alla nascita, Europa Occ. (fonte: United Nations, DESA).

Questo fenomeno, dovuto per lo più all'effetto del generale miglioramento delle condizioni igienico-sanitarie e della qualità di vita della popolazione, costituirebbe indubbiamente un'ottima notizia se, tra gli effetti legati alla maggior aspettativa di vita, non vi fosse l'aumento delle patologie neurodegenerative a carattere progressivo, il cui sintomo più rilevante è costituito dal decadimento cognitivo. Anche il numero di persone destinate a ricevere diagnosi di demenza è infatti in rapida crescita: attualmente, a livello globale, circa 50 milioni di pazienti vivono in tale condizione, ma la cifra è destinata a triplicare entro il 2050. Conducendo una meta-analisi dei dati epidemiologici disponibili, l'Alzheimer's Disease International (2018) ha stimato 9.9 milioni di nuovi casi ogni anno, ovvero in media una nuova diagnosi ogni 3 secondi. Nonostante a partire dagli anni '80 siano stati condotti numerosi studi di popolazione per stimare la prevalenza e l'incidenza delle demenze, dati epidemiologici certi sono ancora carenti, soprattutto nei contesti rurali e con basso livello di istruzione (tabella 1).

Prevalenza	Incidenza
5.05% (95% CI, 4.73-5.39)	11.08 per 1000 persone-anno (95% CI, 10.30-11.89)
M= 3.31% (95% CI, 2.85-3.80)	M= 7.02 per 1000 persone-anno (95% CI, 6.06-8.05)
F=7.13% (95% CI, 6.56-7.72)	F=13.25 per 1000 persone-anno (95% CI, 12.05-14.51)

Tabella 1. Epidemiologia della Demenza di Alzheimer, Europa Occ. (Niu et al. 2017).

Tale situazione rappresenta una reale emergenza per il Sistema Sanitario e Assistenziale pubblico: una popolazione più longeva richiede infatti una redistribuzione più efficiente delle cure e metodologie più accurate di prevenzione e diagnosi precoce (European Commission 2012).

2.2 Demenze e Mild Cognitive Impairment: caratteristiche cliniche

Con il termine “demenza”, assai diffuso nell’uso comune, si fa riferimento dal punto di vista medico ad una sindrome clinica caratterizzata dalla progressiva erosione delle funzioni cognitive, che in maniera ingravescente interferisce con le attività sociali e lavorative quotidiane: i pazienti che ne sono affetti manifestano sintomi cognitivi (es. disturbi di memoria, linguaggio, orientamento visuo-spaziale) in associazione a sintomi di natura comportamentale e psicologica (Cerejeira et al. 2012), come alterazioni dell’umore (depressione, labilità emotiva, euforia), ansia, sintomi psicotici (deliri, allucinazioni, misidentificazioni o falsi riconoscimenti) o neurovegetativi (alterazioni del ritmo sonno-veglia, dell’appetito, del comportamento sessuale), disturbi della personalità (indifferenza, apatia, disinibizione, irritabilità) e dell’attività psicomotoria (vagabondaggio, affaccendamento afinalistico, acatisia). Questa condizione, altamente invalidante, è determinata da una disfunzione acquisita delle funzioni cerebrali. In tal senso le demenze possono essere distinte in:

- *Primarie*, ovvero dovute a processi neurodegenerativi che comportano una perdita progressiva e irreversibile di cellule neuronali (es. Alzheimer, Demenza Fronto-Temporale, demenza a corpi di Levy).
- *Secondarie*, cioè legate a condizioni trattabili (e spesso potenzialmente reversibili) che inducono deterioramento cognitivo: ad esempio disturbi endocrini e/o metabolici (es. ipo- e ipertiroidismo, insufficienza renale cronica, ipoglicemia), malattie infettive e infiammatorie del Sistema Nervoso Centrale (es. meningiti, encefaliti, HIV, malattia di Creutzfeld-Jacob), stati carenziali (es. carenza di tiamina – Sindrome di Korsakoff, carenza di folati), sostanze tossiche (es. alcool, metalli pesanti, farmaci), neoplasie e traumi cranico-encefalici.

Tra le condizioni neurologiche del primo tipo, caratterizzate da decadimento cognitivo progressivo non reversibile legato ad una graduale perdita di funzione neuronale, il primo pensiero va di solito alla Malattia di Alzheimer (AD), di gran lunga la causa più nota e frequente di demenza, responsabile di circa il 60% dei casi diagnosticati; alta incidenza nella popolazione presenile e senile ha però anche la “demenza vascolare” (20% circa), legata a danni corticali o sottocorticali di natura cerebrovascolare (es. episodi micro-ictali); meno diffuse sono invece la demenza con corpi di Levy (LBD, 15%) e la demenza fronto-temporale (FTD, 5%), a sua volta ulteriormente classificata, sulla base delle prevalenti compromissioni all’*onset*, in Afasia Progressiva Primaria (PPA, con sintomatologia di natura linguistica) oppure variante comportamentale della FTD (bvFTD).

Ad oggi non esiste, purtroppo, una terapia farmacologica in grado di prevenire o di far regredire il processo degenerativo che porta alla morte neuronale. La possibilità di identificare tempestivamente la malattia si pone perciò come un imprescindibile obiettivo scientifico (Calzà et al. 2015): in primis, una diagnosi precoce permetterebbe di ottimizzare gli interventi assistenziali nei confronti di pazienti e *caregiver*, consentendo ad esempio una miglior definizione del quadro clinico e la predisposizione di programmi di prevenzione sempre più efficaci, e migliorando la *compliance* al percorso terapeutico, la risposta ai cambiamenti dello stile di vita e la riduzione dei rischi secondari della patologia; consentirebbe inoltre al paziente di pianificare la propria vita in una fase in cui

le sue capacità decisionali sono ancora perlopiù intatte; non ultimo, agevolerebbe la ricerca e la sperimentazione di terapie in grado di modificare il decorso clinico della patologia, fornendo al contempo ai pazienti la possibilità di accedere precocemente ai trattamenti farmacologici attualmente disponibili, che purtroppo si sono dimostrati efficaci soltanto se somministrati nelle fasi subcliniche e precliniche della malattia.

È noto che il processo neuropatologico che conduce alla demenza ha inizio molti anni prima dell'esordio dei sintomi: vi è però una difficoltà oggettiva nel determinare l'entità e la natura delle modificazioni delle funzioni cognitive correlate all'età o a condizioni metaboliche reversibili rispetto a quelle imputabili a un disturbo di tipo dementigeno, soprattutto in ragione dell'elevata variabilità interindividuale che caratterizza il declino delle attività intellettive nel corso della terza e della quarta età.⁴ Non esiste, infatti, una soglia sul piano biologico che delimiti il passaggio da una situazione di normalità cognitiva a una di demenza severa e conclamata; si tratta piuttosto di un continuum cognitivo, che va dall'invecchiamento "tipico" alla demenza passando per una fase di compromissione cognitiva lieve. Il Mild Cognitive Impairment (acronimo MCI; in italiano "Decadimento Cognitivo Lieve"), come entità clinica, individua appunto questa zona grigia (Petersen 2004; 2011; Winblad et al. 2004).

Nata per indicare la fase preclinica della malattia di Alzheimer, l'etichetta clinica MCI copre oggi uno spettro di condizioni molto eterogenee, che possono prevedere il coinvolgimento di uno o più domini cognitivi (MCI *single – multiple domain*), con o senza interessamento delle funzioni mnestiche (*amnesic – non amnesic* MCI), rappresentando la fase prodromica delle differenti demenze primarie poc'anzi descritte.

Qual è dunque l'identikit del paziente con MCI? I criteri diagnostici sono stati riassunti in tre punti (Winblad et al. 2004):

1. il paziente non è né normale, né demente: pur manifestando deficit cognitivi, non soddisfa i criteri per la diagnosi di demenza;
2. il paziente mostra un declino cognitivo. Tale declino:
 - a. è avvertito dal paziente stesso e/o confermato da un parente ed è dimostrato da un deficit in test psicometrico standardizzato;
 - b. è progressivamente ingravescente; il peggioramento del disturbo nel tempo è dimostrabile mediante somministrazioni successive di test cognitivi oggettivi;
3. le attività di base del vivere quotidiano sono preservate e il decadimento delle funzioni strumentali complesse è minimo.

Si tratta quindi di una condizione diagnosticata a individui che manifestano deficit cognitivi maggiori rispetto a quelli attesi per età e livello di istruzione, ma che non interferiscono significativamente con le attività giornaliere.

⁴ Si fa riferimento, in tal senso, al concetto di Riserva Cognitiva (CR), ovvero al costrutto teorico che descrive "l'insieme di abilità cognitive, capacità strategiche e conoscenze acquisite che, nell'arco della vita, le esperienze vissute hanno permesso di accumulare" (Mondini 2013). È ampiamente dimostrato in letteratura (Katzman et al. 1988; Hua et al. 2008; Stern, 2009; 2012) che "le differenze individuali nei processi cognitivi, e quindi nei sottostanti network neurali, conseguenti alle diverse esperienze a cui il soggetto è stato esposto (es. educazione, occupazione lavorativa e attività ricreative intellettualmente stimolanti), consentono ad alcuni pazienti di fronteggiare meglio modificazioni cerebrali strutturali e funzionali, costituendo dunque un fattore protettivo per il decadimento cognitivo" (Corsi & Gagliardi 2019: 163-164). Grazie alla CR il cervello può infatti, in maniera attiva, mettere in atto meccanismi compensatori che gli consentono di aggirare i limiti imposti dalla patologia.

Estremamente complesso è tracciare un quadro epidemiologico: la patologia, diagnosticata soprattutto nei centri specialistici, ha una prevalenza del 16.1% (Di Carlo et al. 2007; 2016). Bisogna però considerare che questo dato sicuramente sottostima il problema: la maggior parte dei pazienti, infatti, non riceve una diagnosi di demenza in questa fase così precoce di malattia (Mecocci & Ferracci 2017).

Mentre le modalità e gli strumenti per la diagnosi e la stadiazione della demenza sono ormai consolidati (es. Mini Mental State Examination - MMSE, Folstein et al. 1975; Velayudhan et al. 2014; Tsoi et al. 2015; Montreal Cognitive Assessment - MoCA, Nasreddine et al. 2005; Conti et al. 2015), in letteratura permangono ampi margini di incertezza sugli indicatori neuropsicologici più idonei per identificare i deficit cognitivi associati al MCI, ed in particolare su quale sia la testistica più idonea da proporre ai pazienti durante lo screening iniziale (Di Nuovo & Castellano, 2012).

Sarebbe cruciale poter disporre di test psicometrici sensibili e specifici per il disturbo: allo stato dell'arte, tuttavia, non esiste purtroppo un unico strumento sufficientemente accurato da poter essere utilizzato su larga scala per lo screening cognitivo dei potenziali pazienti, oppure linee guida *evidence-based* che possano supportare geriatri e neuropsicologi nella valutazione oggettiva delle prestazioni cognitive nelle fasi precliniche (Mortamais et al. 2017).

2.3 OPLON: *OPportunities for active and healthy LONGevity*

Numerosi studi scientifici hanno evidenziato che, nello spettro eterogeneo dei sintomi associati alle fasi iniziali e precliniche della demenza, spesso vi è un declino subliminale della facoltà di linguaggio (Taler & Philips 2008; Boschi et al. 2017). In tale fase, di solito, i deficit verbali non hanno ripercussioni evidenti sulla correttezza formale delle produzioni e non intaccano le capacità linguistico-comunicative di base come la comprensione in contesto ecologico, la partecipazione a conversazioni di routine e la lettoscrittura.

I disturbi linguistici si manifestano seguendo modalità e tempi diversi a seconda delle caratteristiche individuali del paziente (in particolare della sua riserva cognitiva - CR) e dello specifico fenotipo della malattia. Pur essendo presenti nei soggetti diversi anni prima dell'esordio clinico della patologia, sono però difficilmente identificabili mediante i tradizionali test neuropsicologici "carta e matita", poco ecologici se applicati all'elicitazione e alla valutazione di strutture testuali complesse. Il terreno più fertile per indagarne le caratteristiche (a livello qualitativo e quantitativo) o per seguirne le traiettorie di erosione è infatti l'osservazione dei parlanti in contesto spontaneo o semi-strutturato, *setting* che consente non soltanto di individuare eventuali difficoltà nella rievocazione di lessemi o veri e propri deficit morfo-sintattici, ma anche di intercettare impoverimenti a livello informativo e difficoltà di organizzazione e pianificazione testuale.

