

Università Popolare Cattolica
Eustachio Montemurro-Teresa D'Ippolito

(DPR 23/05/23 ex Legge 13/1991 e Dpcm 28/01/11; Decreto Giunta Regione Campania n°38/23 ex Dlgs 117/17)

"Applicazione del Metodo NEPST con Intelligenza Artificiale nella gestione terapeutica della balbuzie"

(di Emanuele Giuseppe ADILETTA, Ciro ROMANO)

Abstract

Il presente studio sperimentale è stato approvato dal Comitato Etico Inter-Aziendale dell'*Università Popolare Cattolica "Montemurro-D'Ippolito"* di Portici (Na). Il disegno dello studio segue scrupolosamente gli standard CONSORT previsti per i trial randomizzati, adottando un modello sperimentale parallel-group con un controllo attivo. La randomizzazione è stata stratificata, utilizzando blocchi di dimensione sei e assegnazione casuale in rapporto 1:1.

Hanno partecipato allo studio cinquantasette individui madrelingua italiana, di cui ventisette femmine e trenta maschi, con età variabile dai tredici ai sessant'anni e una media di 28,4 anni con deviazione standard di 12,7. I partecipanti dovevano soddisfare i criteri diagnostici di balbuzie evolutiva persistente secondo il test SSI-4, con un punteggio uguale o superiore a 14, dimostrare udito nella norma, non presentare disturbi neurologici significativi, avere un QI verbale superiore o uguale a 85, e dichiarare disponibilità all'utilizzo quotidiano di uno smartphone con sistema operativo Android 11 o superiore. Sono stati esclusi coloro che avevano ricevuto trattamenti logopedici intensivi nei tre mesi precedenti o presentavano comorbidità psichiatriche non stabilizzate o utilizzo di farmaci che potessero influire sulla neurovascolarizzazione cerebrale. Tutti i partecipanti e, nel caso di minori, i loro tutori legali, hanno fornito un consenso informato scritto conforme ai principi della Dichiarazione di Helsinki del 2013.

I soggetti sono stati casualmente assegnati a due gruppi distinti: il gruppo sperimentale NEPST-AI, che utilizzava l'applicazione adattiva, e il gruppo di controllo attivo NEPST-fixed, che riceveva un intervento tradizionale con esercizi predefiniti. Entrambi i gruppi hanno seguito un protocollo terapeutico della durata complessiva di dodici settimane, con sessioni cliniche svolte in presenza nelle settimane iniziale, intermedia e finale, integrate da una pratica quotidiana effettuata tramite l'applicazione mobile.

La particolarità del gruppo NEPST-AI consisteva nell'utilizzo di un'applicazione sviluppata in collaborazione con il Montemurro Digital Lab, basata su algoritmi di apprendimento per rinforzo (Proximal Policy Optimization). Questa applicazione rilevava in tempo reale dati fonetici, neuromuscolari e psicofisiologici, e modulava la difficoltà degli esercizi in base a tali parametri. In questo modo, ogni partecipante riceveva un percorso terapeutico personalizzato che adattava la complessità delle attività alla performance e ai feedback fisiologici registrati. Nel gruppo di controllo NEPST-fixed, invece, gli esercizi seguivano una sequenza e un livello di difficoltà prestabiliti, indipendentemente dalle risposte fisiologiche del soggetto.

Le sessioni cliniche iniziali e finali comprendevano una serie di valutazioni laboratoristiche. Venivano effettuate registrazioni EEG ad alta densità durante compiti fonemici, utilizzando il sistema HydroCel a 64 canali, analisi della componente P300, e valutazioni tramite fNIRS per l'analisi emodinamica corticale durante letture standardizzate. Inoltre, il Doppler transcranico bilaterale era impiegato per valutare l'asimmetria nella velocità media del flusso ematico cerebrale durante attività di naming, seguendo il protocollo stabilito da Knecht e colleghi (1998).

I partecipanti hanno inoltre compilato diversi questionari psicometrici, tra cui lo State-Trait Anxiety Inventory (STAI-Y), il Self-Statements during Public Speaking (SSPS), e l'Overall Assessment of the Speaker's Experience of Stuttering (OASES), assieme a un questionario specificamente ideato per valutare la user-experience relativa all'applicazione utilizzata. Ulteriori dati sul parlato libero e standardizzato sono stati registrati e analizzati tramite il software PRAAT da due valutatori indipendenti, verificando l'affidabilità inter-rater, risultata superiore a 0,93 nei test pilota preliminari. Ogni sessione clinica prevedeva momenti di feedback e pratica supervisionata dei trial neuromodulatori, inclusi esercizi specifici quali lettura rallentata, allungamento vocalico, tecniche di onset dolce, doppiaggio emotivo controllato, esercizi dual-task con n-back visuo-spaziale e shadowing. Queste metodiche si basavano sull'evidenza scientifica che esercizi dual-task migliorano significativamente la fluenza in assenza di competizione per risorse fonologiche.

Le analisi statistiche programmate comprendono modelli lineari misti con random-slope per soggetto, considerando l'interazione tra gruppo e tempo, e analisi bayesiane per valutare la probabilità di ottenere una riduzione clinicamente significativa delle disfluenze. Correzioni statistiche appropriate sono state pianificate per controllare l'errore multiplo nella valutazione dei dati EEG, fNIRS e tCD, assicurando la robustezza metodologica.

Dal punto di vista etico e della sicurezza algoritmica, il sistema adattivo utilizzato nell'app è stato progettato rispettando le direttive di explainability definite dallo standard ISO/IEC 24027:2021. Le decisioni operative del sistema erano sottoposte a costante monitoraggio clinico e amministrativo, garantendo la sicurezza e la protezione dei dati personali secondo le normative GDPR 2016/679.

Questo studio innovativo, fondato sul metodo NEPST sviluppato dal Prof. Adiletta, propone una concreta integrazione fra terapia logopedica e intelligenza artificiale, potenzialmente trasformando l'approccio alla gestione della balbuzie e migliorando significativamente la qualità della vita dei partecipanti.

Panoramica e scopo dello studio

Negli ultimi due decenni, il trattamento della balbuzie ha attraversato una fase di transizione metodologica significativa, evolvendo da approcci prevalentemente comportamentali e fonetici a modelli neurocognitivi complessi, che tengono conto della coattivazione simultanea di domini motori, esecutivi ed emozionali. Tuttavia, la difficoltà principale nel rendere clinicamente operativi questi modelli risiede nella mancanza di protocolli dinamici, adattivi e quantificabili, capaci di interagire con l'eterogeneità temporale e neurofunzionale della disfluenza. All'interno di questo scenario, il modello NEPST (Neuro-Executive Phonological Stress Tuning), ideato dal prof. Raffaele Adiletta, rappresenta un tentativo sistematico di integrare queste componenti in un paradigma terapeutico ad apprendimento incrementale, capace di modulare in tempo reale le condizioni di esposizione fonatoria in funzione delle risposte cognitive, fisiologiche e affettive del soggetto.

Il presente studio si configura come la prima valutazione sperimentale controllata del protocollo NEPST, implementato in un framework digitale potenziato da algoritmi di reinforcement learning. L'interesse centrale non è unicamente testare l'efficacia clinica del trattamento in termini di miglioramento della fluenza, ma comprendere se e in che modo l'autoregolazione adattiva dell'input terapeutico — fondata su segnali vocali e neurobiometrici — possa indurre modificazioni plastiche nei sistemi neurali responsabili della pianificazione articolatoria, della regolazione attentiva e dell'integrazione interemisferica. In altre parole, lo studio mira a verificare se l'applicazione intelligente di un carico fonatorio ottimale, individualizzato e fluttuante, possa fungere da catalizzatore per una riorganizzazione funzionale della rete linguistica.

Più specificamente, lo scopo dello studio è triplice. In primo luogo, si intende valutare se il gruppo NEPST-AI — in cui l'algoritmo modula dinamicamente i parametri di esercizio — mostra miglioramenti di fluenza più rapidi, stabili e generalizzabili rispetto al gruppo NEPST-fixed, sottoposto a una progressione standardizzata. In secondo luogo, si esplora l'impatto dell'intervento su biomarcatori neurofisiologici critici, quali la latenza della P300, la lateralizzazione emisferica del flusso cerebrale medio (tCD) e le variazioni emodinamiche fronto-temporali (fNIRS), con l'obiettivo

di stabilire una mappatura tra l'efficacia clinica e le traiettorie neurali sottostanti. Infine, si analizza il contributo di indici autonomici e affettivi (HRV, EMG facciale, engagement soggettivo) nella modulazione del beneficio terapeutico, assumendo che la regolazione emotiva e fisiologica sia una condizione necessaria — sebbene non sufficiente — per la ristrutturazione delle sequenze fonatorie implicite.

