

studi AltLA 14

FARE LINGUISTICA APPLICATA
CON LE *DIGITAL HUMANITIES*

a cura di
JACOPO SATURNO - LORENZO SPREAFICO

Milano 2022

L'AItLA pubblica una collana di monografie e di collettanee sui diversi temi della linguistica applicata. I manoscritti vengono valutati con i consueti processi di revisione di pari per assicurarne la conformità ai migliori standard qualitativi del settore. I volumi sono pubblicati nel sito dell'associazione con accesso libero a tutti gli interessati.

Comitato scientifico

Cecilia Andorno, Giuliano Bernini, Simone Ciccolone, Silvia Dal Negro, Maria Elena Favilla, Anna De Meo, Laura Gavioli, Antonietta Marra, Natacha S.A. Niemants, Elena Nuzzo, Jacopo Saturno, Lorenzo Spreafico, Marilisa Vitale.

© 2022 AItLA - Associazione Italiana di Linguistica Applicata
Via Cartoleria, 5
40100 Bologna - Italy
email: info@aitla.it
sito: www.aitla.it



Edizione realizzata da
Officinaventuno
Via Flli Bazzaro, 18
20128 Milano - Italy
email: info@officinaventuno.com
sito: www.officinaventuno.com

ISSN: 2724-4830

ISBN edizione cartacea: 978-88-97657-50-7

ISBN edizione digitale: 978-88-97657-51-4

edizione digitale distribuita secondo la licenza Creative Commons CC BY NC ND



Indice

Introduzione	5
FRANCESCO CUTUGNO - EMANUELA CAMPISI Linguistica applicata alle esperienze di cultura digitale. Analisi multimodale nel progetto CHROME	9
CHRISTOPH DRAXLER Automatic Transcription of Spoken Language Using Publicly Available Web Services	27
VALERIA CARUSO - ROBERTA PRESTA Approcci digitali alla lessicografia: sviluppo e valutazione di un dizionario per smartphone	49
SIMONE CICCOLONE - GIULIA GROSSO Macchine parlanti o macchine apprendenti? Osservazioni preliminari sull'uso dell'etero-ripetizione in produzioni di chatbot e nell'interlingua	69
VALENTINA DE IACOVO - MARCO PALENA La variazione prosodica in italiano: l'utilizzo di un chatbot Telegram per la didattica assistita per apprendenti di italiano L2 e nella valutazione linguistica delle conoscenze disciplinari	85
GIOVINA ANGELA DEL ROSSO - SILVIA BRAMBILLA L'accuratezza della trascrizione ASR sul parlato non-standard. L'italiano nell'OH Portal	99
MARTA MAFFIA - MASSIMO PETTORINO - ANNA DE MEO - ROSA DE MICCO - GIOACCHINO TEDESCHI - ALESSANDRO TESSITORE Ritmo e malattia di Parkinson. Prime riflessioni sull'annotazione automatica del parlato di pazienti in stadi iniziali	117
LUISA REVELLI - ANTONIO MASTROPAOLO - DANIELE PAOLO RADICIONI La <i>sottodeterminazione</i> nei testi giuridici: verso un'analisi linguistico- computazionale	131
FLAVIA SCIOLETTE - EMILIANO GIOVANNETTI Un modello per domarli tutti: verso una rappresentazione del testo come esplicitazione di documento, lingua e contenuto	145
LORENZO SPREAFICO Fare linguistica applicata con le <i>digital humanities</i> : interfaccia di metodo o di comodo uso?	159
Indice autori	173

Ritmo e malattia di Parkinson.

Prime riflessioni sull'annotazione automatica del parlato di pazienti in stadi iniziali

Abstract

Speech disorders resulting from Parkinson's disease (PD), typically classified as hypokinetic dysarthria, include the alteration of speech rhythm, even at the very onset of the disease. The purpose of this study is to test an automatic annotation procedure for the calculation of two rhythmic variables – %V, vowel percentage, and VtoV, the mean interval between two consecutive vowel onset points – on a corpus of read speech collected from 20 patients with early-stage PD and from 20 healthy controls, sex-matched and similar age. The results of the automatic procedure, compared to those previously obtained with the manual labelling, were found to be effective in the characterization of PD speech rhythm, with significantly higher values of %V in patients' productions than in healthy control speech. Some critical issues are discussed in details.

1. Introduzione

Il Parkinson è la seconda malattia neurodegenerativa più diffusa al mondo dopo l'Alzheimer, in particolare nella popolazione con più di 65 anni (de Lau & Breteler 2006). Numerosi studi sperimentali condotti su diverse lingue hanno evidenziato come i cambiamenti anatomici e fisiologici correlati alla perdita di dopamina nel sistema nervoso centrale abbiano effetti sul sistema respiratorio, fonatorio e articolatorio dei soggetti affetti da tale morbo (tra gli altri, Darley et al. 1969a, Ramig et al. 2008). Nel parlato parkinsoniano la disartria ipocinetica si manifesta, infatti, in alterazioni a livello segmentale, come la riduzione dell'area vocalica e il fenomeno del *target undershooting*, il mancato raggiungimento del target articolatorio nella produzione sia dei suoni vocalici sia di quelli consonantici, collegato alla riduzione/alterazione dell'ampiezza e della velocità di movimento degli organi fonatori (Gili Fivela et al. 2014, Rusz et al. 2013, Sapir et al. 2010, Skodda et al. 2011). Sul piano soprasegmentale si assiste, inoltre, a una più scarsa modulazione del parlato, con un *range* tonale ridotto rispetto a quello osservato in parlanti di controllo (Darley

¹ Dipartimento di Studi Letterari, Linguistici e Comparati, Università degli Studi di Napoli "L'Orientale".