Come anticipato nel § 1, tali alterazioni, spesso impercettibili perfino per il paziente, in quanto manifestazioni prodromiche del processo neuropatologico in corso, possono essere utilizzate come veri e propri biomarker per lo screening e la diagnosi precoce della malattia. Le attività condotte presso il Dipartimento di Filologia Classica e Italianistica (FICLIT) ed il Centro di Ricerca Industriale "Scienze della vita e tecnologie per la salute" dell'Università di Bologna in collaborazione con l'unità di Neuropsicologia clinica dell'Arcispedale Santa Maria Nuova IRCSS di Reggio Emilia nell'ambito del progetto

OPLON - “Opportunities for active and healthy LONgevity” (finanziato dal MIUR nell’ambito della linea di intervento “Smart Cities and Communities”, DD 391/RIC) si sono inserite in questa linea: l’obiettivo era infatti realizzare un primo prototipo di *pipeline* in grado di estrarre automaticamente indici linguistici quantitativi dalle registrazioni delle verbalizzazioni semi-spontanee di soggetti di età presenile e senile, e identificare la presenza di un eventuale danno cognitivo grazie all’analisi computazionale di tali indicatori. A lungo termine, il progetto si propone di mettere a punto un dispositivo *low-cost* altamente “ecologico” da distribuire ai medici di base, che rintracci in maniera completamente automatica eventuali indici di impoverimento della produzione verbale correlati a decadimento cognitivo, consentendo di avviare un percorso diagnostico di approfondimento neurologico (Beltrami et al., 2016; 2018).

La tecnica sta acquisendo favore crescente nella comunità NLP internazionale (Vincze et al. 2016; Asgari et al. 2017; König et al 2018; Tóth et al. 2018; Themistocleous et al. 2018; Gosztolya et al. 2019; Fraser et al. 2019a; 2019b). Al momento, tuttavia, non risultano in letteratura studi simili in lingua italiana.

Nei paragrafi che seguono verrà illustrata la metodologia adottata e verranno presentati i risultati fin qui conseguiti. Lo studio, condotto in accordo con i principi etici enunciati nella dichiarazione di Helsinki, è stato approvato dal Comitato Etico dell’Azienda Ospedaliera di Reggio Emilia (protocollo n. 2013/0013438).

2.3.1 Costruzione del corpus ed estrazione degli indici linguistici

Sono stati reclutati 96 soggetti (48 M; 48 F), di età compresa tra i 50 ed i 75 anni, di cui 48 controlli (GC), con normale invecchiamento, e 48 affetti da declino cognitivo (PG). Tra questi ultimi, 32 hanno ricevuto diagnosi di Mild Cognitive Impairment; sono ulteriormente distinguibili in due sottogruppi:

- “a-MCI”, MCI single domain di tipo amnesico: 16 pazienti (8 M, 8 F);
- “md-MCI”, MCI multiple-domain: 16 pazienti (8 M, 8 F ; 12 amnesic-md-MCI; 4 non-amnesic-mdMCI)

16 sono invece risultati affetti da demenza in fase iniziale, “eD” (8 M, 8 F). L’indagine clinica ha individuato quattro tipologie di demenza nel campione: probabile demenza di Alzheimer (AD, 6 casi); Demenza Fronto-Temporale (FTD, 3 casi; di cui 1 Afasia Progressiva Primaria-PPA, 1 variante comportamentale-bvFTD, 1 Demenza Semantica), Demenza Mista (MD, 6 casi) e Demenza con Corpi di Levy (LBD, 1 caso). I criteri di inclusione e le caratteristiche demografiche del campione sono sintetizzati rispettivamente in tabella 2 e 3.

Gruppo di Controllo	Pazienti con Declino Cognitivo
- MMSE $\geq$ 24	- MMSE $\geq$ 18
- MoCA $\geq$ 18	- Diagnosi di Mild Cognitive Impairment (MCI) o Demenza in fase iniziale (eD)
- Assenza di patologie neurologiche	- Assenza di deficit neurosensoriali
- Assenza di deficit neurosensoriali	- Nessun riscontro di disabilità intellettiva in età evolutiva
- Nessun riscontro di disabilità intellettiva in età evolutiva	- Adeguata comprensione verbale (che consenta al soggetto di comprendere le istruzioni fornite dal neuropsicologo)
- Adeguata comprensione verbale	
- Assenza di deficit percettivi/attentivi	
- Assenza di patologie psichiatriche	

- Assenza di familiarità per patologie di natura dementigena con esordio precoce (< di 65 anni)	- Assenza di patologie psichiatriche
	<i>Ulteriori criteri per la diagnosi di MCI:</i>
	- Autonomia preservata
	<i>Ulteriori criteri per la diagnosi di a-MCI:</i>
	- Rilevamento di un deficit in un test standardizzato di memoria
	<i>Ulteriori criteri per la diagnosi di md-MCI:</i>
	- Due o più aree cognitive interessate dal deterioramento cognitivo
	<i>Ulteriori criteri per la diagnosi di eD:</i>
	- Due o più aree cognitive interessate dal deterioramento cognitivo
	- Parziale riduzione dell'autonomia

Tabella 2. Criteri di inclusione (Beltrami et al. 2018).

		Età		Scolarità
		n	(media $\pm$ ds)	(media $\pm$ ds)
Gruppo di controllo		48	61.60 $\pm$ 6.93	13.00 $\pm$ 3.92
Pazienti con Declino Cognitivo	a-MCI	16	64.19 $\pm$ 7.44	11.00 $\pm$ 4.00
	md-MCI	16	64.50 $\pm$ 7.47	11.56 $\pm$ 4.79
	eD	16	66.38 $\pm$ 6.70	9.38 $\pm$ 4.01

Tabella 3. Caratteristiche demografiche del campione (Beltrami et al. 2018).

Tutti i partecipanti coinvolti nello studio hanno completato un questionario di tipo anamnestico (es. informazioni anagrafiche, lavorative, rete familiare, familiarità per malattie neurodegenerative, storia clinica e farmacoterapia) e si sono sottoposti ad una valutazione neuropsicologica completa. La batteria di test somministrata indaga estesamente diversi domini cognitivi che vanno incontro precocemente a modificazioni in conseguenza di un processo neurodegenerativo, risultando dunque efficaci nella discriminazione tra soggetti normali, con demenza e Mild Cognitive Impairment. In particolare, sono stati somministrati:

- MMSE - Mini-Mental State Examination (Folstein et al. 1975; Measso et al. 1993)
- MoCA - Montreal Cognitive Assessment (Nasreddine et al. 2005; Conti et al. 2015)
- GPCog - General Practitioners assessment of Cognition (Brodaty et al. 2002)
- CDT - Clock Drawing Test (Freedman et al. 1994; Lee et al. 2011)
- Fluenze Verbal, fonemica (PF) e semantica (SF) (Novelli et al. 1986; Mioshi et al. 2006; Carlesimo et al. 1996)

In aggiunta ai tradizionali test carta e matita, sono stati proposti ai pazienti il Paired Associate Learning PAL, subtest del Cambridge Neuropsychological Test Automated

Battery – CANTAB (Junkkila et al. 2012), e un questionario di valutazione dell'indice di Riserva Cognitiva, il CRI-q, Cognitive Reserve Index Questionnaire (Nucci et al. 2011).

I risultati dell'esame neuropsicologico tradizionale (Beltrami et al. 2018), di cui si riportano mediane e distanze interquartiliche, sono sintetizzati in figura 4.

Le differenze sono state valutate dal punto di vista statistico mediante i test non parametrici di Mann-Whitney (GC vs. eD) e di Kruskal-Wallis con post-hoc di Dunn (GC vs. aMCI vs. mdMCI); i livelli di significatività considerati sono: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Come previsto, i punteggi medi del gruppo di controllo sono risultati significativamente più alti di quelli ottenuti dai pazienti con eD in tutti i test somministrati ($p < 0.0001$). Il test che meglio è riuscito a discriminare i soggetti a-MCI rispetto ai controlli è il GPCog: ciò è dovuto con alta probabilità al fatto che gran parte del test è basata su task con forte coinvolgimento mnesico.

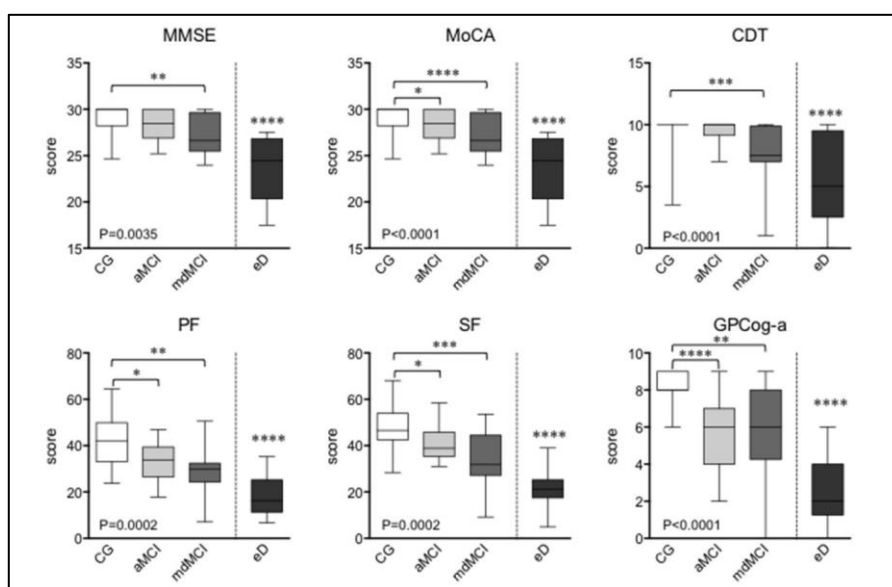


Figura 4. Risultati dei test neuropsicologici tradizionali somministrati al campione di studio (immagine tratta da Beltrami et al. 2018).

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

- Task FIGURA. Frase elicitante: “Descrivi questa immagine”. Lo stimolo visivo, tratto dal test 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 (figura 5).
- Task LAVORO. Frase elicitante: “Descrivi una giornata lavorativa”.
- Task SOGNO. Frase elicitante: “Descrivi l'ultimo sogno che ricordi”.

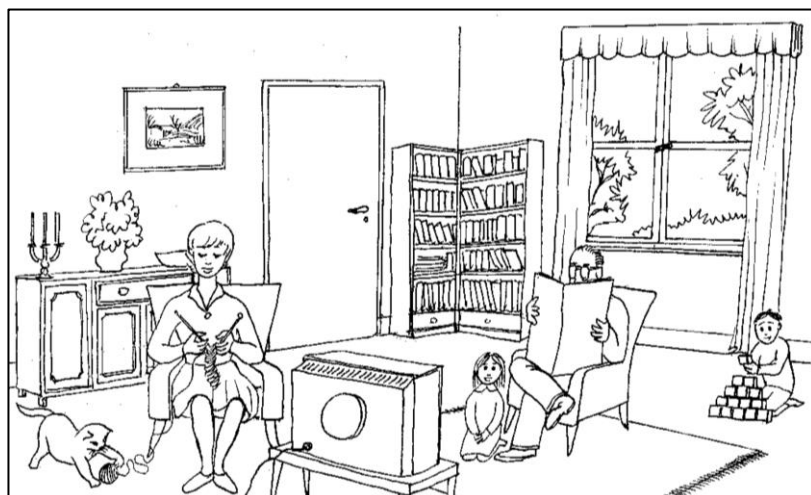


Figura 5. Stimolo visivo per la “descrizione di figura complessa” del test Esame del Linguaggio II (immagine tratta da Ciurli et al. 1996).

Oltre a stimolare la produzione verbale semi-spontanea dei soggetti, consentendo la successiva analisi linguistica, questi compiti permettono di valutare l’eventuale compromissione delle funzioni mnesiche; infatti, in tutte e tre le prove è indispensabile ricordare ciò che si sta dicendo (memoria di lavoro), ciò che si è già riferito (memoria episodica) o si vorrà raccontare (memoria prospettica); inoltre, lo svolgimento del compito richiede che si conosca il significato dei lessemi utilizzati (memoria semantica) e che si sia in grado di rievocare, su richiesta, ricordi personali riferiti ad un passato più o meno remoto (memoria episodica autobiografica, recente e remota).

I campioni di linguaggio parlato sono stati raccolti al termine delle sedute nella forma di file .WAV (segnale mono-canale digitalizzato a 44.1KHz di frequenza, range dinamico di 16 bit) mediante un registratore Olympus Linear PCM Recorder 233LS-5 posizionato sul tavolo di fronte al soggetto, e successivamente sottoposti a trascrizione ortografica manuale utilizzando il software Transcriber⁵ (figura 6). Come unità di allineamento è stato scelto l’enunciato, definito e identificato sulla base di criteri intonativi e illocutivi come il corrispettivo di un atto linguistico, ovvero “espressione linguistica interpretabile pragmaticamente” (Austin 1962; Cresti 2005; Cresti & Moneglia 2018).