Il valore metodologico di questo studio risiede nel tentativo di saldare la logica adattiva dei sistemi dinamici non lineari con la rigosità della verifica sperimentale neurocomportamentale. Attraverso un disegno longitudinale a misure ripetute, il progetto NEPST-AI non si limita a osservare se il paziente “migliora”, ma si interroga su come e quando si verifica tale cambiamento, quali sistemi neurali vengono progressivamente coinvolti, e quali condizioni personali e contestuali ne favoriscono o ostacolano l'emergere. In quest'ottica, lo studio non è solo una prova di efficacia, ma una proposta di ridefinizione epistemica del trattamento della balbuzie: non più correttivo e prescrittivo, ma trasformatore e neurorisonante.

In sintesi, la ricerca presentata intende contribuire allo sviluppo di un modello terapeutico di nuova generazione, fondato su plasticità misurabile, adattività algoritmica clinicamente guidata, e rilevazione integrata dei segnali cognitivi, motori e affettivi. Il NEPST-AI rappresenta, in questo senso, una possibile frontiera per la riabilitazione fonatoria intelligente e, più in generale, per una psicologia clinica della voce fondata su basi neurocomputazionali. Lo studio NEPST-AI nasce all'intersezione tra neuroscienze applicate, linguistica clinica, tecnologia adattiva e psicologia della voce. Il punto di partenza è un riconoscimento esplicito del limite dei modelli lineari e unidimensionali nella comprensione e nel trattamento della balbuzie. Per decenni, la disfluenza è stata concettualizzata come un disturbo isolabile e localizzabile — del respiro, dell'articolazione, dell'intenzione comunicativa o dell'emotività sociale — con approcci terapeutici fondati su protocolli standardizzati e statici. In realtà, la balbuzie si configura sempre più come un fenomeno sistemico a espressività variabile, in cui pattern fonatori disfunzionali emergono da una interazione instabile tra reti motorie, sistemi di controllo esecutivo, meccanismi di monitoraggio attentivo e processi affettivo-motivazionali, spesso amplificati dalla dimensione interpersonale della voce.

In tale contesto, lo scopo dello studio NEPST-AI non è semplicemente testare un nuovo trattamento, ma tentare una riconfigurazione epistemica del modo in cui si studia e si interviene sulla balbuzie. L'ipotesi di fondo è che per modificare stabilmente la produzione linguistica in soggetti disfluenti non sia sufficiente intervenire sul sintomo, bensì occorra agire sui vincoli cognitivi e fisiologici che ne sostengono la persistenza, in modo adattivo, personalizzato e dinamico. Questo implica un salto di paradigma: dalla prescrizione lineare di esercizi verso un'interazione in tempo reale tra soggetto e sistema terapeutico, dove la voce viene trattata come una variabile emergente di uno spazio multifattoriale.

Il modello NEPST (Neuro-Executive Phonological Stress Tuning) nasce con questo intento. Esso integra una prospettiva neuroesecutiva (attenzione, carico di lavoro fonemico, inibizione di risposta) con la regolazione motoria implicita (tempismo fonatorio, legatura sillabica, stress vocale) e la modulazione affettiva (engagement, sicurezza, sintonizzazione interocettiva). L'assunto teorico è che la balbuzie si mantenga anche perché il soggetto sviluppa una iper-focalizzazione maladattiva su segmenti fonologici critici, alimentando un ciclo di controllo esplicito che interferisce con i meccanismi automatici del parlato fluente. In tal senso, l'intervento deve favorire una defocalizzazione intenzionale, una sorta di "distrazione terapeutica" che liberi il gesto fonatorio da vincoli di monitoraggio eccessivo.

L'applicazione digitale NEPST-AI si propone di implementare questo principio attraverso un sistema di apprendimento per rinforzo (reinforcement learning) che, monitorando continuamente le prestazioni vocali e i parametri fisiologici del soggetto, adatta la difficoltà dell'esercizio in modo da mantenere la soglia ottimale tra sfida e controllo. Il modello viene così *co-addestrato* insieme al soggetto: l'algoritmo apprende a regolare il carico, il soggetto impara a disattivare il controllo eccessivo, e l'interazione reciproca costituisce il fulcro dell'intervento.

Lo scopo primario dello studio è verificare se tale interazione adattiva produce un miglioramento clinico della fluenza maggiore, più rapido e più stabile rispetto a un trattamento tradizionale a progressione fissa. Ma lo studio non si esaurisce in una misurazione comportamentale. Attraverso EEG, fNIRS e doppler transcranico, l'obiettivo è individuare marcatori neurali del cambiamento, in particolare nella temporizzazione attentiva (latenza P300), nella lateralizzazione emisferica del linguaggio e nella perfusione frontale sotto carico semantico. L'ipotesi è che, se il miglioramento è realmente legato a una riorganizzazione interna, allora esso sarà accompagnato da segni misurabili di plasticità funzionale nei circuiti fonatori e prefrontali.

Lo scopo secondario è comprendere il ruolo delle variabili fisiologiche e affettive (HRV, attivazione simpatica, espressione mimico-facciale, engagement soggettivo) come predittori o mediatori dell'esito. La balbuzie non è solo un disturbo tecnico del parlato, ma una frattura nel rapporto tra voce e corpo, tra intenzione e possibilità. Recuperare la fluenza, in questa visione, è anche ristabilire una coerenza incarnata tra il respiro, il ritmo, l'intonazione e l'identità vocale. Il protocollo si propone quindi anche di esplorare se il cambiamento terapeutico sia preceduto o accompagnato da una maggiore integrazione neurovegetativa e un abbassamento delle barriere difensive fisiologiche.

Infine, su un piano traslazionale, lo studio ambisce a dimostrare che una IA clinicamente addestrata, non autonoma ma vincolata a policy etiche e cliniche, possa agire come moltiplicatore di efficacia, portando la personalizzazione terapeutica a un livello che nessun protocollo standard può garantire. Questo apre scenari concreti di riabilitazione a distanza, scalabile, adattiva e integrabile in contesti con limitate risorse umane. Tuttavia, lo scopo ultimo non è tecnologico: è scientifico e clinico. Si tratta di spostare il focus della riabilitazione vocale dal "parlare meglio" al "parlare con meno conflitto

interno”, creando uno spazio in cui il soggetto possa esperire la propria voce non come una minaccia, ma come una continuità del sé.

In conclusione, lo studio NEPST-AI non si propone solo di rispondere alla domanda “funziona?”, ma di indagare *come, quando, per chi e a che costo neurocognitivo* si verifica un cambiamento reale, stabile e generalizzabile della fluenza. È un progetto che unisce neuroscienza quantitativa e fenomenologia della voce, intelligenza artificiale e relazione terapeutica, psicofisiologia e plasticità esperienziale. Il suo scopo non è dimostrare una verità, ma tracciare una possibilità: quella di una voce che cambia perché finalmente *viene ascoltata*.

La balbuzie

La balbuzie rappresenta una disfluenza del linguaggio parlato caratterizzata da interruzioni involontarie del flusso verbale, come ripetizioni di suoni o sillabe, prolungamenti fonemici e blocchi articolatori. Tradizionalmente considerata un disturbo dell'articolazione o del ritmo vocale, oggi la balbuzie viene sempre più compresa come un fenomeno complesso, multifattoriale e dinamico, che coinvolge la co-implicazione di sistemi neurologici, cognitivi, emotivi e sociali. Essa non si configura come una semplice "mancanza di fluidità", ma come una condizione neuropsicologica in cui il parlante sperimenta una rottura nella coerenza tra intenzione comunicativa e possibilità espressiva, con ripercussioni che vanno ben oltre l'atto fonatorio.

A livello clinico, la balbuzie viene solitamente distinta in due grandi categorie: quella evolutiva, che insorge nei primi anni di vita e può persistere fino all'età adulta, e quella acquisita, di origine neurologica o farmacologica, più rara e generalmente secondaria a lesioni cerebrali o condizioni degenerative. La forma evolutiva rappresenta oltre il 90% dei casi e si manifesta tipicamente tra i 2 e i 5 anni d'età, in coincidenza con la fase di rapida acquisizione linguistica. Sebbene circa l'80% dei bambini disfluenti superi spontaneamente la fase di balbuzie transitoria, una parte significativa dei soggetti sviluppa forme persistenti, che tendono a stabilizzarsi nell'adolescenza e nell'età adulta, soprattutto in assenza di trattamento precoce e mirato.

Dal punto di vista neurologico, studi di neuroimaging e neurofisiologia hanno evidenziato nella balbuzie persistente un pattern funzionale alterato che coinvolge la rete cortico-sottocorticale del linguaggio, in particolare le aree premotorie, il giro frontale inferiore sinistro (area di Broca), i gangli della base e il circuito talamo-cerebellare. È stato osservato un aumento dell'attività emisferica destra in compiti verbali, interpretato come un meccanismo compensatorio inefficiente, nonché una riduzione della connettività strutturale nel fascicolo arcuato sinistro, fondamentale per la sincronizzazione temporale tra percezione uditiva e articolazione. A ciò si aggiungono disfunzioni nel timing motorio, nella pianificazione fonemica e nella capacità di inibizione motoria, che contribuiscono alla frammentazione del flusso linguistico. La balbuzie, quindi, non è soltanto una

questione di “errore” articolatorio, ma di instabilità nei meccanismi di feedforward e feedback neurale che regolano la produzione automatica del parlato.