² Dipartimento di Scienze Mediche e Chirurgiche avanzate, Università degli Studi della Campania "Luigi Vanvitelli".

et al. 1969b, Goberman & Coelho 2005, MacPherson et al. 2011). Anche le caratteristiche ritmiche e la velocità di articolazione risultano alterate (Liss et al. 2009), sebbene non vi siano dati uniformi nella letteratura sul tema: in alcuni casi i soggetti parkinsoniani presentano il fenomeno della tachilalia, per cui parlano molto più velocemente rispetto a quelli sani (Ackerman et al. 1997, Hirose et al. 1982); in altri casi il parlato patologico è descritto come molto rallentato (Logemann et al. 1978, Ludlow et al. 1987); in altri ancora non si riscontra nessuna differenza tra parlato parkinsoniano e parlato sano (Duez 2006, Skodda & Schlegel 2008).

È attestato, comunque, un fenomeno di globale *dysrhythmia*, che non interessa solo il piano della comunicazione verbale ma che si evidenzia anche in compiti non linguistici come nel tenere il tempo con l'andatura o con il movimento di un dito (Cochen De Cock et al. 2018, Puyjarinet et al. 2018).

Negli stadi più avanzati della malattia, il parlato parkinsoniano è fortemente esitante, caratterizzato da una bassa velocità di eloquio, con frequenti e lunghe pause silenti e numerose disfluenze. Negli stadi iniziali, invece, i cambiamenti nei parametri acustici della voce, seppur strumentalmente rilevabili, sono percettivamente meno rilevanti e non intaccano l'intelligibilità della comunicazione. Tuttavia, diversi studi hanno ipotizzato che proprio l'osservazione di tali variazioni possa rivelarsi uno strumento sostenibile e non invasivo di supporto alla diagnosi clinica e al monitoraggio dell'avanzamento della patologia (tra gli altri, Cohen 2003, Harel et al. 2004).

1.1 Il progetto di ricerca

Il presente contributo è parte di un progetto di ricerca più ampio dal titolo "Analisi vocale per la diagnosi precoce del morbo di Parkinson: sviluppo e validazione di uno *speech tool* non invasivo", finanziato dai fondi PON-AIM 2014-2020. L'obiettivo a lungo termine del progetto è lo sviluppo di un dispositivo che, a partire da una traccia vocale (una registrazione di parlato letto), possa automaticamente calcolare alcuni parametri acustici.

In particolare, l'attenzione è posta su due variabili ritmiche, la durata media del VtoV (Vowel to Vowel), ossia l'intervallo tra il punto di inizio di un fono vocalico e il punto di inizio del fono vocalico successivo, e la percentuale di porzione vocale (%V) sul tempo totale dell'enunciato, escludendo le pause silenti e le eventuali disfluenze³. In studi precedenti, infatti, la metrica %V/VtoV è risultata efficace nel rendere conto delle variazioni ritmiche del parlato parkinsoniano rispetto al parlato sano in diverse lingue (Pettorino et al. 2016, Pettorino et al. 2017, Pettorino

³ Prendere in considerazione il VtoV permette di superare i limiti di una segmentazione basata su una "tradizionale" divisione in sillabe. I confini sillabici non sono infatti sempre chiaramente identificabili, come già discusso in Pettorino et al. 2013. Al contrario, i punti di inizio dei foni vocalici, i *VOPs* (*Vowel Onset Points*), sono facilmente individuabili sul segnale acustico in quanto caratterizzati da un repentino e notevole incremento del valore dell'intensità. Questo tipo di segmentazione, basata su caratteristiche acustiche, riflette, quindi, la reale percezione di discontinuità ritmiche nel segnale da parte di chi ascolta.

et al. 2018) e anche negli stadi iniziali della malattia (Maffia et al. 2020, Maffia et al. 2021).

Parallelamente alla raccolta di un sempre più ampio corpus di parlato patologico, il raggiungimento dell'obiettivo del progetto di ricerca prevede che la procedura di analisi acustica del parlato, fino ad ora manuale, diventi automatica. L'automaticità è legata da un lato alla necessità di ridimensionare il carico di lavoro dell'annotazione manuale, nell'ottica di ottimizzare i tempi e le energie della ricerca nel suo complesso, dall'altro alla prospettiva di effettuare in futuro tali misurazioni senza la presenza dell'operatore, nel contesto delle consuete valutazioni cliniche o anche da remoto.

1.2 Obiettivo del lavoro

Partendo dai risultati degli studi precedentemente condotti sulle caratteristiche ritmiche del parlato parkinsoniano, in questo contributo si intende valutare l'applicazione della procedura di annotazione automatica nel contesto del parlato disartrico di persone allo stadio iniziale della malattia di Parkinson. Si presenteranno, inoltre, alcune osservazioni preliminari di carattere qualitativo sui risultati di tale procedura.

2. Metodologia

2.1 I partecipanti e le valutazioni cliniche

Sono stati finora coinvolti nella ricerca 20 pazienti con malattia di Parkinson (PD), di cui 12 donne, dai 41 agli 81 anni (età media di circa 64 anni), tutti reclutati presso l'ambulatorio "Parkinson e altri disordini del movimento" della I Clinica neurologica e neurofisiopatologia dell'Università della Campania "Luigi Vanvitelli" (Napoli).