I turni di parola del soggetto sottoposto a test sono estratti dalle trascrizioni, emendati manualmente dalle disfluenze e sottoposti ad annotazione morfosintattica semi-automatica; in particolare, gli enunciati selezionati sono stati sottoposti a *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 italiana (Tamburini 2007; Attardi & Simi 2009). Il *tagset* adottato è quello del TUT - Turin University TreeBank (Bosco et al. 2000); l’annotazione morfosintattica è prodotta nel formato tabulare standard CoNLL (Buchholz & Marsi 2006).

⁵ <http://trans.sourceforge.net>

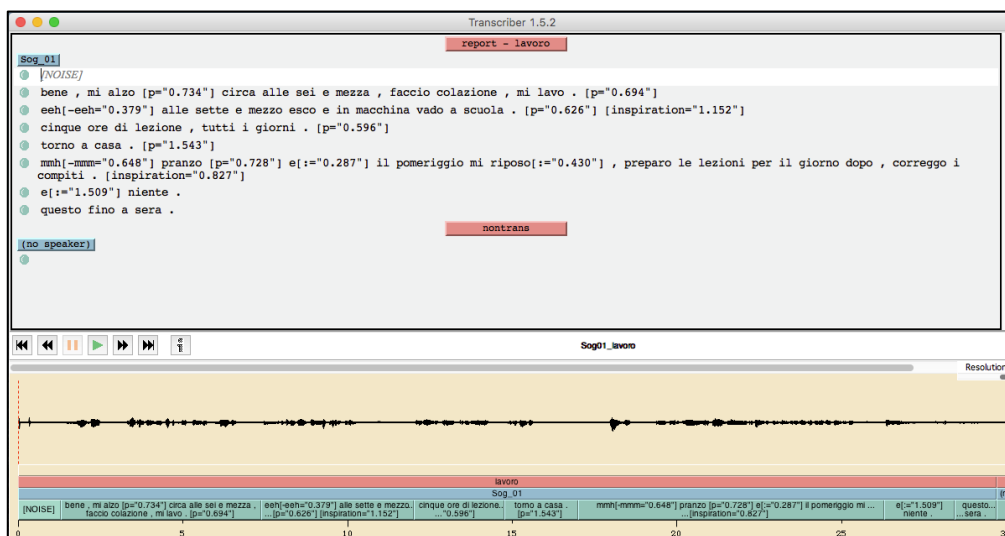


Figura 6. Screenshot del software Transcriber, trascrizione di un piccolo campione di parlato (dall'alto: editor di testo per l'annotazione e la segmentazione dei turni, pulsanti per la navigazione del file audio, waveform, tier di annotazione).

L'annotazione è stata ricontrollata manualmente utilizzando per la revisione il software *open source* DGA – *Dependency Grammar Annotator*, strumento grafico per l'annotazione sintattica di testi secondo il quadro teorico delle Grammatiche a dipendenze (figura 7).

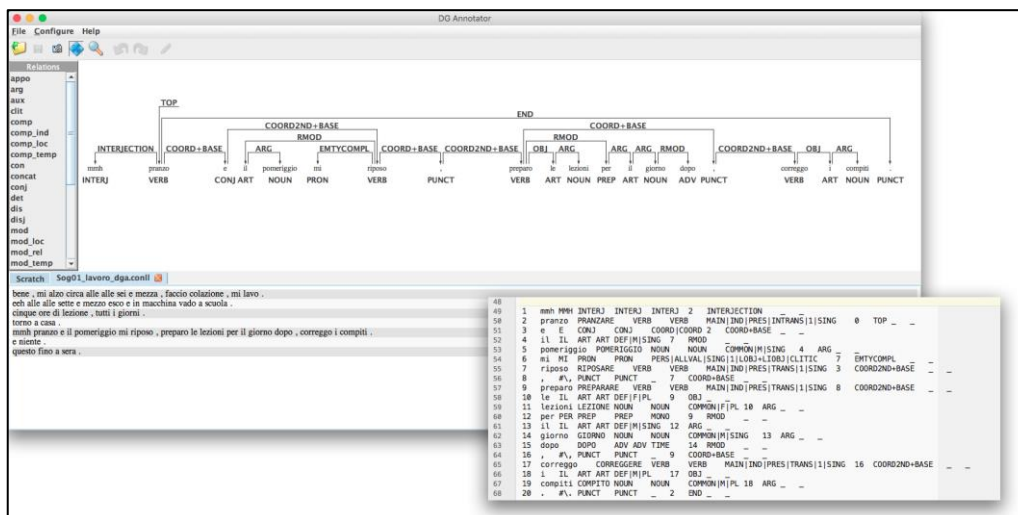


Figura 7. Screenshot del software DGA, correzione morfosintattica di un enunciato in ambiente grafico.

2.3.1 Verso un “profilo linguistico” del MCI e delle demenze: identificazione degli indici linguistici discriminativi

Sui campioni di parlato raccolti e annotati manualmente sono stati calcolati in maniera automatica indici quantitativi di tipo linguistico e stilometrico, già descritti e utilizzati in letteratura oppure creati ex novo. Le *feature* utilizzate, di cui si fornisce una sintesi in tabella 4, sono riferibili a quattro livelli di analisi: aspetti acustici e ritmici del segnale vocale, parametri di complessità lessicale e morfosintattica. Per una descrizione analitica si rinvia a Beltrami et al. (2018).

Indici acustici	
SPE_SILMEAN SPE_SILMEDIAN SPE_SILSD	Silence segments duration: media, mediana, deviazione standard (Satt et al. 2013)
SPE_SPEMEAN SPE_SPESD SPE_SPEMEDIAN	Speech segments duration: media, mediana, deviazione standard (Satt et al. 2013)
SPE_TRVSD	Temporal regularity of voiced segment (Satt et al. 2013)
SPE_VR	Verbal Rate (Singh et al. 2001; Roark et al. 2011)
SPE_TPR	Transformed Phonation Rate (Singh et al. 2001; Roark et al. 2011)
SPE_SPT	Standardized Phonation Time (Singh et al. 2001; Roark et al. 2011)
SPE_SPR	Standardized Pause Rate
SPE_RMSEM SPE_RMSED	Root Mean Square Energy: media, deviazione standard (López-de-Ipiña 2013)
SPE_PITCHM SPE_PITCHSD	Pitch: media, deviazione standard (López-de-Ipiña 2013)
SPE_SPCENTRM SPE_SPCENTRSD	Spectral Centroid: media, deviazione standard (López-de-Ipiña, 2013)
SPE_HFractDM SPE_HFractDSD	Higuchi Fractal Dimension: media, deviazione standard (López-de-Ipiña, 2013)
Indici ritmici	
RHY_%V	Percentage of vocalic intervals (Ramus et al. 2009)
RHYDeltaV RHYDeltaC	Standard deviation of vocalic and consonantal intervals (Ramus et al. 2009)
RHY_VnPVI RHY_CrPVI	Pairwise Variability Index, raw /normalized (Grabe & Low 2002)
RHY_VarcoV RHY_VarcoC	Variation coefficient for ΔV and ΔC (Delwo 2006)
Indici Lessicali	
LEX_ContDens	Content Density (Roark et al. 2011)
LEX_PoS *	Part-of-Speech rate (Holmes & Singh 1996; Bucks et al., 2000)
LEX_RefRRReal	Reference Rate to Reality (Vigorelli 2004)
LEX_PDEIXIS LEX_SDEIXIS LEX_TDEIXIS	Personal, Spatial and Temporal Deixis rate (March et al. 2006; Cantos-Gómez et al. 2009)
LEX_RPRO LEX_NEGADV	Relative pronouns and negative adverbs rate
LEX_TTR LEX_BrunetW LEX_HonoreR	Lexical Richness: Type-Token Ratio, W - Brunet's Index and R - Honoré's Statistic (Brunet 1978; Honoré 1979; Holmes & Singh 1996)
LEX_ACTVRB	Action Verbs rate (Gagliardi 2014)
LEX_DM_F	Frequency-of-use tagging (De Mauro 2000)
LEX_IDEAD	Propositional Idea Density (Snowdon et al. 1996; Roark et al. 2011)
LEX_NW	Number of words

Indici Sintattici	
SYN_NPLENM SYN_NPLENSD	Number of dependent elements linked to the noun: media e deviazione standard
SYN_GRAPHDISTM SYN_GRAPHDISTSD	Global Dependency Distance (Roark et al. 2011)
SYN_ISynCompl	Syntactic complexity (Szmrecsányi 2004)
SYN_MAXDEPTHM SYN_MAXDEPTHSD	Syntactic embeddedness: maximum depth of the structure
SYN_SLENM SYN_SLENSD	Utterance length

Tabella 4. Indici linguistici considerati nello studio (Beltrami et al. 2018).

Le misure acustiche e ritmiche sono state estratte direttamente dal segnale vocale: utilizzando il *Voice Activity Detector* “ssvad” (Mak & Yu 2014) il file audio è stato segmentato e sono state identificate le regioni contenenti voci umane rispetto a silenzi e rumori. Per ottenere una trascrizione fonetica automatica temporalmente allineata, necessaria al calcolo degli indici ritmici, è stato utilizzato il sistema di *forced alignment* del pacchetto Kaldi-DNN-ASR,⁶ addestrato sul corpus italiano APASCI (Angelini et al. 1994). I tratti lessicali e sintattici sono stati invece derivati dall’annotazione revisionata.

Per ragioni tecniche legate all’eccessivo rumore presente nelle registrazioni, è stato necessario scartare dal campione effettivamente analizzato 4 sessioni riferite a soggetti di controllo: il campione definitivo è perciò composto da 92 soggetti.

Anche in questo caso le differenze osservate nel campione sono state valutate dal punto di vista statistico mediante i test non parametrici di Mann-Whitney (GC vs. eD) e di Kruskal-Wallis (GC vs. aMCI vs. mdMCI); i livelli di significatività considerati sono: *p < 0.05; **p < 0.01; ***p < 0.001; ****p < 0.0001. I risultati sono sintetizzati in tabella 5.

	Task FIGURA		Task LAVORO		Task SOGNO	
	MCI	eD	MCI	eD	MCI	eD
SPE_SILMEAN	**	***	****	****		***
SPE_SILMEDIAN						
SPE_SILSD						
SPE_SPEMEAN	**	****	**	****	***	****
SPE_SPESD						
SPE_SPEMEDIAN						
SPE_TRVSD						
SPE_VR		****		***		****
SPE_TPR	***	***	***	****	***	****
SPE_SPT						
SPE_SPR	*	****	**	****	***	****
SPE_RMSEM						
SPE_RMSED						
SPE_PITCHM						
SPE_PITCHSD						
SPE_SPCENTRM	*	**				
SPE_SPCENTRSD						
SPE_HFractDM	**		**		**	
SPE_HFractDSD						
RHY_%V						

⁶ <http://kaldi.sourceforge.net/about.html>

RHYDeltaV						
RHYDeltaC					**	
RHY_VnPVI					*	
RHY_CrPVI					*	
RHY_VarcoV	**					
RHY_VarcoC					*	
LEX_ContDens	***	****				****
LEX_PoS_*						
ADJ	**		*			**
ADV				**		
ART				**		
CONJ					*	*
INTERJ		*		*		
NOUN				****		*
PHRAS						*
PREDET				*		
PREP		**				*
PRON		**		**		
VERB						
LEX_RefRReal			*	**		
LEX_PDEIXIS		*		**		
LEX_SDEIXIS						
LEX_TDEIXIS						
LEX_RPRO						
LEX_NEGADV						
LEX_TTR						
LEX_BrunetW	*			*		**
LEX_HonoreR		***				
LEX_ACTVRB		**		*		
LEX_DM F						
LEX_IDEAD	*					****
LEX_NW		*		**		
SYN_NPLENM		*		**		
SYN_NPLENSD						
SYN_GRAPHDISTM	*	**	*	***	**	***
SYN_GRAPHDISTSD						
SYN_ISynCompl				*		***
SYN_MAXDEPTHM		***	***	****	*	***
SYN_MAXDEPTHSD						
SYN_SLENM		****	****	****	**	****
SYN_SLENSD						

Tabella 5: Significatività statistica degli indici considerati nello studio (tabella adattata da Beltrami et al. 2018).

Sebbene il campione analizzato sia troppo ristretto per consentire di trarre indicazioni definitive, è possibile osservare come molti tra gli indici calcolati si siano dimostrati in grado di discriminare tra i gruppi oggetto dello studio; i domini maggiormente intaccati dalla patologia sembrano essere quello acustico, quello lessicale e quello sintattico. In particolare, risultano alterati alcuni parametri temporali della fonazione (es. *verbal rate*, *transformed phonation rate*, *standardized phonation time* e *standardized pause rate*) e alcune caratteristiche spettroacustiche della voce (es. *Spectral centroid*, *Higuchi Fractal Dimension*); le produzioni appaiono inoltre formalmente corrette ma impoverite, sia a livello lessicale (es. *Content Density*) che a livello sintattico (es. *Global Dependency*).