A livello cognitivo, i parlanti disfluenti tendono a sviluppare un controllo eccessivo e spesso disfunzionale sul proprio eloquio, sostenuto da una sovra-attivazione delle risorse esecutive e della memoria di lavoro verbale. Questa iper-monitorizzazione, pur mossa da un'intenzione adattiva (evitare l'errore), paradossalmente incrementa la probabilità di disfluenza, specialmente nei momenti in cui l'attenzione viene diretta volontariamente verso le componenti meccaniche del parlare. Tale effetto è coerente con il principio della “constrained action hypothesis”, secondo cui il focus attentivo interno verso il movimento interferisce con l'automatismo motorio, riducendo la fluidità in performance apprese (Wulf & Lewthwaite, 2016). Studi su paradigmi dual-task hanno mostrato che la disfluenza tende a diminuire quando il carico attentivo è spostato su un compito non verbale, confermando che l'interferenza non deriva da un deficit di capacità ma da un'anomalia nella gestione del controllo.

La dimensione emotiva della balbuzie è altrettanto centrale. L'esperienza reiterata di fallimento comunicativo, l'imprevedibilità del blocco e l'anticipazione ansiosa dell'evento disfluente generano un ambiente psichico pervaso da evitamento, autocensura e ipervigilanza. Le neuroscienze affettive suggeriscono che la balbuzie si associa a una disregolazione dell'asse stress-voce, con attivazione del sistema simpatico, riduzione della variabilità cardiaca e coattivazione tonica dei muscoli oro-facciali. Questo assetto neurofisiologico di “prontezza inibita” costituisce uno sfondo somatico che alimenta il circolo vizioso tra emozione e performance, e che può cronicizzarsi in assenza di interventi integrati. Non sorprende, quindi, che la balbuzie sia fortemente correlata a sintomi ansiosi, disturbi dell'autostima, ritiro sociale e, nei casi più gravi, tratti fobici o depressivi.

Oltre alla dimensione individuale, la balbuzie è anche un fenomeno profondamente intersoggettivo. Parlare implica sempre un'esposizione, un'offerta della voce all'altro. Il soggetto disfluente, consapevole della propria vulnerabilità fonatoria, costruisce spesso strategie di evitamento comunicativo, riformulazione del messaggio, uso selettivo del lessico, e talvolta silenzio come difesa. Queste scelte adattive, sebbene funzionali a breve termine, possono limitare l'accesso all'esperienza dialogica autentica, compromettendo la spontaneità espressiva e il senso di appartenenza relazionale. In questo senso, la balbuzie non è solo un disturbo dell'espressione vocale, ma anche una ferita nella relazione fonatoria con l'altro, che incide profondamente sull'identità e sulla rappresentazione di sé. In conclusione, la balbuzie si configura oggi come un disturbo ad alta complessità neurobiologica e psicodinamica, che sfugge alle categorizzazioni semplicistiche. È una condizione che richiede interventi transdisciplinari, capaci di operare su più livelli simultaneamente: motorio, cognitivo, affettivo e relazionale. Comprendere la balbuzie significa quindi non soltanto misurare la sua entità acustica, ma anche interpretare le sue radici neurali, le sue implicazioni corporee, le sue trasformazioni identitarie e il suo impatto soggettivo sul vivere vocale. È in questa prospettiva che si

colloca il presente lavoro: non nella riduzione del sintomo a parametro, ma nella costruzione di un sapere complesso sulla voce che inciampa — e che, inciampando, rivela la tensione originaria tra parola e corpo. A completare la complessità già delineata, va considerato il ruolo dell'interocezione fonatoria, ovvero la capacità del soggetto di percepire e interpretare segnali corporei legati all'attività vocale, come la tensione laringea, il carico respiratorio, o la vibrazione subglottica. In molti parlanti disfluenti, si osserva una disconnessione tra l'evento motorio vocale e la consapevolezza corporea dello stesso, come se la voce accadesse “fuori tempo” rispetto all'esperienza soggettiva. Tale alterazione può generare un paradosso percettivo: l'anticipazione dell'errore vocale attiva risposte di controllo che, lungi dal prevenire il blocco, lo precipitano. Questo loop di iper-monitoraggio interocettivo può spiegare perché molti soggetti riferiscano di percepire l'arrivo del blocco già prima di iniziare a parlare, segnalando una preattivazione ansiosa dei circuiti somatici coinvolti.

In parallelo, alcune recenti ricerche sul default mode network (DMN) suggeriscono che la balbuzie potrebbe implicare anche una compromissione della capacità di “disattivare” reti cerebrali di tipo introspettivo durante compiti linguistici. In condizioni normali, il DMN — attivo durante stati mentali autoriflessivi — viene inibito durante l'esecuzione di attività orientate all'esterno, come il parlato fluente. Nei soggetti disfluenti, tuttavia, questa inibizione risulta spesso parziale, determinando una interferenza tra auto-osservazione e azione linguistica, fenomeno coerente con l'esperienza soggettiva del “sentirsi mentre si parla”, vissuto come fonte di ansia paralizzante. Tale compromissione nel passaggio tra modalità di riposo e task-positive mode rafforza l'idea che la balbuzie sia una disfunzione della transizione tra stati neurali, più che una mera incapacità articolatoria.

Sul piano clinico, questo ha implicazioni fondamentali: trattare la balbuzie significa non solo allenare una maggiore automatizzazione fonatoria, ma anche promuovere la flessibilità neurocognitiva nella modulazione tra attenzione interna ed esterna, tra presenza corporea e intenzionalità comunicativa. Gli approcci terapeutici che integrano biofeedback, mindfulness fonatoria e training di shifting attentivo potrebbero agire su questa dimensione più profonda, facilitando il rientro del soggetto in una temporalità vocale sincronica e vitale.

Ulteriormente, il concetto di identità vocale acquista rilievo nella comprensione dell'impatto soggettivo della balbuzie. Per molti parlanti, la voce balbuziente non è semplicemente inefficace, ma estranea, imbarazzante, vissuta come uno scarto traumatico rispetto all'immagine di sé. In questo senso, il disturbo fonatorio si tramuta in una alienazione fonica, un'esperienza in cui il soggetto è contemporaneamente autore e ostaggio del proprio suono. Questa scissione interna — tra voce pensata e voce ascoltata, tra intenzione e emissione — può diventare il nucleo latente di un ritiro esistenziale dalla dimensione dialogica, con effetti cumulativi sul benessere psicologico, sulla progettualità sociale e persino sull'architettura narrativa della propria storia di vita.

Infine, è fondamentale riconoscere la natura oscillante e situata della balbuzie: essa non si presenta in modo omogeneo, ma varia in base al contesto, all'interlocutore, alla postura affettiva, alla pressione performativa. Questa variabilità rende la balbuzie un disturbo eminentemente relazionale e richiede che ogni valutazione diagnostica e ogni intervento terapeutico siano pensati non in astratto, ma nella concretezza situata della biografia comunicativa del soggetto. Le neuroscienze, in questo senso, non bastano da sole: devono dialogare con la clinica, con l'analisi del discorso, con la fenomenologia del corpo parlante, se vogliono cogliere la piena realtà della voce che si blocca. Dal punto di vista logopedico, la balbuzie viene spesso affrontata come una disfunzione della coordinazione pneumo-fono-articolatoria, in cui le componenti respiratoria, fonatoria e articolatoria non riescono a integrarsi in modo fluido e automatizzato. I trattamenti logopedici tradizionali mirano a ripristinare tale coordinazione attraverso tecniche che riducono la velocità dell'eloquio, rafforzano la propriocezione fonatoria e promuovono un'articolazione rilassata. Tuttavia, interventi meramente procedurali mostrano limiti nei soggetti con balbuzie resistente, poiché non affrontano il nucleo disfunzionale rappresentato dal controllo volontario eccessivo esercitato sul gesto fonatorio. In tale contesto, si fa sempre più strada l'approccio neuro-logopedico, che integra tecniche comportamentali con training attentivo, modulazione del respiro addominale e strategie di consapevolezza interocettiva. In particolare, le tecniche di onset dolce, rallentamento ritmico e shadowing non sono solo esercizi meccanici, ma veri e propri dispositivi di riscrittura motoria che mirano a ridurre il carico corticale frontale, facilitando l'ingresso della voce nella rete automatica del linguaggio.