La diagnosi è stata posta sulla base dei criteri diagnostici modificati della *UK Parkinson's Disease Society Brain Bank* (Gibb & Lees 1988). I criteri di inclusione utilizzati sono i seguenti: (1) esordio della malattia dopo i 40 anni, per escludere parkinsonismi a esordio precoce; (2) un punteggio alla scala di Hoehn & Yahr modificato (mH&Y) ≤ 2.5 ; (3) durata della malattia ≤ 4 anni. I criteri di esclusione sono stati i seguenti: (1) deterioramento cognitivo clinicamente rilevante come da criteri diagnostici; (2) depressione maggiore, depressione minore, e disturbo distimico come da criteri del DSM-IV; (3) altre malattie neurologiche o patologie sistemiche clinicamente rilevanti o instabili.

Sono stati inoltre reclutati 20 soggetti sani con funzione di gruppo di controllo (GC), comparabili ai pazienti per caratteristiche biografiche (8 donne, dai 54 ai 77 anni, età media 65 anni), senza storia di malattie neurologiche e/o disturbi del linguaggio.

Tutti i partecipanti alla ricerca, residenti in Campania, hanno dichiarato l'italiano come lingua materna. Un questionario, inoltre, ha permesso di raccogliere ulte-

riori informazioni di natura sociolinguistica sull'uso delle eventuali diverse lingue conosciute.

I 20 pazienti affetti dalla malattia di Parkinson sono stati sottoposti a una estesa valutazione clinica motoria e non motoria per la quale sono state utilizzate scale validate e specifiche. La severità dei sintomi motori associati alla malattia è stata valutata con l'utilizzo della scala mH&Y e della *Unified Parkinson's Disease Rating Scale* parte III (UPDRS III). La scala UPDRS III è stata anche utilizzata per valutare la presenza di disturbo dell'articolazione del linguaggio clinicamente rilevante, utilizzando un *cut-off* del punteggio all'item 3.1 (Speech) ≥ 1 . La compromissione del disturbo della parola è valutata dal neurologo secondo una scala di severità che va da 0 (normale) a 4 (compromissione severa). Le funzioni cognitive globali sono state valutate con l'utilizzo della scala *Montreal Cognitive Assessment* (MoCA – Flostein et al. 1975). Inoltre, la presenza e severità di sintomi depressivi sono state valutate con l'utilizzo della scala *Beck Depression Inventory* (Beck et al. 1961). Infine, la dose equivalente di terapia dopaminergica giornaliera è stata calcolata utilizzando l'algoritmo di Tomlinson (2010).

In tabella 1 si riportano le caratteristiche demografiche e cliniche di tutti i partecipanti.

Tabella 1 - *Caratteristiche dei partecipanti*

		GC (n=20) media \pm dev.st.	PD (n=20) media \pm dev.st.	p-value*
Dati biografici	Età	64.8 \pm 5.9	63.8 \pm 10.9	0.713
	Sesso (M/F)	8/12	12/8	0.205
	Durata della malattia (in mesi)	-	31.9 \pm 17.1	-
Dati clinici	mH&Y	-	2.1 \pm 0.4	-
	UPDRS III	-	24.0 \pm 7.6	-
	Item 3.1 (Speech)	-	0.7 \pm 0.5	-
	MoCA total	-	22.4 \pm 4.0	-
	BDI	-	5.2 \pm 3.6	-
	LEDD total	-	202.5 \pm 157.8	-
	LEDD DA	-	30.0 \pm 76.6	-

H&Y: Hoehn & Yahr; UPDRS: Unified Parkinson's Disease Rating Scale; MoCA: Montreal Cognitive Assessment; BDI: Beck Depression Inventory; LEDD: Levodopa Equivalent Daily Dose; DA: Dopamine-Agonist. *Test del Chi quadrato.

2.2 La raccolta dei dati

La raccolta dei dati ha avuto luogo nel corso di una visita di *follow-up* clinico e ha previsto la registrazione di parlato letto da parte di ciascun partecipante. Per il *task* di lettura, è stato proposto a ciascun soggetto uno stesso testo di 175 parole e di circa 350 sillabe su un argomento comune (le abitudini alimentari del passato e quelle dei nostri giorni), con un alto livello di leggibilità (indice Gunning Fog: 6; indice Gulpease: 70), semplici strutture morfosintattiche e un lessico prevalentemente ad

alta frequenza (94% dei *tokens*). La scelta di un tale testo input è stata dettata dalla volontà di evitare incertezze nel parlato dovute alla non comprensione di quanto letto, in particolare nell'eventualità di casi di debole alfabetizzazione tra i partecipanti e/o scarsa dimestichezza con simili procedure di raccolta dati.

A tutti i soggetti coinvolti è stato chiesto di leggere ad alta voce il testo nel modo più naturale possibile, al volume e alla velocità a loro più congeniali. Le registrazioni sono state effettuate in una stanza dell'Università "Luigi Vanvitelli", utilizzando il software Praat (Boersma & Weenink 2021) e con frequenza di campionamento pari a 44100 Hz. In totale sono stati registrati circa 50 minuti di parlato letto.