Distance). Ulteriori evidenze sperimentali sono tuttavia necessarie per confermare tali osservazioni.

Le operazioni descritte nei paragrafi precedenti sono propedeutiche alla costruzione e al *testing* di un sistema in grado di classificare automaticamente le produzioni verbali dei pazienti, attualmente in fase di sviluppo.

3. Biomarker linguistici per la diagnosi dell'Anoressia Nervosa: PMLAN

In prospettiva, la *pipeline* sviluppata è applicabile con minimi adattamenti anche all'analisi delle produzioni scritte e/o orali di parlanti affetti da altre patologie che comportino una compromissione, o comunque un'alterazione, della competenza comunicativa verbale, sia per la realizzazione di profili linguistici dettagliati che in funzione diagnostica. Il primo dominio a cui la metodologia presentata è stata estesa sono i Disturbi della Nutrizione e dell'Alimentazione, oggetto dei paragrafi che seguono.

3.1 Anoressia nervosa: sintomatologia, diagnosi ed epidemiologia della malattia

Sebbene abbiano suscitato crescente interesse nei media e nell'opinione pubblica, i Disturbi della Nutrizione e dell'Alimentazione, categoria diagnostica che raggruppa patologie caratterizzate “da un persistente disturbo dell'alimentazione o di comportamenti collegati con l'alimentazione che determinano un alterato consumo o assorbimento di cibo e che danneggiano significativamente la salute fisica o il funzionamento psicosociale” (APA 2013), rappresentano ancora un enigma per i ricercatori e una sfida per i clinici, sia sotto il profilo diagnostico, sia sul versante del trattamento terapeutico (AIDAP 2015).

L'eziopatogenesi, ad oggi perlopiù sconosciuta, è infatti con alta probabilità multifattoriale: origine e sviluppo della malattia sono cioè dovuti all'interazione complessa tra predisposizione biologica e fattori di rischio/precipitanti ambientali solo parzialmente noti (es. adesione a modelli culturali disfunzionali, ambiente familiare caratterizzato da dinamiche patogene, familiarità per problemi psicologici o psichiatrici, traumi).

Tra queste condizioni cliniche particolare interesse riveste, per caratteristiche demografiche, incidenza e prognosi, l'Anoressia Nervosa: il disturbo colpisce in larga maggioranza il genere femminile (in rapporto 10:1, APA 2013), ha prevalenza di 370/100.000 nelle giovani donne (Hoek 1993; Smink et al. 2012) e, sebbene in alcuni adolescenti abbia breve durata e guarisca senza alcuna cura o con un breve trattamento, nel 10-20% dei casi evolve in una condizione cronica che persiste per l'intera vita, danneggiando gravemente il funzionamento interpersonale del paziente. La patologia esibisce inoltre il più alto tasso di mortalità tra le malattie psichiatriche, pari a 5.1 ogni 1000 persone all'anno, 1.3 dei quali per suicidio (Harris & Barraclough 1998; Arcelus et al. 2011). Colpisce soprattutto la popolazione occidentale, mentre è rara nei Paesi in via di sviluppo, in cui non esiste una forte pressione sociale verso la magrezza.

Il disturbo, che spesso si presenta in comorbidità con altre patologie della sfera psichica, si caratterizza per una marcata perdita ponderale e una percezione alterata della propria immagine corporea. Secondo il DSM 5 (APA 2013) i criteri per la diagnosi sono:

- i. restrizione dell'introito energetico rispetto al fabbisogno tale da condurre a un peso corporeo significativamente basso in rapporto all'età, al sesso, alla traiettoria evolutiva e alla salute fisica. Si definisce significativamente basso un peso che sia inferiore a quello minimo normale o, nel caso dei bambini e degli adolescenti, inferiore al peso minimo atteso per l'età e il sesso;
- ii. intensa paura di aumentare di peso o di diventare grassi, o persistere in comportamenti che interferiscono con l'aumento di peso anche quando questo è significativamente basso;
- iii. alterazione del modo in cui vengono vissuti il peso o le forme del corpo, influenza indebita del peso o delle forme del corpo sulla valutazione di sé, o persistente mancanza di riconoscimento della gravità del sottopeso corporeo attuale.

L'indice di massa corporea (BMI - *Body Mass Index*), calcolato dividendo il peso del paziente (in kg) per il quadrato dell'altezza espressa in metri, è la misura più comune per valutare il criterio i. (WHO 1993b; 1995). In particolare, si considera patologico un BMI ≤ 18.5 (Cole et al. 2007). La gravità del disturbo è considerata lieve con BMI $\geq 17 \text{ kg/m}^2$, moderata con BMI $16-16,99 \text{ kg/m}^2$, grave con BMI $15-15,99 \text{ kg/m}^2$ ed estrema con BMI $< 15 \text{ kg/m}^2$. Il livello può però essere aumentato in base ai sintomi clinici (es. amenorrea, problemi ematologici, disfunzioni ormonali), al grado di disabilità funzionale e al bisogno di assistenza e sorveglianza. Due sono i sottotipi identificati dal DSM per la patologia:

- i. “tipo restrittivo” (ICD-10-CM: F50.01) se, negli ultimi 3 mesi, l'individuo non ha presentato ricorrenti episodi di abbuffate o condotte di eliminazione (per es. vomito autoindotto o uso inappropriato di lassativi, diuretici o enteroclistmi). La perdita di peso è quindi ottenuta principalmente attraverso la dieta, il digiuno e/o l'attività fisica eccessiva;
- ii. “con abbuffate/condotte di eliminazione” (ICD-10-CM: F50.02) se, negli ultimi 3 mesi, l'individuo ha presentato ricorrenti episodi di abbuffata o condotte di eliminazione.

I soggetti affetti da AN sono generalmente caratterizzati da rigidità di pensiero e vivono seguendo una disciplina ferrea, all'interno della quale prendere peso rappresenta un fallimento della propria capacità di autocontrollo. Tra i tratti di personalità caratteristici vi è il perfezionismo, che li porta ad esigere da sé stessi risultati maggiori rispetto a quelli richiesti dalla situazione, spesso irrealistici (es. “non sbagliare mai”). La volontà di raggiungere gli altissimi standard che si sono auto-imposti fa sì che si sottopongano a continui sforzi: quasi sempre questo atteggiamento di dedizione e sacrificio nasconde però una bassa autostima e una profonda insicurezza, che esprimono il timore di non essere accettati dagli altri per quello che si è. Nel complesso, tali aspetti predispongono i pazienti ad incorrere in un'attenzione selettiva per gli errori e nel “pensiero dicotomico” (“*black and white thinking*”), *bias* cognitivo che tende a schematizzare la realtà facendo distinzioni rigide e permanenti (es. bene/male, successo/fallimento, bello/brutto). In aggiunta, chi soffre di AN non ne è generalmente consapevole e non percepisce la gravità della sua situazione, anche in presenza di grave deperimento organico.

Sfortunatamente l'esordio dei sintomi è spesso insidioso, pertanto più della metà dei casi sfugge alla diagnosi nel contesto dell'assistenza sanitaria di base (Becker et al. 1999). Dal momento che più precocemente viene iniziato il trattamento, maggiore è la sua efficacia, i clinici sottolineano l'importanza di un'attenta individuazione delle situazioni

a rischio e di una diagnosi tempestiva, per intraprendere un intervento terapeutico adeguato fin dall'esordio.

3.2 Il progetto PMLAN

Affronta la questione il progetto PMLAN, condotto in collaborazione dal Centro Regionale per i Disturbi del Comportamento Alimentare (UO di Neuropsichiatria Infantile) del Policlinico Sant'Orsola – Malpighi e dai Dipartimenti di Scienze Mediche e Chirurgiche e di Filologia Classica e Italianistica (FICLIT) dell'Università di Bologna, sotto la supervisione delle prof.ssa Parmeggiani. Lo studio, approvato dal Comitato Etico dell'Azienda Ospedaliero-Universitaria di Bologna, Policlinico Sant'Orsola-Malpighi (protocollo n. 683/2019/Oss/AOUBo) e tuttora in corso, si propone infatti di individuare biomarker linguistici per lo screening e la diagnosi dell'AN, somministrabili nelle fasce di età critiche per l'insorgenza del disturbo in *setting* ecologici.

Sulla base delle evidenze sperimentali raccolte per altre patologie (§1) e della fenotipizzazione clinica dei soggetti con AN, si è ipotizzato che l'alterata percezione dell'immagine corporea, la rigidità di pensiero e la modificazione del tono dell'umore associate alla malattia trovino una manifestazione nelle produzioni verbali dei pazienti, e che tali alterazioni siano rilevabili mediante un'analisi computazionale del linguaggio, come già sperimentato nel progetto OPLON. Per verificare tale assunto, vengono misurati e confrontati dal punto di vista statistico indici linguistici quantitativi estratti dalle produzioni verbali di pazienti di età evolutiva affetti da AN, in cura presso i servizi ambulatoriali, di *day hospital* e di degenza ordinaria dell'UOS per i Disturbi del Comportamento Alimentare del Policlinico, e di coetanei normopeso reclutati presso istituti scolastici del territorio bolognese; i criteri di inclusione nel campione sono sintetizzati in tabella 6.

Gruppo di Controllo	Pazienti con AN
- Età: 14-18	- Età: 14-18
- BMI $\geq$ 18.5	- Diagnosi di Anoressia Nervosa, in accordo con le linee guida internazionali (es. DSM-5)
- Buona proprietà della lingua italiana (verificata mediante questionario sulla “storia linguistica” del soggetto)	- Buona proprietà della lingua italiana (verificata mediante questionario sulla “storia linguistica” del soggetto)
- Ottenimento del consenso informato	- Ottenimento del consenso informato

Tabella 6. Criteri di inclusione PMLAN.

Data l'età tipica di esordio, che si colloca in fascia scolare, si è deciso di raccogliere le produzioni scritte dei partecipanti invece della voce. Il protocollo richiede dunque di produrre tre brevi testi scritti (di circa 10-15 righe), in risposta ai seguenti stimoli:

1. task personale: “Parla di te: come sei, fisicamente e caratterialmente, cosa ami e non ami fare”;
2. task neutro: “Cosa fai solitamente quando incontri i tuoi amici?”;
3. descrizione di figura complessa: lo stimolo visivo proposto è la celebre figura della batteria BDAE - *Boston Diagnostic Aphasia Examination Battery*, nota come *Cookie theft* (Goodglass et al. 2001)



Figura 8: “Cookie theft” (immagine tratta da Goodglass et al. 2001).

I testi raccolti vengono digitalizzati e sottoposti ad annotazione morfosintattica semi-automatica mediante il parser a dipendenze TULE (§2.3.1); oltre all'estrazione dei tratti sintattici e lessicali già presentati in tabella 4, viene effettuata l'analisi delle produzioni con LIWC (*Linguistic Inquiry and Word Count*), software per l'analisi testuale che consente di calcolare l'incidenza percentuale di specifiche categorie lessicali al fine di “catturare” stati psicologici e variabili sociali dei parlanti (Chung & Pennebaker 2007; Tausczik & Pennebaker 2010). Infine, le differenze osservate tra i due gruppi vengono valutate dal punto di vista statistico mediante il test non parametrico di Kolmogorov-Smirnov.

Indici LIWC	
Processi linguistici	Numero totale di parole (WC), numero di parole per frase (WPS), parole composte da più di 6 grafemi (SIXLTR), parole del Dizionario LIWC (DIC)
Parole funzione	1 persona singolare (1PS), 1 persona plurale (1PP), 2 persona singolare (2PS), 2 persona plurale (2PP), 3 persona singolare (3PS), 3 persona plurale (3PP), negazioni (NEG), tempo passato (PST), tempo presente (PRES), tempo futuro (FUT), gerundio (GER), modo condizionale mood (COND), passivo (PASS), participio passato
Processi affettivi (AFFP)	Emozioni positive (+EMO), Emozioni negative (-EMO), Ansia (ANX), Rabbia (ANG), Tristezza (SAD)
Processi cognitivi (COGP)	Introspezione (INS), causa (CAU), discrepanze (DISCR), possibilità (TENT), certezza (CERT), inibizione (INH), inclusione (INCL), esclusione (EXCL)
Processi percettivi (PERCP)	Vista (SEE), udito (HEAR), percezione (FEEL)
Processi biologici (BIOP)	Corpo (Boby), salute/cura (Hlt), ingestione (Ing)
Preoccupazioni personali (PERSC)	Lavoro (WORK), morte (DEATH), realizzazione (ACH), piacere (LEIS), casa (HOME)
Processi psicologici (PSYP)	Famiglia (FAM), amici (FR), esseri umani (HUM), processi sociali (SOC)

Tabella 7. Indici LIWC.