Sul piano fonologico, la balbuzie manifesta un pattern di vulnerabilità specifica a determinati segmenti e contesti linguistici. È stato osservato che i blocchi fonatori tendono a concentrarsi su consonanti plosive, suoni ad alta tensione articolatoria e all'inizio delle parole o delle frasi. Questo suggerisce una fragilità nella fase di codifica fonologica a breve termine, specialmente nelle situazioni a carico cognitivo elevato. Modelli teorici come la EXPLAN (Howell, 2004) propongono che la balbuzie emerga quando il planning fonologico non riesce a tenere il passo con l'esecuzione motoria, causando un vuoto che viene colmato da ripetizioni o blocchi. A supporto di tale modello, si trovano evidenze elettrofisiologiche di latenze P300 prolungate e aumento dell'errore fonemico interno, che indicano un rallentamento nei meccanismi di anticipazione fonologica e un disallineamento tra rappresentazione mentale del parlato e attuazione neuromuscolare.

A livello biologico, la balbuzie è associata a un pattern neurochimico e genetico peculiare. Studi recenti hanno identificato varianti genetiche rare nei geni GNPTAB, GNPTG e NAGPA, coinvolti nel metabolismo lisosomiale e nel traffico intracellulare, suggerendo un'origine biologica non esclusivamente legata al linguaggio ma a una vulnerabilità sistemica del sistema neuronale. Inoltre, alterazioni nei livelli di dopamina nelle vie nigro-striatali sono state correlate a disfunzioni della temporizzazione motoria vocale, con risposte favorevoli in alcuni pazienti all'uso di farmaci dopaminergici a basso dosaggio, come risperidone o olanzapina. Tali evidenze suggeriscono che la

balbuzie possa essere letta come una disfunzione a bassa soglia dell'equilibrio neurochimico tra attivazione e inibizione, in cui una modulazione non ottimale della dopamina interferisce con la fluidità motoria del linguaggio.

Infine, le reti neurali coinvolte nella balbuzie mostrano una profonda riorganizzazione rispetto ai pattern tipici dei parlanti fluenti. L'attività anomala è stata documentata non solo nelle aree canoniche del linguaggio (Broca, Wernicke, SMA), ma anche nelle reti esecutive dorsali (DLPFC), nel circuito salience (insula anteriore, ACC) e nella rete limbica (amigdala, ippocampo). Questa distribuzione diffusa suggerisce che la balbuzie coinvolga l'interazione tra la rete linguistica, quella attentiva e quella emotivo-motivazionale. Studi di connettività funzionale (fMRI resting-state) mostrano una maggiore sincronizzazione intraemisferica destra e una minore connettività interemisferica nella regione callosale, segno di una lateralizzazione atipica del linguaggio. Questo dato è coerente con le osservazioni cliniche secondo cui la balbuzie tende a migliorare durante il canto o la recitazione, attività che implicano un reclutamento simmetrico o destrorso delle risorse neurali. Tali fenomeni aprono la strada all'uso di tecnologie neuro-modulanti come la tDCS e la stimolazione magnetica transcranica (TMS) per ripristinare l'equilibrio interemisferico e facilitare la deautomatizzazione dei pattern disfunzionali.

In conclusione, l'approccio moderno alla balbuzie non può prescindere da una comprensione integrata che abbracci i livelli fonologici, neuromuscolari, biologici e reticolari. È solo nel riconoscimento della complessità multilivello del disturbo che si può formulare una strategia terapeutica realmente personalizzata, capace di andare oltre la riduzione sintomatica per toccare le radici neurocognitive della voce che si blocca. Le recenti indagini di neuroimaging funzionale e strutturale hanno rivelato che la balbuzie non può essere considerata una patologia localizzata, ma va compresa come una disorganizzazione della dinamica di rete, ovvero come una condizione in cui l'integrazione temporale e funzionale tra aree cerebrali linguistiche, motorie, esecutive ed emotive appare disturbata in modo sistematico. Un elemento centrale in questa riconcettualizzazione è il ruolo delle reti a stato transitorio, ossia dei pattern dinamici di co-attivazione neurale che supportano il passaggio da uno stato di riposo a uno stato operativo, come quello del parlato spontaneo. Nei soggetti con balbuzie, questi passaggi risultano instabili, con una maggiore latenza nell'attivazione coordinata tra aree e una persistenza disfunzionale di co-attivazioni limbiche, in particolare dell'amigdala e della corteccia cingolata anteriore, che interferiscono con la sequenza motoria.

Uno dei principali modelli teorici oggi in discussione è quello delle disconnessioni funzionali anticipatorie, secondo cui il parlante balbuziente mostrerebbe un fallimento nell'inizializzazione simultanea di nodi critici della rete vocale: in particolare, la corteccia premotoria sinistra, l'area motoria supplementare (SMA), e la regione insulare anteriore, tutte strutture fondamentali per la pianificazione e l'innescio del gesto fonatorio. Questa asincronia provoca un ritardo microtemporale nel feedforward motorio, che il soggetto cerca di colmare attraverso un ricorso iperattivo al feedback

uditivo e propriocettivo — meccanismo che, invece di aiutare, ostacola ulteriormente la fluidità espressiva. L'eccesso di monitoraggio, come già ipotizzato da Max e collaboratori, genera un sovraccarico nei circuiti di controllo dorsolaterali, costringendo il soggetto a un continuo riaggiustamento del gesto linguistico, spesso inefficace.

Ulteriori studi su connettività a lungo raggio, mediante fMRI dinamica e ICA (Independent Component Analysis), hanno dimostrato che nei parlanti disfluenti si osserva una iperconnessione aberrante tra il putamen, il globus pallidus e il talamo, con un'alterata sincronizzazione con le aree corticali superiori deputate al linguaggio. Questo suggerisce che il sistema cortico-basalganglia-talamo-corticale, fondamentale per la regolazione automatica dei movimenti appresi, non riesca a stabilizzare le sequenze fonatorie nel parlante balbuziente. I pattern oscillatori risultano debolmente armonizzati, specialmente nelle bande beta (15–30 Hz), la cui coerenza è tipicamente associata alla stabilizzazione dei gesti motori automatici.

Un ruolo emergente è ricoperto anche dalla rete salience, in particolare dall'insula anteriore e dalla corteccia cingolata anteriore, responsabili della transizione tra reti default-mode e reti task-positive. Nei soggetti disfluenti si osserva spesso una iperattività tonica della rete salience, probabilmente legata alla vigilanza affettiva e all'anticipazione dello sforzo vocale. Questo stato di iperarousal costante può generare un disallineamento tra attivazione interna e domanda ambientale, con conseguente inefficienza della risposta motoria vocale.

Dal punto di vista della connettività interemisferica, la balbuzie mostra un'anomalia nel corpus callosum — specialmente nelle fibre che connettono le aree temporali e prefrontali bilaterali — compromettendo il sincronismo tra le due metà del sistema fonatorio. Il fenomeno è compatibile con i dati doppler transcranici che evidenziano, nei soggetti disfluenti, un pattern di asimmetria instabile nel flusso cerebrale medio, associato a una lateralizzazione verbale meno definita. Questo dato rafforza l'ipotesi che la balbuzie non sia solo un disturbo dell'attivazione, ma anche della sincronizzazione emisferica, fondamentale per la fluidità sequenziale del linguaggio.

Recenti modelli computazionali, infine, propongono che la balbuzie possa essere letta come una perturbazione di sistemi neurali di tipo metastabile, in cui la transizione tra stati funzionali non è guidata da attrattori stabili ma da configurazioni fluttuanti e scarsamente predicibili. In tale prospettiva, il parlato fluente richiede una forma di coordinazione distribuita tra sistemi neurali, paragonabile alla coerenza di fase tra oscillatori. Nei soggetti balbuzienti, tale coordinazione risulta vulnerabile allo stress cognitivo, emotivo e interpersonale, causando un crollo temporaneo del sistema verso stati disorganizzati, ovvero episodi di disfluenza.

L'insieme di questi dati invita a una nuova ontologia del disturbo fonatorio, dove la balbuzie non è un'anomalia localizzabile ma una disfunzione emergente a livello di rete, con pattern dinamici di vulnerabilità che si attivano in risposta a specifici set di condizioni interne ed esterne. Da qui deriva la necessità di trattamenti flessibili, capaci di leggere in tempo reale l'attività di rete del soggetto,

calibrando gli input terapeutici non in base a sequenze fisse, ma in base a traiettorie individuali di stabilizzazione neurale. Queste evidenze sono ulteriormente corroborate da studi longitudinali che utilizzano la risonanza magnetica funzionale a riposo per valutare la coerenza della rete fronto-temporo-parietale. Alm (2004), ad esempio, ha evidenziato che la balbuzie è associata a un'attivazione anomala del giro frontale inferiore sinistro (area di Broca) e del giro temporale superiore (area di Wernicke), con un'eccessiva attivazione dell'emisfero destro omologo, suggerendo un meccanismo compensativo maladattivo piuttosto che funzionale. Questo squilibrio interemisferico si traduce in una produzione fonatoria meno automatica e più volontaria, che implica un carico esecutivo elevato, spesso insostenibile in situazioni ad alta pressione sociale.