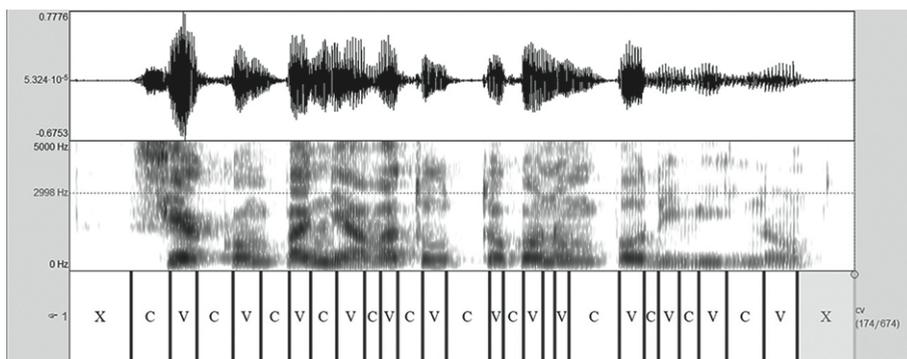
I pazienti hanno effettuato le registrazioni sotto l'effetto dei consueti farmaci dopaminergici. Tutti i soggetti hanno letto il foglio informativo sulle finalità della raccolta dati e firmato il relativo modulo per il consenso informato.

2.3 L'annotazione dei dati

Il corpus di registrazioni è stato sottoposto a due diverse procedure di annotazione.

In primo luogo esso è stato analizzato con Praat. Il segnale acustico è stato segmentato manualmente in intervalli vocalici (V) e consonantici (C), come mostrato in figura 1.⁴

Figura 1 - *Spettrogramma dell'enunciato "si accompagnava tutto con un po' di vino" (voce maschile). Annotazione manuale. C: intervallo consonantico, V: intervallo vocalico, X: pausa silente*



L'annotazione manuale dei più di 20.000 intervalli V e C è stata condotta da due esperti, i quali hanno svolto periodiche sessioni di standardizzazione dei risultati

⁴ La scelta di non riportare la trascrizione fonetica degli enunciati in tutte le didascalie alle immagini di spettrogrammi è legata all'utilizzo di una metodologia di analisi che intende superare le difficoltà relative all'individuazione dei singoli foni realmente pronunciati, siano essi consonantici o vocalici. Il tipo di segmentazione che si propone si basa, infatti, sulle caratteristiche del segnale acustico, in particolare dei parametri di intensità e periodicità: con l'etichetta V (o *Vowel*) si segnala la completa apertura del canale fonatorio, con libero passaggio del flusso di aria e assenza di turbolenze; al contrario, l'etichetta C (o *Consonant*) identifica la presenza di una chiusura, una stretta o un qualsiasi tipo di impedimento che, ostacolando il passaggio dell'aria nel canale fonatorio, provochi fruscio o frizione. La trascrizione fonetica sarà invece riportata in quei casi in cui si vorrà mostrare una realizzazione dell'enunciato molto lontana da quella attesa.

(in fase iniziale, intermedia e finale)⁵. In presenza di differenze o dubbi nelle due annotazioni, si è proceduto a un riesame caso per caso e si è stabilita una procedura comune.

In particolare, nell'annotazione manuale del segnale acustico, le criticità riscontrate sono state le seguenti:

- La segmentazione di alcune porzioni di parlato, soprattutto disartrico, caratterizzate da un forte fenomeno di ipoarticolazione. Un esempio è il tracciato riportato in figura 2. Il testo letto dal parlante è la frase "si faceva la spesa tutti i giorni". Come si vede la sequenza si riduce dalle 11 sillabe previste alle 7 sillabe realmente prodotte. In casi come questo la segmentazione ha tenuto conto delle caratteristiche acustiche del segnale, indipendentemente dal modello atteso, per cui l'intero enunciato è stato suddiviso in 7 intervalli consonantici e 7 vocalici. Va anche detto che in casi più problematici, in cui non è stato possibile identificare i singoli segmenti, si è deciso di etichettare come sequenze non comprensibili tali porzioni di enunciato, non considerandole poi nel calcolo della %V e del VtoV.
- L'individuazione del confine nel nesso V+C nasale. Laddove visibile, la porzione nasalizzata della vocale è stata assegnata all'intervallo V. Un esempio particolarmente interessante è quello riportato in figura 3. La sequenza si riferisce alla frase "secondo diversi studi". In corrispondenza del nesso [ondo] nella parola "secondo" è infatti possibile identificare cinque distinti tratti: un primo tratto vocalico, un tratto in cui la vocale si nasalizza, come risulta dalla riduzione dell'ampiezza del segnale, un tratto consonantico nasale, con ulteriore caduta dell'intensità, un tratto di occlusiva sonora, anch'essa parzialmente nasalizzata, e finalmente la vocale. Nella segmentazione i tratti risultano solo tre, V-C-V. L'inclusione della porzione corrispondente alla vocale nasalizzata nel primo intervallo vocalico (e non nella consonante nasale) si basa su considerazioni che tengono conto della meccanica articolatoria richiesta nella sua produzione. La vocale nasalizzata, infatti, si distingue dalla vocale orale esclusivamente per l'apertura del diaframma rinovelare durante la sua realizzazione. Ma tale apertura diaframmatica non è dovuta a una partecipazione attiva degli articolatori, nel caso specifico del velo e quindi del muscolo *levator palatini* ma, al contrario, è il risultato di un rilassamento di tale muscolo. Sul piano dell'attività neuromuscolare di controllo degli articolatori, quindi, la vocale nasalizzata richiede ancor meno impegno rispetto a una vocale orale, in cui è necessaria la contrazione del *levator palatini* per tutta la sua durata. Per questo motivo la vocale nasalizzata è da considerare come parte vocalica e non consonantica. Naturalmente le consonanti nasali, pur presentando lo stesso rilassamento del *levator palatini*, sono però caratterizzate dalla contrazione di uno dei muscoli orali (ad esempio per la

⁵ Come di norma in procedure di *standard setting*, nell'ottica di incrementare l'affidabilità dell'annotazione manuale, durante le sessioni i due valutatori hanno effettuato dei controlli incrociati sulla segmentazione di alcune porzioni di parlato scelte come campione.

pronuncia della bilabiale [m] sarà contratto l'*orbicularis oris*) e quindi classificate senza dubbio come consonanti.