Come anticipato, lo studio è attualmente ancora in corso: i risultati preliminari ottenuti su un campione bilanciato per sesso, età e scolarità di 51 ragazze (17 AN e 34 controlli) suggeriscono che alcuni dei parametri linguistici considerati sono effettivamente in grado di distinguere le pazienti dalle coetanee normopeso (Minori et al. 2020). In particolare, la semplificazione sintattica sembra il tratto più rilevante che caratterizza le produzioni verbali delle ragazze con AN: tra gli indici significativi vi sono infatti la lunghezza media delle frasi, il numero medio di elementi dipendenti dalla testa nei sintagmi nominali e la *Global Dependency Distance*. Dal punto di vista lessicale, invece, sono risultati distintivi *Content Density*, *R – Honoré’s statistic*, percentuale di avverbi e congiunzioni, deissi personale. A livello semantico, infine, i dati mostrano una minor incidenza di lessemi riferiti a processi percettivi nelle pazienti rispetto ai controlli. Le evidenze preliminari necessitano però di ulteriore conferma ampliando la coorte e seguendo longitudinalmente i soggetti reclutati.

4. Verso l’identificazione e l’estrazione computazionale di biomarker per i disturbi del neurosviluppo

Tra le condizioni cliniche che trarrebbero particolare beneficio dalla disponibilità di strumenti per la diagnosi precoce vi sono i disturbi del neurosviluppo, condizioni cliniche che comportano una compromissione del funzionamento cognitivo e/o sociale con esordio nell’infanzia. Il parlato patologico di età evolutiva pone però particolari difficoltà all’analisi computazionale rispetto alle produzioni verbali adulte: non soltanto non esistono, almeno per l’italiano, corpora linguistici bilanciati di grandi dimensioni che documentino in maniera affidabile le traiettorie di acquisizione della L1 in condizioni normotipiche, rispetto a cui valutare le devianze fonetico-fonologiche e morfosintattiche, ma i principali strumenti di annotazione automatica sono allenati su verbalizzazioni adulte e quindi mostrano performance molto al di sotto dello stato dell’arte se applicati agli enunciati prodotti dal bambino. Per superare tali ostacoli, sono state intraprese due linee di ricerca:

1. sviluppo di una *pipeline* per la trascrizione assistita, la correzione e la generazione di statistiche sui testi orali raccolti (Gagliardi, Gregori & Ravelli 2020);
2. raccolta, trascrizione e analisi quantitativa manuale delle produzioni verbali di bambini affetti da differenti disturbi del neurosviluppo, propedeutiche all’individuazione di marker linguistici da estrarre computazionalmente.

Vengono qui presentati, in maniera necessariamente sintetica, i risultati dei due primi casi di studio dedicati alla costruzione di profili linguistici in bambini con Disturbo Primario del Linguaggio (DPL) e Autismo, condotti in collaborazione con i servizi di Riabilitazione logopedica della AUSL Toscana Centro. Segnaliamo tuttavia che sono attualmente in corso ricerche affini su condizioni cliniche meno indagate, almeno in lingua italiana: il disturbo disprassico (orale e/o verbale), i disturbi fonetico-fonologici (“Speech Sound Disorder”) e il DPL in situazioni di multilinguismo.

4.1 Caratteristiche acustiche, lessicali e pragmatiche di bambini con Disturbo Primario del Linguaggio

Con l’etichetta clinica “Disturbo Primario del Linguaggio” - DPL (APA 2013; Leonard 2014; Bishop et al. 2017; FLI & CLASTA 2019) si fa riferimento ad un disordine dello sviluppo abbastanza frequente in età evolutiva (prevalenza: 5-7% in età prescolare), che comporta la compromissione o l’evoluzione deviante di una o più componenti del sistema linguistico non attribuibili a condizioni congenite o acquisite (es. deficit sensoriali o comorbidità con altri disturbi). Il DPL come categoria diagnostica non si riferisce a una condizione omogenea (Laws & Bishop 2003; FLI & CLASTA 2019): i deficit osservabili possono infatti limitarsi alla produzione oppure, nei casi più gravi, estendersi ai processi di comprensione del linguaggio verbale. Possono inoltre interessare aspetti diversi dell’elaborazione del messaggio (Wallace et al. 2015): la forma del linguaggio (elaborazione fonetico-fonologica e morfosintattica), il contenuto semantico-lessicale e/o il suo uso (elaborazione pragmatica e discorsiva). Le difficoltà linguistiche incontrate dal bambino hanno spesso sequele evidenti a partire dai primi anni della scuola primaria: il DPL si associa infatti a un elevato rischio di problemi di apprendimento scolastico (es. lettoscrittura e calcolo).

Dati l’alta prevalenza del disturbo nella popolazione e il suo forte l’impatto sullo sviluppo cognitivo e socio-emotivo del bambino, le linee guida internazionali suggeriscono l’utilità di identificare precocemente i soggetti a rischio promuovendo interventi riabilitativi nella fascia d’età prescolare, in cui l’efficacia è massima (CADTH 2013; FLI & CLASTA 2019).

Il caso di studio presentato (Beraldi, Gagliardi e Innocenti 2018; Gagliardi *in stampa*) si propone di tratteggiare un profilo completo delle abilità comunicative di un campione di 8 bambini in età prescolare con DPL espressivo già in trattamento presso i servizi territoriali, confrontandolo con le competenze verbali di coetanei con sviluppo cognitivo normotipo. A tal fine, durante lo svolgimento di tre task (figura 9), ne sono stati videoregistrati, trascritti e analizzati gli elinqui (tabella 8):

1. versione italiana del “Bus Story test” (Renfrew 2015; Mozzanica et al. 2016), prova standardizzata di *retelling* sollecitato da tavole cartacee figurate;
2. *retelling* della celebre fiaba “I tre porcellini”, proposta attraverso un libro cartaceo illustrato;
3. *retelling* del cortometraggio “Orso polare”, proposto mediante un tablet.

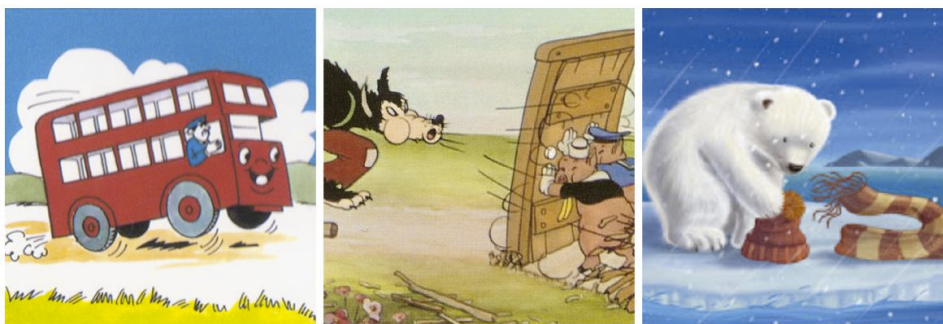


Figura 9. I tre task proposti nello studio.

Indici considerati nello studio	
Indici acustici/fluency	<i>Total locution time</i> (TLT), <i>Total phonation time</i> (TPT), numero, durata e percentuale di pause vuote (#EP, -EP, %EP), numero di pause piene (#FP), numero di false partenze e <i>retracting</i> (#R), <i>Verbal rate</i> (VR), <i>Standardized Phonation Time</i> (SPT), <i>Standardized Pause Rate</i> (SPR)
Indici lessicali e morfosintattici	Numero di parole (#W), enunciati (#E) e turni (#T), lunghezza media dell'enunciato (LME), numero di parole per turno (W/T) numero di enunciati per turno (E/T), numero di sequenze interrotte (#I, %V) enunciati <i>verbless</i> (#V, %V) e frasi (#S, %S), numero e percentuali di frasi principali, coordinate e subordinate (#MC, %MC, #CC, %CC, #SC, %SC), numero e correttezza dei pronomi clittici (#CLIT, %CLIT, #CLIT-O, #CLIT-D, #CLIT-S), uso del tempo passato (TENSE), numero di errori morfosintattici (#ERR), <i>type/token ratio</i> (TTR)
Indici pragmatici	Numero di espressioni referenziali (SNind., SNdef, PropN, DemPron., TonicPersPron, NullPron), numero di connettivi testuali

Tabella 8. Indici acustici, lessicali, morfo-sintattici e pragmatici considerati nello studio.

La significatività delle differenze riscontrate tra i due gruppi è stata valutata dal punto di vista statistico mediante test non parametrici, in ragione della ridotta numerosità campionaria (Kolmogorov-Smirnov per i dati di tipo quantitativo, χ^2 per i dati qualitativi; livello di significatività *p-value* < 0.05).

A livello acustico sono risultati significativi *Verbal rate* e *Standardized Phonation Time*, a dimostrazione che l'eloquio dei bambini con DPL è ipofluente rispetto ai coetanei normotipo. Rilevante anche la diversa composizione di turni e enunciati: il gruppo DPL produce, rispetto ai controlli, una percentuale più alta di enunciati interrotti a scapito di strutture frastiche ben formate; inoltre, mentre i bambini con sviluppo tipico ricorrono abbondantemente a strategie di coordinazione e subordinazione, organizzando le loro narrazioni in turni di parola strutturalmente articolati, il gruppo con DPL costruisce le proprie narrazioni orali giustapponendo per lo più frasi principali.

In prospettiva, i medesimi indici possono essere estratti in maniera automatica dalle produzioni parlate dei bambini, fornendo un supporto alla redazione del bilancio logopedico e agevolando il monitoraggio delle traiettorie di recupero nel corso del trattamento.

4.2 Caratteristiche prosodiche dell'eloquio dei bambini con Disturbo dello Spettro Autistico

Tra i disturbi del neurosviluppo che negli ultimi anni hanno ricevuto crescente attenzione vi è il Disturbo dello Spettro Autistico (*Autism Spectrum Disorder* - ASD), sindrome comportamentale causata da un disordine dello sviluppo biologicamente determinato che si esprime in modi e livelli di gravità variabili, causando disabilità permanente. Sebbene le caratteristiche individuali di ogni singolo soggetto autistico possano essere molto diverse (situandosi in uno "spettro", appunto), esistono alcune aree di fragilità che aiutano a definire e diagnosticare tale neurodiversità (APA 2013). In particolare, nei soggetti con ASD si riscontrano:

1. deficit persistenti della comunicazione sociale e nell'interazione sociale in molteplici contesti (es. reciprocità emotiva, comportamenti comunicativi non verbali, sviluppo/gestione/comprendimento delle relazioni);
2. pattern di comportamento, interessi o attività ristretti, ripetitivi (es. movimenti, uso di oggetti o eloquio stereotipati; insistenza nella *sameness* e aderenza alla routine priva di flessibilità; iper- o iporeattività in risposta a stimoli sensoriali o interessi insoliti verso aspetti sensoriali dell'ambiente).

Anche limitando l'attenzione allo sviluppo della facoltà linguistica, il quadro è ovviamente molto eterogeneo: i fenotipi comunicativi individuabili includono infatti sia casi in cui la comunicazione è limitata a semplici vocalizzi, sia elocui dotati di buona fluency, seppur talvolta connotati da una dominante dimensione ecolalica (Brandi 2005; Pfanner, Tancredi & Marcheschi 2008). La letteratura scientifica è però concorde nell'individuare tra le caratteristiche peculiari del disturbo, anche nei bambini che esibiscono un buon livello di funzionalità, deficit di natura prosodica. Numerosi lavori hanno documentato atipie riferite agli aspetti soprasegmentali dell'eloquio: l'intonazione e la qualità della voce vengono giudicate percettivamente anomale (es. la prosodia è di solito descritta come "aberrante", "inespressiva", "cantilenante", "robotica", "bizzarra"); anche il ritmo e l'intensità vengono spesso segnalati come "alterati" (McCann & Peppé 2003).

Per lo studio, attualmente in corso di pubblicazione (Biancalani 2019), è stato reclutato un campione di 28 bambini (22 M; 6F) di età compresa tra i 6 e i 10 anni: 14 bambini con diagnosi di Disturbo dello Spettro Autistico, in trattamento presso i servizi di logopedia dei presidi territoriali dell'Azienda USL Toscana Centro e l'Opera Santa Rita Fondazione Onlus di Prato, e 14 pari età con sviluppo cognitivo e linguistico tipico (gruppo di controllo). Sono state analizzate dal punto di vista acustico (tabella 9) le interazioni comunicative tra logopedista e bambino durante la generazione di una storia, presentata disponendo sei immagini colorate e plastificate in sequenza ordinata (figura 10).