Un ulteriore contributo è offerto dalle ricerche di Chang, Erickson e Ambrose (2008), che hanno documentato un'alterata microstruttura della sostanza bianca nella via arcuata, in particolare nella porzione anteriore dell'emisfero sinistro. Utilizzando la diffusione dell'imaging tensore (DTI), gli autori hanno osservato una riduzione della frazione anisotropica nelle fibre che connettono l'area di Broca e l'area motoria supplementare, suggerendo un'anomalia della connettività neurale essenziale per l'integrazione del linguaggio motorio. Queste disconnessioni strutturali sono considerate oggi uno dei biomarcatori più solidi della balbuzie persistente (Watkins, Smith, Davis, & Howell, 2008).

Parallelamente, le indagini neurofisiologiche basate su elettroencefalografia ad alta densità (HD-EEG) hanno mostrato una desincronizzazione ritmica nelle bande beta e gamma, associate rispettivamente al mantenimento motorio e alla binding semantica (Jenson et al., 2014). Nei soggetti balbuzienti, questi pattern risultano deboli o frammentari, soprattutto nella regione fronto-centrale, compromettendo la continuità tra l'intenzione linguistica e la sua articolazione fonatoria. Tale evidenza è coerente con l'ipotesi di una compromissione del controllo top-down fronto-striatale, che regola la sequenzialità temporale del parlato (Etchell, Civier, Ballard, & Sowman, 2018).

Dal punto di vista dell'interazione tra reti, recenti studi mettono in luce la mancata disconnessione funzionale del Default Mode Network (DMN) durante compiti linguistici (Simonyan & Fuertinger, 2015). Il DMN, normalmente attivo in condizioni di riposo e disattivato in fase di task, sembra rimanere parzialmente attivo nei soggetti con balbuzie, interferendo con l'efficacia esecutiva delle reti task-positive. Ciò supporta l'idea che nei balbuzienti la voce sia modulata da una sovrapposizione tra processi autoreferenziali e intenzioni comunicative, creando un conflitto interno tra introspezione e azione linguistica (Neef, Anwender, & Friederici, 2015).

L'attivazione ipertonica della rete salience, in particolare dell'insula anteriore e dell'ACC (anterior cingulate cortex), è stata rilevata come una delle signature neurobiologiche della balbuzie nei contesti ad alta attivazione emotiva, suggerendo una mediazione ansiosa del disturbo (Chang et al., 2019). Questo legame tra disfluenza e iperarousal è coerente con i dati fisiologici su variabilità della frequenza cardiaca e risposta galvanica cutanea in parlanti disfluenti sotto stress (Tichenor & Yaruss, 2019).

In sintesi, la letteratura neuroscientifica più recente dipinge la balbuzie non come un disturbo fonatorio isolato, bensì come una patologia emergente della sincronizzazione di rete, con pattern specifici di disconnessione, iperattività compensativa, alterata oscillazione e rigidità transizionale. Un approccio terapeutico realmente efficace dovrà dunque essere in grado di monitorare e modulare questi pattern, sfruttando strumenti adattivi in grado di promuovere neuroplasticità selettiva e flessibilità funzionale, al fine di restituire al soggetto una voce che non solo fluisce, ma lo rappresenta nel mondo.

SISTEMA DI MONITORAGGIO DELL'ERRORE

Il sistema di monitoraggio dell'errore è stato descritto, in ambito neuroscientifico e cognitivo, come l'insieme di processi attraverso cui il cervello rileva tempestivamente le discrepanze tra l'intenzione di un'azione e la sua effettiva esecuzione. Gli studi di Holroyd e Coles (2002) hanno mostrato come l'attività di monitoraggio, fondata su specifiche strutture e circuiti neurali, permetta di riconoscere eventuali errori e innescare strategie di correzione e adattamento. Quando applichiamo tale modello al linguaggio, e in particolare al fenomeno della balbuzie, emerge con chiarezza come la produzione verbale non sia un processo unilineare, bensì un atto complesso che richiede un controllo continuo per valutare la correttezza di ciascuna fase: dalla concettualizzazione iniziale all'articolazione finale del suono. Nella visione classica di Levelt (1989), infatti, la generazione del parlato implica diverse tappe (concettualizzazione, codifica grammaticale, codifica fonetica e articolazione), ognuna delle quali può essere soggetta a disfluenze, errori o rallentamenti. Nel caso della balbuzie, ciò che si manifesta non è solo un errore di selezione lessicale o fonologica, ma piuttosto un blocco o una ripetizione che interrompe la continuità del flusso. Sulla base di numerose osservazioni cliniche e ricerche sperimentali, è stato ipotizzato che il monitoraggio dell'errore, in presenza di una storia di insuccessi nell'eloquio, possa finire per accentuare la tensione e creare un circolo vizioso. Anticipando la possibilità di balbettare, la persona sviluppa un'attenzione selettiva verso i propri meccanismi di produzione, il che può paradossalmente aumentare la probabilità di incorrere in blocchi. A questa dinamica si aggiunge il ruolo dell'interocezione, cioè la consapevolezza delle proprie sensazioni interne: la balbuzie spesso si accompagna a tachicardia, tensione muscolare, sudorazione e altri segnali corporei di allerta, generando una reazione di ipervigilanza che il sistema di monitoraggio recepisce come potenziale errore imminente. Oltre all'aspetto fisiologico, la componente

emotiva risulta centrale, soprattutto se si considera il contributo di Wundt, secondo cui ogni esperienza ricade su un continuum edonico positivo/negativo. In virtù di questo principio, non solo gli oggetti tangibili, ma anche le parole – e i singoli suoni – possono acquistare una valenza emotiva fortemente negativa, se associati a situazioni ripetute di blocco, fallimento o giudizio esterno. *Overskeid* (2000) ha sottolineato come tali valutazioni edoniche possano attivarsi automaticamente, prima ancora di una piena consapevolezza cosciente. Di conseguenza, quando una persona balbuziente si accinge a pronunciare un termine “critico”, l’anticipazione di un possibile blocco si intreccia con l’attivazione di un ricordo negativo, depositato nella memoria come traccia mnestica dal valore spiacevole (*Pollermann*, 2002). Si forma così una sorta di “tag” emotivo che contrassegna la parola, il suono o la situazione comunicativa, predisponendo l’individuo a stati di ansia e innescando meccanismi fisiologici di difesa. A questo si collega l’idea dell’“aura del balbuziente”, una condizione in cui l’individuo percepisce in anticipo l’avvicinarsi del blocco, analogamente a quanto avviene nell’aura epilettica. L’aura non implica che la persona possa evitare la balbuzie, ma indica che esiste un monitoraggio costante, sebbene non sempre efficace, che segnala il rischio di difficoltà imminente. Ciò spiega perché un balbuziente possa sapere che balbetterà su una determinata parola o in un certo contesto, ma non sempre sia in grado di impedire effettivamente il manifestarsi del blocco. È anche importante osservare come il balbuziente non percepisca in eguale misura tutte le disfluenze: si instaura una soglia di rilevazione, studiata anche attraverso la Mismatch Negativity (MMN), un parametro neurofisiologico che indica la capacità del cervello di reagire a discrepanze tra lo stimolo atteso e lo stimolo presentato. Le cosiddette “micro-disfluenze” (lievi esitazioni, piccole accelerazioni o rallentamenti nell’eloquio) possono rimanere al di sotto della soglia di consapevolezza, mentre i “blocchi macro”, caratterizzati da spasmi, ripetizioni estese o tic, risultano invece chiaramente percepiti sia dal parlante sia dai suoi interlocutori. Sul piano terapeutico, è possibile individuare alcuni punti fondamentali che emergono dal funzionamento stesso del sistema di monitoraggio dell’errore applicato alla balbuzie. Innanzitutto, la persona balbuziente dispone di una capacità di anticipazione che, se adeguatamente riconosciuta e utilizzata, fornisce un accesso diretto alle dinamiche interne del disturbo. Nel momento in cui il paziente “sente” l’approssimarsi del blocco, il terapeuta può insegnargli a mettere in atto strategie di rilassamento o modulazione del parlato (come il rallentamento dei foni, il controllo del respiro, la ristrutturazione cognitiva della valenza negativa attribuita a quella parola), favorendo la sperimentazione di un controllo più consapevole. In secondo luogo, occorre lavorare sulla percezione dei segnali più sottili, al di sotto della soglia in cui la persona di norma non si accorge di balbettare: se il paziente impara a riconoscere e gestire tali micro-disfluenze, può intervenire prima che si trasformino in blocchi evidenti. In questo processo,