- La durata delle consonanti occlusive a inizio di catena fonica. Nel caso delle occlusive sonore, il confine è stato identificato con l'inizio delle vibrazioni glottidali, sempre ben visibili sullo spettro; alle occlusive sorde, invece, è stata assegnata una durata pari al valore medio di un'occlusiva sorda scempia in posizione intervocalica nello stesso enunciato.

Figura 2 - Spettrogramma dell'enunciato "si faceva la spesa tutti i giorni" (voce maschile, PD), realizzato in una sequenza del tipo [sia|fes|sət:itʒon:i]. Annotazione manuale. C: intervallo consonantico, V: intervallo vocalico, X: pausa silente

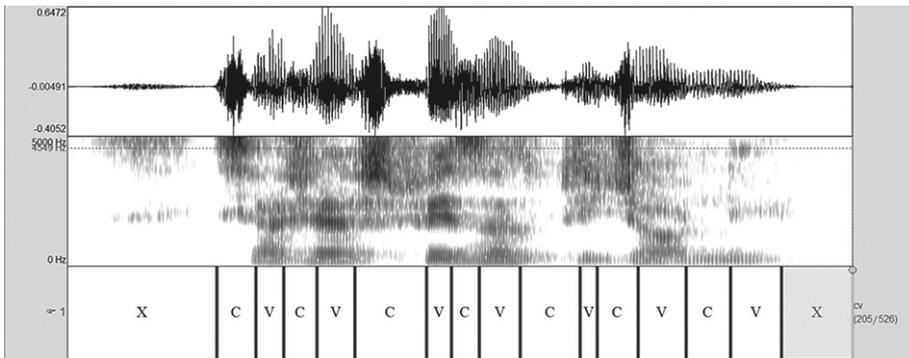
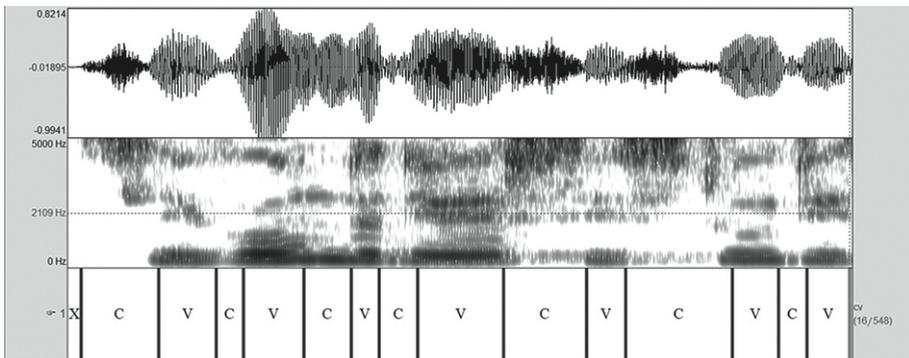


Figura 3 - Spettrogramma dell'enunciato "secondo diversi studi" (voce femminile, PD). Annotazione manuale. C: intervallo consonantico, V: intervallo vocalico, X: pausa silente



Le stesse registrazioni di parlato letto, inoltre, sono state sottoposte a una procedura di annotazione completamente automatica, condotta con il software SPPAS, versione 3.7 (Bigi 2015), già usato per l'annotazione del parlato disartrico in studi sulla lingua francese (Bigi et al. 2015), *open source* e con un'intuitiva interfaccia grafica. A partire da un file audio e dalla trascrizione ortografica degli enunciati, SPPAS è in grado di produrre l'annotazione automatica del segnale su diversi livelli: *IPUs*, cioè *Inter Pausal Units* o catene foniche, parole, sillabe e fonemi. La procedura automa-

tica utilizzata ai fini del presente studio si articola in diverse fasi (riportate in figura 4), che hanno come prodotto una serie di textgrid allineati con il segnale acustico. Ciò permette a SPPAS di dialogare facilmente con Praat.

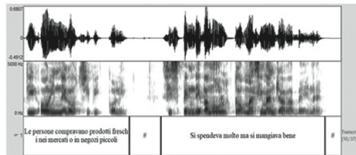
Dopo aver estratto le durate dei segmenti con i due metodi di annotazione, sono stati calcolati per ciascun parlante la %V e il valore medio del VtoV. In entrambe le procedure, sono state annotate anche le disfluenze (false partenze, nasalizzazioni, vocalizzazioni ecc.) e le pause silenti, ma non sono state considerate nel calcolo della %V e del VtoV.