Indici considerati nello studio	
Fluenza	durata del task, Total Phonation Time (TPT), pause piene (numero e durata), pause vuote (numero e durata)
Prosodia	velocità di articolazione, velocità di eloquio, range tonale, frequenza fondamentale (minimo, massimo, media)

Tabella 9. Indici considerati nello studio.



Figura 10. Stimolo visivo proposto nello studio (“Una festa di compleanno” tratta da “Sequenze da raccontare” della casa editrice Shubi).

Le sessioni sono state videoregistrate, trascritte e annotate manualmente; la significatività delle osservazioni è stata quindi valutata mediante test statistici (Kolmogorov-Smirnov per i dati di tipo quantitativo, χ^2 per i dati qualitativi; livello di significatività $p\text{-value} < 0.05$).

Anche in questo caso, numerose sono le differenze rilevate, sia sul piano linguistico, sia sotto il profilo funzionale. In particolare, sono risultati statisticamente significativi ($p\text{-value} < 0.05$) la durata complessiva dello scambio conversazionale, TPT, numero e durata delle pause (sia piena che vuote), velocità di eloquio e di articolazione, F_0 media e *range* tonale. Inoltre, tutti i bambini del gruppo ASD si sono limitati, su sollecitazione della logopedista, a descrivere la sequenza di immagini proposte, non cogliendone la struttura narrativa né passaggi impliciti della storia. Tali risultati, oltre ad avere un chiaro interesse conoscitivo e applicativo in vista della costruzione di una *pipeline* di analisi interamente automatica, suggeriscono la necessità di programmare interventi terapeutici a sostegno dello sviluppo delle competenze micro- e macrolinguistiche risultate deficitarie, possibilmente creando materiali che si adattino alle specifiche esigenze cognitive ed emotive dei bambini con ASD.

4. Conclusioni

A conclusione di questo contributo, ci sentiamo di riaffermare la necessità di una più stretta interazione tra scienze del linguaggio e discipline mediche: la linguistica può infatti rappresentare per i clinici una valida fonte da cui attingere metodi e quadri teorici di riferimento per l'interpretazione delle diverse “grammatiche patologiche” (Dovetto 2017). Al contempo, citando Jakobson (1941: 11), “per il linguista, che si interessa alla struttura del linguaggio pienamente dispiegata, anche la sua nascita e la sua estinzione devono contenere molti insegnamenti”.

La collaborazione sistematica tra linguisti e professionisti della salute che, a vario titolo, si occupano di valutazione e riabilitazione del linguaggio (in primis psicologi,

neuropsicologi e neuropsichiatri) deve necessariamente far fronte a ostacoli contingenti, legati non soltanto ai differenti percorsi di formazione, ma anche all'eccessiva specializzazione dei saperi; riteniamo, sulla scia di Favilla e Ferroni (2009), che la costruzione di un percorso condiviso passi necessariamente dalla condivisione di un linguaggio comune.

Bibliografia

- Associazione Italiana Disturbi dell'Alimentazione e del Peso – AIDAP. 2015. *I disturbi dell'alimentazione. Cosa sono? Come affrontarli?* Verona: AIDAP.
- Alzheimer's Disease International. 2018. *The World Alzheimer Report 2018*. London: Alzheimer's Disease International (ADI).
- American Psychiatric Association – APA. 2013. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: APA.
- Angelini, Bianca, Brugnara, Fabio, Falavigna, Daniele, Giuliani, Diego, Gretter, Roberto, & Omologo, Maurizio. 1994. Speaker independent continuous speech recognition using an acoustic-phonetic Italian corpus. In *Proceedings of ICSLP 94*, 18–22 settembre 1994, Yokohama (Japan), 1391–1394. Grenoble: ISCA.
- Arcelus, Jon, Mitchell, Alex J., Wales, Jackie & Nielsen, Søren. 2011. Mortality rates in patients with Anorexia Nervosa and other Eating Disorders: A Meta-analysis of 36 Studies. *Archives of General Psychiatry* 68(7). 724–731.
- Arias-Vergara, Tomas, Vásquez-Correa, Juan Camilo, Orozco-Arroyave, Juan R. & Nöth, Elmar. 2018. Speaker models for monitoring Parkinson's disease progression considering different communication channels and acoustic conditions. *Speech Communication* 101. 11–25.
- Asgari, Meysam, Kaye, Jeffrey & Dodge, Hiroko. 2017. Predicting Mild Cognitive Impairment from spontaneous spoken utterances. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* 3. 219–228.
- Attardi, Giuseppe & Simi, Maria. 2009. Overview of the Evalita 2009 Part-of-Speech tagging task. In *Poster and Work-shop Proceedings of the 11th Conference of the Italian Association for Artificial Intelligence*, 12 dicembre 2009, Reggio Emilia (Italia).
- Austin, John Langshaw. 1962. *How to Do Things with Words*. Oxford: Clarendon Press.
- Bayestehtashk, Alireza, Asgari, Meysam, Shafran, Izhak & McNames, James. 2015. Fully automated assessment of the severity of Parkinson's disease from speech. *Computer Speech & Language* 29(1). 172–185.
- Beltrami, Daniela, Calzà, Laura, Gagliardi, Gloria, Ghidoni, Enrico, Marcello, Norina, Rossini Favretti, Rema & Tamburini, Fabio. 2016. Automatic identification of Mild Cognitive Impairment through the analysis of Italian spontaneous speech productions. In Calzolari, Nicoletta, Choukri, Khalid, Declerck, Thierry, Goggi, Sara, Grobelnik, Marko, Maegaard, Bente, Mariani, Joseph Mazo, Hélène, Moreno, Asunción, Odijk, Jan, Piperidis, Stelios (a cura di), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, 2086–2093. Paris: ELRA - European Language Resources Association.
- Beltrami, Daniela, Gagliardi, Gloria, Rossini, Favretti, Rema, Ghidoni, Enrico, Tamburini, Fabio & Calzà, Laura. 2018. Speech Analysis by Natural Language

- Processing Techniques: A Possible Tool for Very Early Detection of Cognitive Decline? *Frontiers in Aging Neuroscience* 10. 369.
- Benba, Achraf, Jilbab, Abdelilah & Hammouch, Ahmed (2017). Detecting Multiple System Atrophy, Parkinson and Other Neurological Disorders Using Voice Analysis. *International Journal of Speech Technology* 20(2). 281–288.
- Beraldi, Francesca, Gagliardi, Gloria & Innocenti, Milvia. 2018. Abilità narrative nei bambini con DSL in età prescolare: un’analisi dei comportamenti linguistici e non-verbali nei compiti di retelling (Paper presentato alla 15 Conferenza annuale dell’Associazione Italiana di Scienze Cognitive “The new era of Artificial Intelligence: a cognitive perspective”, Pavia, 17-19 dicembre 2018.)
- Biancalani, Sara. 2019. *Aspetti soprasegmentali, pragmatici e non verbali del linguaggio nel Disturbo dello Spettro Autistico: uno studio pilota*. Firenze: Università degli Studi di Firenze. (Tesi di laurea.)
- Biomarkers Definitions Working Group. 2001. Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical pharmacology and therapeutics* 69(3). 89–95.
- Boorse, Jaclin, Cola, Meredith, Plate, Samantha, Yankowitz, Lisa, Pandey, Juhi, Schultz, Robert T. & Parish-Morris, Julia. 2019. Linguistic markers of autism in girls: evidence of a “blended phenotype” during storytelling. *Molecular Autism* 10. 14.
- Boschi, Veronica, Catricalà, Eleonora, Consonni, Monica, Chesi, Cristiano, Moro, Andrea & Cappa, Stefano F. 2017. Connected Speech in Neurodegenerative Language Disorders: A Review. *Frontiers in Psychology* 8. 269.
- Bosco, Cristina, Lombardo, Vincenzo, Vassallo, Daniela & Lesmo, Leonardo. 2000. Building a Treebank for Italian: a Data-driven Annotation Schema. In Gavrilidou, Maria, Carayannis, George, Markantonatou, Stella, Piperidis, Stelios & Stainhauer, Gregory (a cura di), *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC 2000)*, 31 maggio-2 giugno 2000, Athens (Greece). Paris: ELRA - European Language Resources Association.
- Brandi, Luciana. 2005. Linguaggio e comunicazione: dis/giunzioni autistiche. *Quaderni del Dipartimento di Linguistica - Università di Firenze* 15. 169–192.
- Brodsky, Henry, Pond, Dimity, Kemp, Nicola M., Luscombe, Georgina, Harding, Louise, Berman, Karen & Huppert, Felicia A. 2002. The GPCOG: A New Screening Test for Dementia Designed for General Practice. *Journal of the American Geriatrics Society* 50(3). 530–534.
- Brunet, Etienne. 1978. *Le Vocabulaire de Jean Giraudoux. Structure et Evolution*. Geneve: Slatkine.
- Bucks, Romola S., Singh, Sameer, Cuerden, Joanne M. & Wilcock, Gordon K. 2000. Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology* 14(1), 71–91.
- Buchholz, Sabine & Marsi, Erwin. 2006. CoNLL-X shared task on Multilingual Dependency Parsing. In *Proceedings of CoNLL-X '06: Proceedings of the Tenth Conference on Computational Natural Language Learning*, 149–164. Stroudsburg, PA: ACL - Association for Computational Linguistics.
- Cacheda, Fidel, Fernandez, Diego, Novoa, Francisco J. & Carneiro, Victor. 2019. Early Detection of Depression: Social Network Analysis and Random Forest Techniques. *Journal of Medical Internet Research* 21(6). e12554.

- CADTH. 2013. Screening tools compared to parental concern for identifying speech and language delays in preschool children: A review of the diagnostic accuracy. (https://www.cadth.ca/sites/default/files/pdf/htis/jul-2013/RC0464_RR_RiB_Speech_Delay_Tools_e.pdf) (consultato il 31.03.2020).
- Calzà, Laura, Beltrami, Daniela, Gagliardi, Gloria, Ghidoni, Enrico, Marcello, Norina, Rossini Favretti, Rema & Tamburini, Fabio. 2015. Should we screen for cognitive decline and dementia? *Maturitas* 82(1). 28–35.
- Cantos-Gòmez, Pascual. 2009. Featuring linguistic decline in Alzheimer's disease: A corpus-based approach. In Mahlberg, Michaela, González Díaz, Victorina & Smith, Catherine (a cura di), *Proceedings of the Corpus Linguistics Conference 2009 (CL2009)*, 20-23 luglio 2009, Liverpool (UK).
- Carlesimo, Giovanni Augusto, Caltagirone, Carlo, Gainotti, Guido, Nocentini, Ugo, Fadda, Lucia, Gallassi, Roberto, Lorusso, Sebastiano, Marfia, Girolama Alessandra, Marra, Camillo & Parnetti Lucilla. 1995. Batteria per la valutazione del deterioramento mentale (parte II): standardizzazione ed affidabilità diagnostica nell'identificazione di pazienti affetti da sindrome demenziale. *Archivio di Psicologia Neurologia e Psichiatria* 56. 471–488.
- Cerejeira, Joaquim, Lagarto, Luísa & Mukaetova-Ladinska, Elizabeta Blagoja. 2012. Behavioral and Psychological Symptoms of Dementia. *Frontiers in Neurology* 3. 73.
- Chandran, David, Robbins, Deborah Ahn, Chang, Chin-Kuo, Shetty, Hitesh, Sanyal, Jyoti, Downs, Johnny, Fok, Marcella, Ball, Michael, Jackson, Richard, Stewart, Robert Cohen, Hannah, Vermeulen, Jentien M., Schirmbeck, Frederike, de Haan, Lieuwe & Hayes, Richard. 2019. Use of Natural Language Processing to identify Obsessive Compulsive Symptoms in patients with schizophrenia, schizoaffective disorder or bipolar disorder. *Scientific Reports* 9. 14146.
- Chung, Cindy K, Pennebaker, James W. 2007. The psychological functions of function words. In Fiedler, Klaus (a cura di), *Social Communication*, 343–359. New York: Psychology Press.
- Ciurli, Paola, Marangolo, Paola, & Basso, Anna. 1996. *Esame del Linguaggio II*. Firenze: Giunti Organizzazioni Speciali.
- Clahsen, Harald, Rothweiler, Monika Sterner, Franziska & Chilla, Solveig. 2014. Linguistic markers of Specific Language Impairment in bilingual children: the case of verb morphology. *Clinical Linguistics & Phonetics* 28(9). 709–721.
- Cole, Tim J., Flegal, Katherine M., Nicholls, Dasha & Jackson, Alan A. 2007. Body Mass Index cut offs to define thinness in children and adolescents: international survey. *BMJ: British Medical Journal* 335(7612). 194.
- Conti, Silvia, Bonazzi, Stefano, Laiacona, Marcella, Masina, Marco & Vanelli Coralli, Mirco. 2015. Montreal Cognitive Assessment (MoCA) – Italian version: regression-based norms and equivalent scores. *Neurological Sciences* 36(2). 209–214.
- Corsi, Giulia & Gagliardi, Gloria. 2019. La riserva cognitiva in ambito afasiologico. Potenzialità e limiti. In: Gagliardi Francesco & Cruciani, Marco (a cura di), *Medicina, Filosofia e Cognizione*, 163–203. Roma: Aracne Editrice.
- Cresti, Emanuela. 2005. Enunciato e frase: teoria e verifiche empiriche. In Biffi, Marco, Calabrese, Omar & Salibra, Luciana (a cura di), *Italia Linguistica: discorsi di scritto e di parlato. Scritti in onore di Giovanni Nencioni*, 249–260. Siena: Prolagon.
- Cresti, Emanuela & Moneglia, Massimo. 2018. Chapter 13. The illocutionary basis of information structure: The Language into Act Theory (L-Act). In Adamou, Evangelia, Haude, Katharina & Vanhove, Martine (a cura di), *Information Structure in Lesser-*