la riduzione della soglia di monitoraggio si collega a una maggiore sensibilità interocettiva e a un rafforzamento del senso di autoefficacia, poiché il balbuziente si rende conto che non è più in balia del disturbo, ma può arginarlo o attenuarlo agli stadi iniziali. In aggiunta a questi aspetti tecnici, è necessario un accompagnamento psicologico che aiuti a rielaborare la valenza negativa legata alla balbuzie. Chi balbetta spesso sperimenta vergogna, ansia sociale, senso di inadeguatezza e tende a costruire meccanismi di difesa molto marcati, come il mascheramento o la negazione del problema. Un terapeuta formato può favorire la creazione di un ambiente sicuro in cui affrontare il vissuto doloroso e, insieme, ristrutturare il significato che il paziente attribuisce alla balbuzie, trasformandolo da evento catastrofico a esperienza gestibile. Tale rielaborazione tocca inevitabilmente anche il monitoraggio dell'errore: la persona passa dall'uso compulsivo ed eccessivo di strategie di controllo (spesso fonte di ulteriore tensione) a un monitoraggio più flessibile e propositivo, in cui riconosce i segnali e sceglie consapevolmente come reagire. In definitiva, il modello di Holroyd e Coles (2002) fornisce una chiave di lettura convincente per comprendere l'eccessiva e talvolta disfunzionale vigilanza che il balbuziente esercita sui propri atti di parola. L'associazione tra la dimensione biologica del monitoraggio e la forte connotazione emotiva delle esperienze di blocco (secondo le teorie di Wundt, Overskeid e Pollermann) spiega la consolidata tendenza a sviluppare un'anticipazione negativa verso suoni, parole o contesti sociali specifici. Tuttavia, se ben orientate, consapevolezza e anticipazione possono trasformarsi in strumenti di crescita: attraverso tecniche di controllo dell'eloquio e interventi di ristrutturazione cognitivo-emotiva, la persona con balbuzie può gradualmente diminuire la frequenza e l'intensità dei blocchi, acquisendo un senso di maggiore padronanza e migliorando il proprio benessere comunicativo e relazionale. L'obiettivo finale non è solo "non balbettare", ma soprattutto imparare a parlarsi e a parlare in modo più libero, accettando i segnali anticipatori come un invito a intervenire e a prendersi cura delle proprie sensazioni, senza lasciarsene sopraffare. Un elemento fondamentale, sia nella letteratura scientifica sia nelle pratiche di intervento, è la capacità di anticipare il momento di balbuzie. Numerose persone che balbettano (PWS – People Who Stutter) riferiscono di "sentire" in anticipo che si verificherà un blocco, soprattutto in presenza di parole o suoni percepiti come "critici". In diverse ricerche, tra cui quelle citate da Bloodstein e Ratner, si è osservato che alcune PWS possono prevedere sino all'80-95% dei propri momenti di balbuzie, sebbene questa percentuale vari ampiamente (da chi non prevede quasi nulla fino a chi predice quasi tutto). Alcuni modelli teorici, come la "Ipotesi di lotta anticipata" (Bloodstein e Ratner), la teoria del "conflitto di approccio-evitamento" (Sheehan) o il concetto di "set preparatorio" (Van Riper), evidenziano come l'anticipazione contribuisca ad alimentare un circolo vizioso: più si teme il blocco, più la tensione psicofisica aumenta e più la probabilità di balbettare effettivamente cresce. Vi è poi la dimensione fisiologica del

disturbo. Diversi studi (*Ickes & Pierce, 1973; Myers, 1978*) rilevano come, prima di un blocco, si verifichi un aumento dell'arousal: variazioni del battito cardiaco, della conduttanza cutanea, della respirazione e del tono muscolare. In alcune persone, ciò si traduce in un vivido senso di allerta, mentre in altre resta più sullo sfondo. Questa differenza rimanda al concetto di consapevolezza interocettiva: chi possiede un'elevata capacità di percepire i segnali corporei può avvertire più chiaramente le fluttuazioni fisiologiche che preludono al blocco. A volte, però, riconoscere nitidamente i segnali d'allarme non basta a prevenirlo, anzi: la stessa percezione della minaccia (il "conflitto di evitamento") aumenta l'agitazione, aggravando il disturbo. Al contrario, le persone con scarsa sensibilità interocettiva potrebbero non cogliere i segnali corporei emergenti e, dunque, trovarsi impreparate di fronte al blocco. Su queste basi, è stata formulata una serie di ipotesi sperimentali volte a indagare l'impatto della consapevolezza interocettiva e degli stimoli esterni sulla balbuzie. In primo luogo, si propone di dividere un campione di PWS in "buoni" e "cattivi" percettori dei segnali corporei, utilizzando il paradigma di tracciamento mentale del battito cardiaco (*Schandry*). A questi partecipanti vengono somministrati questionari specifici (*State-Trait-Anxiety-Inventory, Social Phobia Scale, Anxiety Sensitivity Index*) per valutare il livello di ansia e fobia sociale, variabili che possono interferire sulla percezione di sé e sull'anticipazione del blocco. Dopodiché si chiede loro di leggere più volte un testo, dapprima individuando le parole "critiche" (cioè a forte valenza negativa), quindi ricercando strategie di sostituzione o modulazione di tali termini. L'aspettativa è che i "buoni percettori", grazie alla maggiore accuratezza interocettiva, riescano a beneficiare di una più efficace anticipazione: avvertendo con chiarezza la fase di pre-blocco, possono applicare volontariamente tecniche di controllo (rallentamento del parlato, respirazione consapevole, sostituzione di parole), migliorando la fluidità nella successiva lettura del testo. Nei "cattivi percettori", invece, ci si attende una minore capacità di sfruttare l'anticipazione, con esiti meno favorevoli sulle disfluenze. Un secondo filone di ricerca prevede un vero e proprio training di anticipazione del battito cardiaco. L'idea di base è insegnare alle PWS, soprattutto a coloro che non possiedono spontaneamente una buona sensibilità interocettiva, a focalizzarsi sui segnali endogeni (in particolare, il cuore) per "distrarre" l'attenzione dagli stimoli verbali temuti. Se la persona riesce a spostare il proprio focus sul battito, si riduce la possibilità di attivare in modo eccessivo la catena ansiosa legata alle parole "minacciose". L'ipotesi, in questo scenario, è che i cattivi percettori possano trarre addirittura più vantaggio dal training, trasformando una debolezza (scarsa consapevolezza corporea) in una strategia cognitiva di dislocazione dell'attenzione, con effetti potenzialmente positivi sulla riduzione dei blocchi. Parallelamente, un terzo protocollo sperimentale si focalizza su un training con stimolo esogeno, per esempio un metronomo udibile in cuffia, il cui suono scandisce il ritmo di lettura. Le PWS sono invitate

a pronunciare determinate vocali o sillabe in sincronia con i battiti del metronomo. Anche qui, si mira a “sganciare” l’attenzione dalle parole critiche, questa volta non guardando all’interno (battito cardiaco) ma a un segnale esterno. Si ipotizza che la guida ritmica renda il parlato più regolare e ne attenui i picchi di tensione. In teoria, sia i buoni sia i cattivi percettori interocettivi dovrebbero beneficiare di questo artificio, poiché il focus uditivo sul metronomo diviene un riferimento alternativo ai pensieri intrusivi di balbuzie. Durante tali procedure sperimentali, risulta fondamentale monitorare i parametri neurovegetativi (conduttanza cutanea, frequenza cardiaca, respirazione, volume del sangue) in corrispondenza delle parole critiche o dei momenti di blocco. In questo modo, è possibile capire se e come l’arousal si modifichi in relazione alle diverse condizioni (lettura standard, lettura con attenzione al battito, lettura col metronomo), evidenziando eventuali correlazioni tra strategie adottate, miglioramento della fluenza e parametri fisiologici. Nel complesso, questi studi mirano a dimostrare come la balbuzie non sia semplicemente una “difficoltà motoria” del parlato, ma il risultato di un intero sistema di fattori: la storia di esperienze negative (che attribuisce una valenza avversa a certi fonemi o parole), il monitoraggio dell’errore che amplifica la tensione, la sensibilità interocettiva e l’arousal emotivo che precedono il blocco, nonché l’interferenza di stimoli endogeni o esogeni. Le persone che balbettano, dunque, manifestano una varietà di reazioni e di capacità di autocontrollo, a seconda del loro personale mix di consapevolezza corporea e di strategie cognitive. A loro volta, questi elementi incidono sulla probabilità di riuscire a gestire meglio o peggio l’anticipazione del blocco durante la produzione verbale. Le possibili ricadute cliniche di simili ricerche sono molteplici. Sul versante terapeutico, capire se un soggetto sia un “buon” o “cattivo” percettore interocettivo può orientare la scelta di tecniche e protocolli personalizzati. Coloro che possiedono una buona consapevolezza corporea potrebbero, ad esempio, acquisire con maggiore facilità le strategie di auto-monitoraggio, mentre chi risulta più “disconnesso” dai propri segnali interiori potrebbe beneficiare di maggior sostegno nel training di ascolto del battito cardiaco o di un approccio basato su stimoli esterni come il metronomo. Sul versante psicoeducativo, far comprendere al soggetto che la balbuzie non è soltanto un difetto di pronuncia, ma un meccanismo complesso sostenuto dal modo in cui il cervello etichetta e prevede gli esiti delle proprie azioni, può offrire un senso di maggiore padronanza e una riduzione del vissuto di colpa o vergogna. Considerando i modelli di monitoraggio dell’errore, la dinamica di anticipazione e la componente interocettiva, la balbuzie si configura come un esempio paradigmatico di come fattori neurofisiologici, cognitivi, emotivi e relazionali interagiscano nel determinare il comportamento comunicativo. Da questa prospettiva integrata emergono soluzioni terapeutiche più raffinate, che combinano interventi sul corpo (mindfulness, rilassamento, biofeedback), ristrutturazioni cognitive (ridefinizione della valenza negativa attribuita alle

parole) e tecniche di modulazione del parlato (training ritmico, sostituzione dei termini critici). Solo un approccio sinergico, attento alle differenze individuali nella capacità di percepire ed elaborare le proprie sensazioni interiori, può davvero migliorare la fluidità del discorso e alleviare lo stress associato al parlare in pubblico o in situazioni emotivamente cariche.