Figura 4 - *La procedura di annotazione automatica in SPPAS*

file audio + trascrizione ortografica



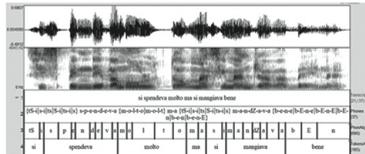
Fill-in IPUs



segmentazione in catene foniche



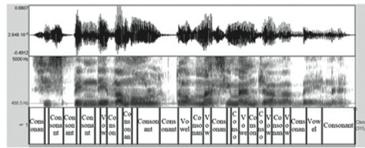
*Text normalization
Phonetization
Alignment*



segmentazione in parole e fonemi



*Plugin
Classification of phonemes*



segmentazione in vocali e consonanti

2.4 L'analisi statistica

T-test per campioni indipendenti sono stati utilizzati per verificare la significatività statistica delle differenze tra i dati dei parlanti parkinsoniani con quelli del gruppo di controllo ricavati con le due annotazioni. Test del *Chi quadrato* sono stati utilizzati per comparare le variabili categoriali. L'analisi statistica è stata condotta con SPSS (versione 23) e la soglia di significatività è stata fissata a 0.05.

3. Risultati

Nella tabella 2 si riportano i valori medi di %V e di VtoV ottenuti nei due gruppi di parlanti, GC e PD, con le due procedure di annotazione.

Tabella 2 - Valori di %V e VtoV nei due gruppi e con le due annotazioni

		GC (n=20) Media±dev.st.	PD (n=20) Media±dev.st.	p-value
annotazione manuale	%V	44.7±1.3	49.8±1.7	<0.001
	VtoV	0.186±0.02	0.185±0.02	0.889
annotazione automatica – SPPAS	%V	43±4.3	45.8±2.6	0.017
	VtoV	0.208±0.03	0.196±0.03	0.201

Si osserva innanzitutto come l'annotazione automatica sia in grado di distinguere efficacemente i due gruppi sulla base del parametro %V, così come avviene con l'annotazione manuale: i valori relativi al parlato patologico risultano significativamente più alti rispetto a quelli del GC, sebbene la significatività sia decisamente più bassa nel caso dei dati ottenuti con SPPAS.

In relazione al VtoV, entrambe le annotazioni non producono risultati che cambiano significativamente nei due gruppi.

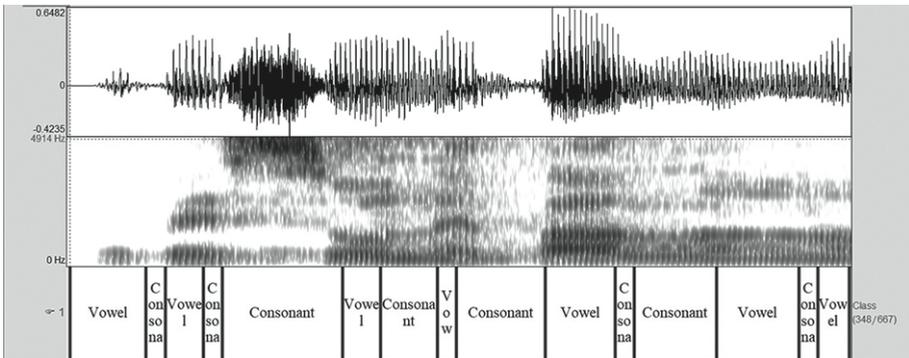
Se la procedura automatica sembra confermare le osservazioni ottenute con la segmentazione manuale del segnale, si nota, tuttavia, che i valori di %V ottenuti con SPPAS sono in media più bassi rispetto a quelli dell'annotazione manuale, con uno scarto maggiore nel caso del gruppo PD (4%) rispetto ai parlanti GC (1.7%).

Per cercare di capire le motivazioni di tali risultati e verificare in che modo SPPAS “perda” porzioni vocaliche o “guadagni” porzioni consonantiche rispetto all'annotazione manuale, si è proceduto a un'analisi qualitativa delle segmentazioni, che si è focalizzata in particolare sulle produzioni dei parlanti che ottengono risultati più divergenti con le due procedure di annotazione. Si riportano di seguito alcune osservazioni preliminari.

Il software riscontra alcune criticità laddove la qualità della registrazione risulta più bassa: un rumore di fondo o il riverbero ambientale, facilmente identificabili nell'analisi fatta da esperti, sono senz'altro elementi che non permettono sempre al sistema automatico di riconoscere con buona approssimazione i diversi foni e di individuarne i confini.

Nei casi, inoltre, di accentuata ipoarticolazione nel parlato disartrico, come quello riportato in figura 5, il software è costretto forzatamente a riconoscere tutti i fonemi rappresentati nella trascrizione ortografica, anche quando essi non siano effettivamente presenti nel segnale acustico: le due consonanti vibranti sono del tutto assenti nell'enunciato prodotto ma il software assegna loro due segmenti, sacrificando una porzione delle vocali che le precedono.

Figura 5 - Spettrogramma dell'enunciato "le persone parlano" realizzato in una sequenza del tipo [epes:o:nepal:ano] (voce maschile, PD). Annotazione automatica in consonanti e vocali



La minore accuratezza articolatoria nelle produzioni dei PD rispetto a quelle dei GC potrebbe, quindi, dar conto dello scarto maggiore nel confronto con i risultati della procedura manuale di annotazione⁶.

Tale difficoltà può essere risolta con una *enriched transcription*, prevista in SPPAS (Bigi & Meunier 2018) ma ancora non sperimentata dagli autori, che tenga conto dell'effettiva pronuncia di ogni singola parola. In una prospettiva a lungo termine di completa automatizzazione della procedura, la soluzione può essere l'ampliamento del dizionario di pronunce a partire dal quale il software applica la procedura di conversione grafema-fonema.