- described Languages. Studies in prosody and syntax*, 360–402. Amsterdam – Philadelphia: John Benjamins.
- Delwo, Volker. 2006. Rhythm and speech rate: a variation coefficient for deltaC. In: Karnowski, Paweł & Szigeti, Imre (a cura di), *Language and Language-Processing. Proceedings of the 38th Linguistics Colloquium*, 27-29 agosto 2003, Piliscsaba (Hungary), 231–241. Frankfurt am Main: Peter Lang.
- De Mauro, Tullio. 2000. *Il dizionario della lingua italiana*. Torino: Paravia.
- Di Carlo, Antonio, Lamassa, Maria, Baldereschi, Marzia, Inzitari, Marco, Scafato, Emanuele, Farchi, Gino & Inzitari, Domenico. 2007. CIND and MCI in the Italian elderly. Frequency, vascular risk factors, progression to dementia. *Neurology* 68(22). 1909–1916.
- Di Carlo, Antonio, Baldereschi, Marzia, Lamassa, Maria, Bovis, Francesca, Inzitari, Marco, Solfrizzi, Vincenzo, Panza, Francesco, Galluzzo, Lucia, Scafato, Emanuele, Inzitari, Domenico & the Italian Longitudinal Study on Aging Working Group. 2016. Daily Function as Predictor of Dementia in Cognitive Impairment, No Dementia (CIND) and Mild Cognitive Impairment (MCI): An 8-Year Follow-Up in the ILSA Study. *Journal of Alzheimer's disease* 53(2). 505–515.
- Di Nuovo, Santo & Castellano, Sabrina. 2012. Dall'invecchiamento normale a quello patologico: indicatori di Mild Cognitive Impairment e strumenti per rilevarli. *Life Span and Disability / Ciclo Evolutivo e Disabilità* XV(2). 247–269.
- Dovetto, Francesca Maria (a cura di). 2017. *Lingua e patologia. Le frontiere interdisciplinari del linguaggio*. Roma: Aracne.
- Dudy, Shiran, Bedrick, Steven, Asgari, Meysam & Kain, Alexander. 2018. Automatic analysis of pronunciations for children with speech sound disorders. *Computer Speech & Language* 50. 62–84.
- European Commission. 2012. *Report on the public consultation on eHealth Action Plan 2012-2020*. (<https://ec.europa.eu/digital-single-market/en/news/report-public-consultation-ehealth-action-plan-2012-2020>) (consultato il 31.03.2020)
- Favilla, Maria Elena & Ferroni, Lucia. 2009. *Disturbi del linguaggio e neurolinguistica*. Perugia: Guerra Edizioni.
- FLI (Federazione Logopedisti Italiani) & Associazione CLASTA (Communication & Language Acquisition Studies in Typical and Atypical Populations). 2019. *Consensus Conference sul Disturbo Primario del Linguaggio*. (<https://www.disturboprimariolinguaggio.it/wp-content/uploads/2019/11/Documento-Finale-Consensus-Conference-2.pdf>) (consultato il 31.03.2020)
- Folstein, Marshal F., Folstein Susan E. & McHugh Paul R. 1975. Mini-Mental State. A practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research* 12(3). 189–198.
- Fraser, Kathleen C., Lundholm Fors, Kristina, Eckerström, Marie, Öhman, Fredrik & Kokkinakis, Dimitrios. 2019a. Predicting MCI Status From Multimodal Language Data Using Cascaded Classifiers. *Frontiers in Aging Neuroscience* 11. 205.
- Fraser, Kathleen C., Lundholm Fors, Kristina & Kokkinakis, Dimitrios. 2019b. Multilingual word embeddings for the assessment of narrative speech in Mild Cognitive Impairment. *Computer Speech & Language* 53. 121–139.
- Freedman, Morris, Leach, Larry, Kaplan, Edith, Winocur, Gordon, Shulman, Kenneth I. & Delis, Dean C. 1994. *Clock Drawing: A neuropsychological analysis*. Oxford: Oxford University Press.

- Gagliardi, Gloria. 2014. *Validazione dell'ontologia dell'azione IMAGACT per lo studio e la diagnosi del Mild Cognitive Impairment*. Firenze: Università degli Studi di Firenze (Tesi di dottorato).
- Gagliardi, Gloria. *in stampa*. Acquisizione di strutture sintattiche complesse e sviluppo della coesione nei bambini con disturbo del linguaggio in età prescolare: uno studio pilota. *AION Linguistica* 9. 136–160.
- Gagliardi, Gloria, Gregori, Lorenzo & Ravelli, Andrea Amelio. 2020. An NLP pipeline as assisted transcription tool for speech therapists. In *Proceedings of the 12 th International Conference on Language Resources and Evaluation (LREC 2020)*, 124–130. Paris: ELRA (European Language Resources Association).
- Godino-Llorente, Juan Ignacio & Gómez-Vilda, Pedro. 2004. Automatic Detection of Voice Impairments by Means of Short-Term Cepstral Parameters and Neural Network Based Detectors. *IEEE Transactions on Biomedical Engineering* 51(2). 380–384.
- Goodglass, Harold, Kaplan, Edith & Barresi, Barbara. 2001. *The Boston Diagnostic Aphasia Examination (BDAE)*. Philadelphia: Lippincott Williams & Wilkins.
- Gosztolya, Gábor, Vincze, Veronika, Tóth, László, Pákáski, Magdolna, Kálmán, János & Hoffmann, Ildikó. 2019. Identifying Mild Cognitive Impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. *Computer Speech & Language* 53. 181–197.
- Grabe, Esther, Low, Ee Ling. 2002. Durational variability in speech and the rhythm class hypothesis. In Gussenhoven, Carlos & Warner, Natasha (a cura di), *Papers in Laboratory Phonology* 7, 515–546. Berlin: Mouton de Gruyter.
- Harris, Clare & Barraclough, Brian. 1998. Excess mortality of mental disorder. *British Journal of Psychiatry* 173(1). 11–53.
- Hoek, Hans Wijbrand. 1993. Review of the epidemiological studies of Eating Disorders. *International Review of Psychiatry* 5(1). 61–74.
- Holmes, David I. & Singh, Sameer. 1996. A stylometric analysis of conversational speech of aphasic patients. *Literary and Linguistic Computing* 11(3). 133–140.
- Honoré, Antony. 1979. Some simple measures of richness of vocabulary. *Association of Literary and Linguistic Computing Bulletin* 7. 172–177.
- Hosom, John-Paul, Shriberg, Lawrence & Green, Jordan R. 2004. Diagnostic Assessment of Childhood Apraxia of Speech Using Automatic Speech Recognition (ASR) Methods. *Journal of medical speech-language pathology* 12(4). 167–171.
- Hua, Xue, Leow, Alex D., Parikshak, Neelroop, Lee, Suh, Chiang, Ming-Chang, Toga, Arthur W., Jack, Clifford R. Jr., Weiner, Michael W., Thompson, Paul M. & Alzheimer's Disease Neuroimaging Initiative. 2008. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's Disease: an MRI study of 676 AD, MCI and normal subjects. *NeuroImage* 43(3). 458–469.
- Jakobson, Roman. 1941. *Kindersprache, Aphasie und allgemeine Lautgesetz*. Uppsala: Almqvist & Wiksell (trad. it.: (2006) *Linguaggio infantile e afasia*, Torino, Einaudi).
- Junkkila, Jenny, Oja, Sakari, Laine, Matti & Karrasch, Mira. 2012. Applicability of the CANTAB PAL computerized memory test in identifying amnesic Mild Cognitive Impairment and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorder* 34. 83–89.
- Katzman, Robert, Terry, Robert, DeTeresa, Richard, Brown, Theodore, Davies, Peter, Fuld, Paula, Renbing, Xiong & Peck, Arthur. 1988. Clinical, pathological, and neurochemical changes in Dementia: A subgroup with preserved mental status and numerous neocortical plaques. *Annals of Neurology* 23. 138–144.

- Kazinczi, Ferenc, Mészáros, Krisztina & Vicsi, Klára. 2015. Automatic Detection of Voice Disorders. In: Dediu, Adrian-Horia, Martín-Vide, Carlos & Vicsi, Klára (a cura di), *Statistical Language and Speech Processing. SLSP 2015. Lecture Notes in Computer Science*, vol. 9449, 143–152. Cham: Springer.
- König, Alexandra, Linz, Nicklas, Tröger, Johannes, Wolters, Maria, Alexandersson, Jan & Robert, Phillipe. 2018. Fully Automatic Speech-Based Analysis of the Semantic Verbal Fluency Task. *Dementia and Geriatric Cognitive Disorders* 45(3-4). 198–209.
- Laws, Glynis & Bishop, Dorothy V.M. 2003. A comparison of language abilities in adolescents with Down Syndrome and children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research* 46. 1324–1339.
- Lee, Ji Hee, Oh, Eung Seok, Jeong, Seong Hae, Sohn, Eun Hee, Lee, Tae Yong & Lee, Ae Young. 2011. Longitudinal changes in clock drawing test (CDT) performance according to dementia subtypes and severity. *Archives of gerontology and geriatrics* 53(2). e179–182.
- Lesmo, Leonardo. 2007. Il parser basato su regole del Gruppo NLP dell'Università di Torino. *Intelligenza Artificiale* IV(2). 46–47.
- López-de-Ipiña, Karmele, Alonso, Jesus-Bernardino, Travieso, Carlos Manuel, Solé-Casals, Jordi, Egiraun, Harkaitz, Faundez-Zanuy, Marcos, Ezeiza, Aitzol, Barroso, Nora, Ecay-Torres, Miriam, Martinez-Lage, Pablo & Martinez de Lizardui, Unai. 2013. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors* 13. 6730–6745.
- Mak, Man-Wai & Yu, Hon-Bill. 2014. A study of voice activity detection techniques for NIST speaker recognition evaluations. *Computer, Speech & Language* 28(1). 295–313.
- MacFarlane, Heather, Gorman, Kyle, Ingham, Rosemary, Presmanes Hill, Alison, Papadakis, Katina, Kiss, Géza & van Santen, Jan. 2017. Quantitative analysis of disfluency in children with autism spectrum disorder or language impairment. *PLoS ONE* 12(3). e0173936.
- March, Evrim Gocer, Wales, Roger & Pattison, Pip. 2006. The uses of nouns and deixis in discourse production in Alzheimer's disease. *Journal of Neurolinguistics* 19(4). 311–340.
- McCann, Joanne & Peppé, Sue. 2003. Prosody in autism spectrum disorders: a critical review. *International Journal of Language & Communication Disorders* 38(4). 325–350.
- Measso, Giovanni, Cavarzeran, Fabiano, Zappalà, Giuseppe, Lebowitz, Barry D., Crock, Thomas H., Pirozzolo, Francis J., Amaducci, Luigi A., Massari, Danilo & Grigoletto, Francesco, 1993. The Mini-Mental State Examination: Normative Study of an Italian Random Sample. *Developmental Neuropsychology* 9. 77–95.
- Mecocci, Patrizia & Ferracci, Marika. 2017. Epidemiologia delle demenze e del MCI: World Alzheimer Report 2016. *Focus on brain* 1. 4–6.
- Mioshi, Eneida, Dawson, Kate, Mitchell, Joanna, Arnold, Robert & Hodges, John R. 2006. The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International journal of geriatric psychiatry* 21. 1078–1085.
- Ming, Li, Dengke, Tang, Junlin, Zeng, Tianyan, Zhou, Huilin, Zhu, Biyuan, Chen & Xiaobing, Zou. 2019. An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder. *Computer Speech & Language* 56. 80–94.