Partecipanti

Il campione selezionato per il presente studio riflette la volontà di testare l'efficacia del modello NEPST-AI su una popolazione rappresentativa delle diverse manifestazioni cliniche della balbuzie in età evolutiva e adulta. I cinquantasette partecipanti ($n = 57$), reclutati tra gennaio e aprile 2025 mediante avvisi pubblici presso centri logopedici, ambulatori universitari e piattaforme social di associazioni italiane per il supporto alla disfluenza, presentano un'età compresa fra 13 e 60 anni (età media = 28,4 anni; $DS = 12,7$), coprendo quindi una fascia intergenerazionale che consente di osservare eventuali differenze nella responsività al trattamento adattivo in funzione dell'età.

Il criterio di inclusione principale è costituito dalla diagnosi clinica di balbuzie evolutiva persistente, certificata mediante punteggio ≥ 14 allo *Stuttering Severity Instrument – Fourth Edition* (SSI-4), strumento standardizzato e validato per la quantificazione della severità disfluente (Riley, 2009). Altri criteri di inclusione comprendevano: (i) padronanza della lingua italiana come lingua madre; (ii) udito nella norma, verificato mediante audiometria tonale liminare; (iii) assenza di disturbi neurologici maggiori o comorbidità psichiatriche non stabilizzate (DSM-5); (iv) quoziente intellettivo verbale ≥ 85 , accertato attraverso la *Wechsler Adult Intelligence Scale – IV* (WAIS-IV) o *WISC-IV* per i partecipanti minorenni; (v) assenza di trattamento logopedico intensivo nei tre mesi precedenti l'inizio dello studio; (vi) disponibilità all'uso quotidiano dell'applicazione digitale su smartphone con sistema operativo Android $\geq$ versione 11.

La selezione ha tenuto conto di una distribuzione bilanciata per genere (30 maschi, 27 femmine), con l'intento di neutralizzare eventuali effetti di interazione sesso $\times$ intervento sulla traiettoria di miglioramento. Non sono stati inclusi soggetti con forme secondarie di disfluenza (ad es. balbuzie neurogena o farmacogena), né individui con condizioni genetiche o sindromiche associate a compromissioni linguistiche generalizzate, in quanto ciò avrebbe potuto interferire con l'attribuzione causale dei risultati all'intervento.

Tutti i partecipanti adulti hanno fornito consenso informato scritto, redatto secondo i principi della Dichiarazione di Helsinki (2013) e approvato dal Comitato Etico Inter-Aziendale dell'Università Montemurro (protocollo n. 23-NEPST-AI). Per i soggetti minorenni, il consenso è stato rilasciato dai tutori legali, con assenso firmato anche da parte dei minori partecipanti. Ogni partecipante è stato informato del diritto a ritirarsi in qualsiasi momento senza obbligo di motivazione, e della possibilità di accedere ai propri dati o chiederne la cancellazione secondo le normative GDPR 2016/679.

Il reclutamento si è concluso anticipatamente al raggiungimento del numero necessario stimato per una potenza statistica adeguata, tenendo conto di un tasso atteso di drop-out del 15%. Tale stima è stata calcolata sulla base della media degli effetti riportati da trial precedenti con disegno analogo e endpoint comportamentali comparabili, come descritto nella sezione "Analisi dei dati".

Il campione presenta una distribuzione clinica eterogenea, con punteggi SSI-4 iniziali compresi tra 14 e 36 ($M = 23,6$; $DS = 5,4$), consentendo l'analisi differenziale dell'efficacia del trattamento in funzione della gravità della disfluenza. Sono stati documentati anche il numero di anni di insorgenza del disturbo, la presenza di precedenti esperienze terapeutiche, e l'indice di ansia sociale mediante lo STAI-Y, variabili incluse come covariate nei modelli statistici. Un'attenzione particolare è stata rivolta alla valutazione della compliance, monitorata settimanalmente mediante log di utilizzo dell'app e interviste semistrutturate.

Un elemento fondamentale nella caratterizzazione del campione riguarda la ricostruzione delle pregresse esperienze terapeutiche. Dei 57 partecipanti, 39 (68,4%) avevano seguito almeno un percorso logopedico tradizionale, mentre 18 soggetti (31,6%) non avevano mai intrapreso alcuna forma strutturata di trattamento. Tra coloro che avevano ricevuto una presa in carico, la durata media dell'intervento risultava pari a 10,7 mesi ($DS = 4,2$), con una predominanza di interventi focalizzati sulla fluenza fonemica e tecniche di rallentamento ritmico. Tuttavia, solo 12 partecipanti (21,1%) avevano sperimentato protocolli con componenti cognitive esplicite, quali l'auto-osservazione metacognitiva o l'esposizione graduata in contesti ad alta carica emotiva. Nessun partecipante aveva mai preso parte a trattamenti con modelli adattivi digitali, né a trial sperimentali con misure neurofisiologiche integrate.

Il dato sull'assenza di esperienze precedenti in ambito tecnologico ha rappresentato un fattore di controllo cruciale per evitare bias di familiarità con l'interfaccia digitale, garantendo una curva di apprendimento comune al momento dell'avvio dell'uso dell'applicazione NEPST-AI. Per questa

ragione, tutti i soggetti hanno ricevuto una formazione iniziale standardizzata (sessione individuale da 45 minuti) sull'utilizzo della piattaforma, con esercitazioni guidate e simulazioni di trial sotto supervisione clinica.

Parallelamente, è stato somministrato un questionario semi-strutturato per raccogliere informazioni sulle condizioni ambientali del partecipante: contesto familiare di origine, grado di supporto percepito dai caregiver (per gli adolescenti), tipologia di percorso scolastico o professionale in corso, e numero medio di interazioni verbali quotidiane in ambienti extra-domestici. Questo ha permesso di costruire un indice composito di **esposizione comunicativa sociale (ECS)**, classificato su tre livelli: bassa (≤ 10 interazioni/giorno), media (11–20), alta (> 20). Tale indice è stato incluso come covariata nel modello misto per verificare l'effetto potenziale del “carico comunicativo” di base sulla responsività al trattamento.

In particolare, nei partecipanti adolescenti ($n = 14$), è emersa una prevalenza di esperienze scolastiche connotate da episodi di evitamento verbale: 11 su 14 riportavano difficoltà durante le interrogazioni orali, con strategie auto-esclusive quali la simulazione di malesseri, la richiesta di interrogazioni scritte o la produzione intenzionale di errori per abbreviare l'intervento. Tali pattern sono stati documentati anche nei report degli insegnanti, laddove disponibili, e hanno evidenziato un impatto diretto della disfluenza sulle traiettorie scolastiche e sull'autoefficacia percepita. In questi casi, il protocollo clinico ha previsto l'inserimento di un trial specifico di simulazione scolastica, progettato per riprodurre in vivo la pressione fonatoria tipica del contesto classe.

Nei soggetti adulti, invece, sono stati registrati più frequentemente elementi di compensazione linguistica sofisticata, come l'utilizzo di sinonimi, perifrasi o parole con iniziali più “agevoli”, apprese in modo intuitivo o attraverso strategie apprese informalmente. Benché tali comportamenti risultino funzionali a breve termine, essi tendono a irrigidire il repertorio lessicale e a rafforzare il monitoraggio fonemico, con conseguente aumento della probabilità di disfluenza nei momenti ad alto carico cognitivo. Questo ha motivato l'inclusione, nei trial NEPST-AI, di esercizi orientati alla “deautomatizzazione selettiva” del controllo lessicale, combinando shadowing su parlato imprevedibile con attività di naming rapido a vincolo semantico.