La stessa soluzione può essere applicata anche in alcuni casi di pronuncia marcata, regionale o dialettale, di alcuni termini. In figura 6 si riporta lo spettrogramma dell'enunciato "mangiavamo molto pane", prodotto da una voce femminile del gruppo PD. In questo caso si vede come la segmentazione di SPPAS sia complessivamente corretta, anche se vi è una divergenza con quella manuale relativa alla vocale che segue l'affricata sonora [dʒ] di "mangiavamo". L'errore non è dovuto al software ma all'inserimento di un elemento vocalico di tipo [i] tra l'affricata e la vocale [a], tratto tipico della pronuncia di area campana (la parola è pronunciata [mandʒia'va:mɔ]). Il sistema riconosce un elemento diverso dalla [a] ma lo assegna alla consonante che precede in virtù della minore ampiezza del segnale. Si nota, inoltre, come la porzione nasalizzata della prima [a] di "mangiavamo" non venga considerata nell'intervallo vocalico, come invece avviene nell'annotazione manuale⁷. Infine, anche la [a] di "pane" risulta ridotta nella sua durata.

⁶ Si tratta qui solo di un'ipotesi. Le cause del maggiore scarto riscontrato nel gruppo dei parlanti parkinsoniani tra i valori ottenuti con i due tipi di etichettatura non sono oggetto del presente contributo ma sono sicuramente da approfondire.

⁷ Il nesso V + C nasale, come precedentemente evidenziato, risulta di particolare interesse ai fini del presente studio per le peculiari caratteristiche articolatorie del suono vocalico. Per tale motivo, è assolutamente auspicabile che il software SPPAS possa riconoscere le porzioni nasalizzate delle vocali ed etichettarle come *Vowel*.

Le annotazioni ottenute con il software SPPAS si sono, infatti, dimostrate efficaci nel differenziare i due gruppi di parlanti sulla base del parametro %V sebbene si siano evidenziate alcune criticità e alcune divergenze rispetto alla segmentazione manuale. È possibile ipotizzare che, con una migliore qualità delle registrazioni e con una più accurata trascrizione del parlato, la procedura di annotazione automatica possa funzionare sempre meglio e in maniera sempre più affidabile.

In conclusione, quindi, in questa fase del progetto di ricerca, nell'ottica di ottimizzare i tempi di lavoro, l'utilizzo di una metodologia mista di annotazione, che preveda una prima fase automatica seguita da un controllo manuale, è senz'altro auspicabile. Questo primo test sul corpus di parlato letto, inoltre, ha permesso di raccogliere informazioni utili alla progettazione di un nuovo protocollo di registrazione, più adatto a una procedura di annotazione completamente automatica, che rappresenta l'obiettivo ultimo del progetto.

Ringraziamenti

Un doveroso ringraziamento va a Brigitte Bigi, per il supporto agli autori nell'uso di SPPAS e la grande disponibilità dimostrata.

Bibliografia

- Ackermann, Hermann & Konczak, Jürgen & Hertrich, Ingo. 1997. The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and Language* 56(2). 312-319.
- Beck, Aaron T. & Ward, Calvin H. & Mendelson, Mock & Mock, Jeremiah & Erbaugh, John. 1961. An inventory for measuring depression. *Archives of General Psychiatry* 4(6). 561-571.
- Bigi, Brigitte & Meunier, Christine. 2018. Automatic speech segmentation of spontaneous speech. *Revista de Estudos da Linguagem* 26(4).
- Bigi, Brigitte & Klessa, Katarzyna & Georgeton, Laurianne, & Meunier, Christine. 2015. A syllable-based analysis of speech temporal organization: a comparison between speaking styles in dysarthric and healthy populations. In *Sixteenth Annual Conference of the International Speech Communication Association (INTERSPEECH 2015)*, Dresda. 2977-2981.
- Bigi, Brigitte. 2015. SPPAS – Multilingual Approaches to the Automatic Annotation of Speech. *the Phonetician – International Society of Phonetic Sciences* 111-112 (2015-I-II). 54-69.
- Boersma, Paul & Weenink, David. 2021. *Praat: doing phonetics by computer* [Computer program]. Version 6.1.39, <http://www.praat.org/>.
- Cochen De Cock, Valérie & Dotov, Dobromir & Ihalainen, Petra et al. 2018. Rhythmic abilities and musical training in Parkinson's disease: do they help?. *Parkinson's Disease* 4(8).
- Cohen, Henri. 2003. Disorders of speech and language in Parkinson's disease. In Bédard, Marc-Andre & Agid, Yves & Korczyn, Amos & Lespérance, Paul & Chouinard, Sylvain

(eds.), *Mental and behavioral dysfunction in movement disorders*, 125-134. New York: Humana Press.

Darley, Frederic L. & Aronson, Arnold E., & Brown, Joe R. 1969a. Differential diagnosis patterns of dysarthria. *Journal of Speech and Hearing Research* 12(2). 246-269.

Darley, Frederic L. & Aronson, Arnold E., & Brown, Joe R. 1969b. Cluster of deviant speech dimension in the dysarthrias. *Journal of Speech and Hearing Research* 12(3), 462-469.

de Lau, Lonneke M. & Breteler, Monique M. B. 2006. Epidemiology of Parkinson's disease. *The Lancet. Neurology* 5(6). 525-535.

Duez, Danielle. 2006. Syllable structure, syllable duration, and final lengthening in Parkinsonian French speech. *Journal of Multilingual Communication Disorders* 4. 45-57.

Flostein, 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. 189-198.

Gibb, William R. & Lees, Andrew J. 1988. A comparison of clinical and pathological features of young- and old-onset Parkinson's disease. *Neurology* 38. 1402-1406.