- Minori, Giulia, Gagliardi, Gloria, Cuteri, Vittoria, Tamburini, Fabio, Malaspina, Elisabetta, Gualandi, Paola, Rossi, Francesca, Moscano, Filomena, Francia, Valentina & Parmeggiani, Antonia. 2020. Linguistic Markers of Anorexia Nervosa: Preliminary Data from a Prospective Observational Study. In *Proceedings of the 12 th International Conference on Language Resources and Evaluation (LREC 2020)*, 34–45. Paris: ELRA (European Language Resources Association).
- Mondini, Sara, 2013. La riserva cognitiva: concetti e metodologie di indagine. In Bianchi, Angelo (a cura di), *L'esame neuropsicologico dell'adulto. Applicazioni cliniche e forensi*, 123–129. Firenze: Giunti O. S.
- Montero-Benavides, Ana, Fernández Pozo, Rubén, Torre Toledano, Doroteo, Blanco-Murillo, José Luis, López-Gonzalo, Eduardo & Hernández-Gómez, Luis A. 2014. Analysis of voice features related to obstructive sleep apnoea and their application in diagnosis support. *Computer Speech & Language* 28(2). 434–452.
- Mortamais, Marion, Ash, Jessica A., Harrison, John, Kaye, Jeffrey, Kramer, Joel, Randolph, Christopher, Pose, Carine, Albala, Bruce, Ropacki, Michael, Ritchie, Craig W. & Ritchie, Karen. 2017. Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia* 13. 468–492.
- Mozzanica, Francesco, Salvadorini, Renata, Sai, Elena, Pozzoli, Raffaella, Maruzzi, Patrizia, Scarponi, Letizia, Barillari, Maria Rosaria, Spada, Elena, Ambrogi, Federico & Schindler, Antonio. 2016. Reliability, validity and normative data of the Italian version of the Bus Story test. *International Journal of Pediatric Otorhinolaryngology* 89. 17–24.
- Nasreddine Ziad S., Phillips, Natalie A., Bédirian, Valérie, Charbonneau, Simon, Whitehead, Victor, Collin, Isabelle, Cummings, Jeffrey L. & Chertkow, Howard. 2005. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society* 53(4). 695–699.
- Nguyen, Thin, O'Dea, Bridianne, Larsen, Mark, Phung, Dinh, Venkatesh, Svetha & Christensen, Helen. 2017. Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimedia Tools and Applications*, 76. 10653–10676.
- Niu, Hao, Álvarez-Álvarez, Ismael, Guillén-Grima, Francisco & Aguinaga-Ontoso, Ines. 2017. Prevalence and incidence of Alzheimer's disease in Europe: A meta-analysis. *Neurología* 32(8). 523–532.
- Novelli, G., Papagno, Costanza, Capitani, Erminio, Laiacona, Marcella, Vallar, Giuseppe & Cappa, Stefano F. 1986. Tre test clinici di ricerca e produzione lessicale. Taratura su soggetti normali. *Archivio di Psicologia Neurologia e Psichiatria* 4. 477–506.
- Nucci, Massimo, Mapelli, Daniela & Mondini, Sara. 2011. The Cognitive Reserve Questionnaire (CRIq): a new instrument for measuring the cognitive reserve. *Aging clinical and experimental research* 24. 218–226
- Petersen, Ronald C. 2004. Mild Cognitive Impairment as a diagnostic entity. *Journal of Internal Medicine* 256(3). 183–194.
- Petersen, Ronald C. 2011. Clinical practice. Mild Cognitive Impairment. *New England Journal of Medicine* 364. 2227–2234.
- Pettorino, Massimo, Hemmerling, Daria, Vitale, Marilisa & De Meo, Anna (2018). Towards a Speech-Test for Parkinson's Disease detection: A Diachronic Study on Michael J. Fox. In *41st International Conference on Telecommunications and Signal Processing (TSP)*. Red Hook, NY: Institute of Electrical and Electronics Engineers (IEEE).

- Pfanner, Lucia, Tancredi, Raffaella & Marcheschi, Mara 2008. Comunicazione e linguaggio nei disturbi pervasivi dello sviluppo. *Giornale di Neuropsichiatria dell'Età Evolutiva* 28. 59–74.
- Ramirez-Esparza, Nairan, Chung, Cindy K., Kacewicz, Ewa & Pennebaker, James W. 2008. The Psychology of Word Use in Depression Forums in English and in Spanish: Testing Two Text Analytic Approaches. In Adar, Eytan, Hurst, Matthew, Finin, Tim, Glance, Natalie, Nicolov, Nicolas & Tseng, Belle (a cura di), *ICWSM 2008, Proceedings of the Second International Conference on Weblogs and Social Media*, 30 marzo-2 aprile 2008, Seattle (WA), 102–110. Menlo Park (CA): AAAI Press.
- Ramus, Franck, Nespor, Marina & Mehler, Jacques. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73. 265–292.
- Renfrew, Catherine. 2015. *Bus Story Test. A test of narrative speech*. London: Speechmark
- Ringeval, Fabien, Demouy, Julie, Szaszak, György, Chetouani, Mohamed, Robel, Laurence, Xavier, Jean, Cohen, David & Plaza, Monique. 2011. Automatic Intonation Recognition for the Prosodic Assessment of Language-Impaired Children. *IEEE Transactions on Audio, Speech, and Language Processing* 9(5). 1328–1342.
- Roark Brian, Mitchell, Margaret, Hosom, John-Paul, Hollingshead, Kristy & Kaye Jeffrey. 2011. Spoken language derived measures for detecting Mild Cognitive Impairment. *IEEE Transactions on Audio Speech, and Language Processing* 19(7). 2081–2090.
- Satt, Aharon, Sorin, Alexander, Toledo-Ronen, Orith, Barkan, Oren, Kompatsiaris, Ioannis, Kokonozi, Athina & Tsolaki, Magdalini. 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*, 1692–1696. ISCA.
- Singh, Sameer, Bucks, Romola S. & Cuerden, Joanne M. 2001. An evaluation of an objective technique for analysing temporal variables in DAT spontaneous speech. *Aphasiology* 15(6). 571–583.
- Smink, Frédérique R.E., van Hoeken, Daphne, Hoek, Hans W. 2012. Epidemiology of Eating Disorders: Incidence, prevalence and mortality rates. *Current Psychiatry Reports* 14(4). 406–414.
- Smirnova, Daria, Cumming, Paul, Sloeva, Elena, Kuvshinova, Natalia, Romanov, Dimitri & Nosachev, Gennadii. 2018. Language Patterns Discriminate Mild Depression From normal sadness and euthymic state. *Frontiers in Psychiatry* 9. 105.
- Snowdon, David A., Kemper, Susan J., Mortimer, James A., Greiner, Lydia H., Wekstein, David R. & Markesbery, William R. 1996. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the nun study. *The Journal of the American Medical Association (JAMA)* 275. 528–532.
- Stasak, Brian, Epps, Julien & Goecke, Roland. 2019. An investigation of linguistic stress and articulatory vowel characteristics for automatic depression classification. *Computer Speech & Language* 53. 140–155.
- Stern, Yaakov. 2009. Cognitive reserve. *Neuropsychologia* 47(10). 2015–2028.
- Stern, Yaakov. 2012. Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurology* 11. 1006–1012.
- Strimbu, Kyle & Tavel, Jorge A. 2010. What are Biomarkers? *Current opinion in HIV and AIDS* 5(6). 463–466.

- Świetlicka, Izabela, Kuniszyk-Józkowiak, Wiesława & Smółka, Elżbieta. 2013. Hierarchical ANN system for stuttering identification. *Computer Speech & Language* 27(1). 228–242.
- Szmrecsányi, Benedikt M. 2004. On operationalizing syntactic complexity. In Purnelle, Gérald, Fairon, Cédric & Dister, Anne (a cura di), *Proceedings of JADT 2004 - 7th International Conference on Textual Data Statistical Analysis*, 10-12 marzo 2004, Louvain-la-Neuve (Belgium), 1031–1038. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Taler, Vanessa & Phillips, Natalie A. 2008. Language performance in Alzheimer's disease and mild cognitive impairment: a comparative review. *Journal of Clinical and Experimental Neuropsychology* 30(5). 501–556.
- Tamburini, Fabio. 2007. Evalita 2007: The Part-of-Speech tagging task. *Intelligenza Artificiale* IV(2). 4–7.
- Tausczik, Yla R. & Pennebaker, James W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1). 24–54.
- Themistocleous, Charalambos, Eckerström, Marie & Kokkinakis, Dimitrios. 2018. Identification of Mild Cognitive Impairment From Speech in Swedish Using Deep Sequential Neural Networks. *Frontiers in Neurology* 9. 975.
- Tóth, László, Hoffmann, Ildikó, Gosztolya, Gábor, Vincze, Veronika, Szatlóczki, Gréta, Bánréti, Zoltán, Pákási, Magdolna & Kálmán, János. 2018. A Speech Recognition-based Solution for the Automatic Detection of Mild Cognitive Impairment from Spontaneous Speech. *Current Alzheimer Research* 15. 1–10.
- Tsoi, Kelvin K. F., Chan, Joyce Y.C., Hirai, Hoyee. W., Wong, Samuel Y.S., & Kwok, Timothy C.Y. 2015. Cognitive tests to detect dementia: a systematic review and meta-analysis. *JAMA internal medicine* 175. 1450–1458.
- United Nations, Department of Economic and Social Affairs, Population Division. 2019. *World Population Prospects 2019: highlights*. New York: United Nations (ST/ESA/SER.A/423).
- Uwer, Ruth, Albrecht, Ronald & Von Suchodoletz, Waldemar. 2002. Automatic processing of tones and speech stimuli in children with Specific Language Impairment. *Developmental Medicine and Child Neurology* 44(8). 527–532.
- Velayudhan, Latha, Ryu, Seung-Ho, Raczek, Malgorzata, Philpot, Michael, Lindesay, James, Critchfield, Matthew & Livingstone, Gill. 2014. Review of brief cognitive tests for patients with suspected dementia. *International psychogeriatrics* 26. 1247–1262.
- Vigorelli, Pietro (a cura di). 2004. *La conversazione possibile con il malato Alzheimer*. Milano: Franco Angeli.
- Vincze, Veronika, Gosztolya, Gábor Tóth, László, Hoffmann, Ildikó, Szatlóczki, Gréta, Bánréti, Zoltán, Pákási, Magdolna & Kálmán, János, 2016. Detecting Mild Cognitive Impairment by Exploiting Linguistic Information from Transcripts. In Erk, Katrin & Smith, Noah A. (a cura di), *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 7-12 agosto 2016, Berlino (Germany), 181–187. Stroudsburg, PA: ACL - Association for Computational Linguistics.
- Wallace, Ina F., Berkman, Nancy D., Watson, Linda R., Coyne-Beasley, Tamera, Wood, Charles T., Cullen, Katherine, Lohr, Kathleen N. 2015. Screening for speech and language delay in children 5 years old and younger: A systematic review. *Pediatrics* 136 (2). 1–15.

- WHO World Health Organization. 1993a. International Programme on Chemical Safety. Biomarkers and Risk Assessment: Concepts and Principles. (<http://www.inchem.org/documents/ehc/ehc/ehc155.htm>) (consultato il 31.03.2020)
- WHO World Health Organization. 1993b. *The ICD-10 classification of mental and behavioural disorders*. Geneva: WHO.
- WHO World Health Organization. 1995. *Physical status: the use and interpretation of anthropometry*. Technical report. Geneva: WHO.
- WHO World Health Organization. 2001. International Programme on Chemical Safety. Biomarkers in Risk Assessment: Validity and Validation. 2001. Geneva: WHO. (<http://www.inchem.org/documents/ehc/ehc/ehc222.htm>) (consultato il 31.03.2020)
- Williamson, James R., Young, Diana, Nierenberg, Andrew A., Niemi, James, Helfer, Brian S. & Quatieri, Thomas F. 2019. Tracking depression severity from audio and video based on speech articulatory coordination. *Computer Speech & Language* 55. 40–56.
- Winblad, Bengt, Palmer, Katie, Kivipelto, Miia, Jelic, Vesna, Fratiglioni, Laura, Wahlund, Lars-Olof, Nordberg, Agneta, Bäckman, Lars, Albert, Marilyn S., Almkvist, Ove, Arai, Hiroyuki, Basun, Hans, Blennow, Kaj, de Leon, Mony, DeCarli, Charles, Erkinjuntti, Timo, Giacobini, Ezio, Graff, Caroline, Hardy, John, Jack, Clifford R. Jr., Jorm, Anthony, Ritchie, Karen, van Duijn, Cornelia M., Visser, Pieter J. & Petersen, Ronald C. 2004. Mild Cognitive Impairment – beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine* 256(3). 240–246.