Gili Fivela, Barbara & Iraci, Massimiliano M. & Sallustio, Vincenzo & Grimaldi, Mirko & Zmarich, Claudio & Patrocino, Danilo. 2014. Italian Vowel and Consonant (co) articulation in Parkinson's Disease: extreme or reduced articulatory variability? In *10th International Seminar of Speech Production (ISSP), 5-8 Maggio 2014*, Colonia, Germania. 146-149.

Goberman, Alexander M. & Coelho, Carl A. 2005. Prosodic characteristics of Parkinsonian speech: The effect of levodopa-based medication. *Journal of Medical Speech-Language Pathology* 13. 51-68.

Harel, Brian T. & Cannizzaro, Michael S. & Cohen, Henri & Reilly, Nicole, & Snyder, Peter J. 2004. Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment. *Journal of Neurolinguistics* 17. 439-453.

Hirose, Hajime & Kiritani, Shigeru & Sawashima, Masayuki. 1982. Velocity of articulatory movements in normal and dysarthric subjects. *Folia phoniatrica* 34(4). 210-215.

Liss, Julie M. & White, Laurence & Mattys, Sven L. & Lansford, Kaitlin & Lotto, Andrew J. & Spitzer, Stephanie M. & Caviness, John N. 2009. Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research* 52. 1334-1352.

Logemann, Jeri A. & Fisher, Hilda B. & Boshes, Benjamin & Blonsky, E. Richard. 1978. Frequency and cooccurrence of vocal tract dysfunction in the speech of a large sample of Parkinson patients. *Journal of Speech and Hearing Disorders* 43. 47-57.

Ludlow, Christy L. & Connor, Nadine P. & Bassich, Celia J. 1987. Speech timing in Parkinson's and Huntington's disease. *Brain and Language* 32(2). 195-214.

MacPherson, Megan K. & Huber, Jessica E. & Snow, David P. 2011. The intonation-syntax interface in the speech of individuals with Parkinson's disease. *Journal of Speech, Language, and Hearing Research* 54(1). 19-32.

Maffia, Marta & De Micco, Rosa & Pettorino, Massimo & Siciliano, Mattia & Tessitore, Alessandro & De Meo, Anna. 2021. Speech rhythm variation in early-stage Parkinson's disease: a study on different speaking tasks. *Frontiers in Psychology* 12 (668291).

- Maffia, Marta & De Micco, Rosa & Tessitore, Alessandro & De Meo, Anna. 2020. Acoustic characteristics of Italian Parkinsonian speech: a study on early-stage patients. In Botinis, Antonio (ed.) *Proceedings of the 11th International Conference of Experimental Linguistics – Exling2020*, Atene. 117-120.
- Pettorino, Massimo & Maffia, Marta & Pellegrino, Elisa & Vitale, Marilisa & De Meo, Anna. 2013. VtoV: a perceptual cue for rhythm identification, in Mertens, Piet & Simon, Anne Catherine (eds.), *Proceedings of the Prosody-Discourse Interface Conference 2013 - IDP2013, Leuven*, 101-106.
- Pettorino, Massimo & Busà, Maria Grazia, & Pellegrino, Elisa. 2016. Speech Rhythm in Parkinson's Disease: A Study on Italian. In *Proceedings of the 17th Annual Conference of the International Speech Communication Association*, San Francisco, 1958-1961.
- Pettorino, Massimo & Gu, Wentao & Pólrola, Pawel & Fan, Ping, 2017. Rhythmic Characteristics of Parkinsonian Speech: A Study on Mandarin and Polish. In *Proceedings of the 18th Annual Conference of the International Speech Communication Association*, Stoccolma, 3172-3176.
- 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 *Proceedings of 41st International Conference on Telecommunications and Signal Processing (TSP2018)*, Atene, 1-5.
- Puyjarinet, Frédéric & Bégel, Valentin & Gény, Christian & Pinto, Serge & Dalla Bella, Simone. 2018. Relations between rhythmic impairment in speech and non-verbal domains in Parkinson's disease. *Annals of Physical and Rehabilitation Medicine* 61. e46-e47.
- Ramig, Lorraine O. & Fox, Cynthia & Sapir, Shimon. 2008. Speech treatment for Parkinson disease. *Expert Review of Neurotherapeutics* 8. 299-311.
- Rusz, Jan & Cmejla, Roman & Tykalova, Tereza & Ruzickova, Hana & Klempir, Jiri & Majerova, Veronika, et al. 2013. Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. *The Journal of the Acoustical Society of America* 134(3). 2171-2181.
- Sapir, Shimon & Ramig, Lorraine O. & Spielman, Jennifer L. & Fox, Cynthia. 2010. Formant Centralization Ratio: A proposal for a new acoustic measure of dysarthric speech. *Journal of Speech, Language, and Hearing Research* 53(1). 114-125.
- Skodda, Sabine & Schlegel, Uwe. 2008. Speech rate and rhythm in Parkinson's disease. *Movement Disorders* 23(7). 985-992.
- Skodda, Sabine & Visser, Wenke & Schlegel, Uwe. 2011. Vowel articulation in Parkinson's disease. *Journal of voice: official journal of the Voice Foundation* 25(4). 467-472.
- Tomlinson, Claire L. & Stowe, Rebecca & Patel, Smitaa & Rick, Caroline & Gray, Richard & Clarke, Carl E. 2010. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Movement disorders: official journal of the Movement Disorder Society* 25(15). 2649-2653.