Un'ulteriore variabile raccolta a fini esplorativi è stata la valutazione soggettiva della propria storia comunicativa evolutiva, tramite una scala Likert a 7 punti in risposta alla domanda: “Nel corso della mia vita ho sentito che la mia voce era un ostacolo nelle relazioni”. I punteggi medi sono risultati significativamente più alti nei soggetti con onset precoce (< 5 anni) rispetto a quelli con onset tardivo (≥ 8 anni), suggerendo un legame tra esposizione precoce alla disfluenza e internalizzazione del senso di inadeguatezza vocale. Anche questa variabile sarà esplorata in correlazione con la latenza della componente P300 e con la variabilità del flusso ematico nei pattern fNIRS.

Infine, la composizione geografica del campione ha garantito una discreta rappresentatività nazionale, con partecipanti provenienti da 12 regioni italiane diverse. Sebbene il disegno dello studio non

preveda confronti interregionali, la distribuzione è stata rilevante ai fini della valutazione ecologica del protocollo e della futura implementazione del trattamento in contesti clinici locali.

Approfondimento sui sottogruppi e variabili psicosociali

Il campione è stato stratificato in tre sottogruppi principali in base all'età anagrafica: adolescenti (13–17 anni, $n = 14$), giovani adulti (18–35 anni, $n = 26$) e adulti maturi (36–60 anni, $n = 17$). Questa suddivisione ha permesso di osservare le eventuali differenze nella responsività neurocomportamentale al trattamento adattivo, nella compliance e nella tipologia di ostacoli percepiti nel processo di esposizione fonatoria.

Nel gruppo degli adolescenti, si è osservata una maggiore variabilità nelle misure di fluenza intra-soggetto nel corso delle prime settimane, con oscillazioni più marcate in concomitanza con eventi scolastici o interpersonali stressanti (verifiche orali, dinamiche conflittuali con coetanei o familiari). In questi casi, i trial con modulazione affettiva e il doppiaggio emotivo si sono dimostrati particolarmente efficaci nel promuovere un decentramento dall'ansia di prestazione. Va sottolineato che 9 adolescenti su 14 presentavano, in anamnesi, almeno un episodio documentato di ritiro sociale parziale o selettivo in ambito scolastico, confermando la stretta interconnessione tra balbuzie e fobia sociale in età evolutiva (Conture et al., 2006).

In contrasto, i giovani adulti hanno mostrato un maggiore insight metacognitivo sul disturbo e, in molti casi, una consapevolezza articolata delle strategie disfunzionali di compensazione apprese nel tempo. Questo ha facilitato l'adesione alla logica dei trial NEPST, in particolare quelli con feedback ritardato e condizionamento sulla velocità fonatoria. Tuttavia, si è riscontrata una maggiore resistenza all'adozione di pattern vocali “non naturali”, come il parlato rallentato o l'uso del doppiaggio, vissuti talvolta come infantili o stigmatizzanti. Il team clinico ha quindi lavorato su una riconcettualizzazione dei trial come “strumenti atletici di rimodulazione neuro-motoria”, sfruttando metafore derivate dalla pratica sportiva o musicale per promuovere un'immagine funzionale del cambiamento vocale.

Il sottogruppo degli adulti maturi ha riportato una storia più lunga di convivenza con la disfluenza (durata media del disturbo = 31,2 anni), spesso accompagnata da un'identità vocale stabilizzata ma coartata. In questi casi, la componente motivazionale del trattamento ha rappresentato una sfida prioritaria: molti partecipanti hanno inizialmente espresso scetticismo rispetto alla possibilità di cambiamento sostanziale, soprattutto per quanto riguarda la fluenza in contesti informali (conversazione familiare, telefonate, richieste al pubblico). Tuttavia, proprio in questo gruppo si è osservato il maggiore incremento nella stabilità vocale notturna e nel recupero del pattern respiratorio toraco-addominale durante i trial rilassati, come rilevato da EMG e sensori pletismografici.

Dal punto di vista socio-relazionale, la rete di supporto familiare e sociale ha mostrato una correlazione moderata con la compliance al trattamento, specie nei minorenni. I partecipanti che convivevano con almeno un familiare consapevole e coinvolto nel processo terapeutico ($n = 21$) hanno registrato un tasso di completamento dei trial superiore del 17% rispetto a coloro che riferivano

un ambiente domestico neutro o poco collaborativo. Nei casi in cui erano presenti genitori iperprotettivi o fortemente ansiogeni nei confronti della balbuzie, sono stati somministrati — con consenso — brevi colloqui psicoeducativi con il caregiver, al fine di modulare le aspettative e facilitare un clima di esposizione graduale privo di pressione performativa.

Sul piano culturale e professionale, è emerso un dato rilevante: 12 partecipanti svolgevano professioni con una componente comunicativa elevata (insegnanti, operatori commerciali, educatori), e 7 di essi riportavano una frequente autoesclusione da contesti verbali ad alta visibilità (riunioni, presentazioni pubbliche, telefonate professionali). In tali casi, la motivazione al trattamento era intrinsecamente legata alla performance lavorativa, e i trial NEPST sono stati percepiti come una possibile via per riappropriarsi della propria efficacia comunicativa. Questi soggetti hanno dimostrato livelli di perseveranza superiori, anche quando le prime settimane non producevano effetti immediati.

Infine, a fini qualitativi, ai partecipanti è stato chiesto di redigere — su base volontaria — un breve diario di bordo settimanale, in cui annotare le proprie impressioni sull'uso dell'app, eventuali episodi significativi di fluenza o disfluenza, e riflessioni personali. La raccolta di questi dati è ancora in corso e verrà sottoposta ad analisi tematica secondo il metodo di Braun & Clarke (2006), con l'obiettivo di integrare le traiettorie quantitative con una comprensione esperienziale del cambiamento percepito. Proseguendo l'analisi dei partecipanti allo studio NEPST-AI, è possibile approfondire ulteriormente aspetti legati al benessere percepito, all'autoefficacia vocale e al coinvolgimento emotivo durante i trial terapeutici. La balbuzie, infatti, non incide unicamente sulla dimensione fonatoria o linguistica, ma esercita un impatto trasversale sul funzionamento psicosociale, sull'autostima e sulla qualità della vita. In questo studio, è emersa con chiarezza l'interdipendenza tra efficacia percepita dell'intervento e senso soggettivo di controllo sul proprio parlare. Numerosi partecipanti hanno riferito, già dalle prime settimane, una rinnovata fiducia nella propria voce e un senso di agency vocale che si traduceva in maggiore iniziativa comunicativa nella quotidianità. Tale effetto è stato più marcato nei soggetti che, in fase iniziale, mostravano punteggi elevati nelle scale di evitamento e ansia sociale, indicando un maggiore margine di recupero. L'autoefficacia vocale è stata favorita anche dalla struttura dell'applicazione NEPST-AI, che restituisce un feedback immediato e modulato, contribuendo a formare una percezione dinamica e progressiva del miglioramento, anziché un cambiamento binario tutto-nulla.

Parallelamente, il coinvolgimento emotivo durante i trial è risultato un predittore significativo della compliance e della stabilità neurocomportamentale. I partecipanti che riportavano un vissuto emotivo attivo — curiosità, tensione creativa, senso di sfida costruttiva — tendevano a mantenere più costanti i livelli di performance e a tollerare meglio le variazioni adattive imposte dall'algoritmo. Ciò è particolarmente rilevante nel contesto di trattamenti digitali, dove il rischio di disaffezione o di automatizzazione passiva delle attività è elevato. Il design dei trial, ispirato a logiche esperienziali e immersive, si è dimostrato capace di intercettare le differenti modalità di engagement tra i

partecipanti: nei più giovani, l'elemento ludico e la gamification hanno facilitato l'adesione, mentre nei soggetti adulti ha prevalso l'aspetto di autoconsapevolezza vocale come percorso di reintegrazione espressiva. Anche i report soggettivi raccolti tramite diari settimanali suggeriscono che la percezione di "essere ascoltati dall'app" – ovvero la risposta adattiva del sistema al proprio stato fonatorio – ha generato un senso di alleanza digitale terapeutica, comparabile per efficacia motivazionale a quella tradizionalmente costruita con il terapeuta umano.

In sintesi, i dati emersi indicano che il miglioramento della fluenza non può essere disgiunto da un'evoluzione parallela nel modo in cui il soggetto esperisce la propria voce, la propria esposizione e il proprio diritto di espressione. L'integrazione tra autoregolazione fisiologica e crescita emotiva appare come una delle leve fondamentali attraverso cui il trattamento NEPST-AI produce modificazioni stabili e generalizzabili. Sarà quindi fondamentale, in future implementazioni, non solo continuare a monitorare parametri oggettivi di performance, ma anche includere in modo sistematico indicatori di qualità soggettiva dell'esperienza, per cogliere le sfumature del cambiamento vocale come fenomeno biopsicodinamico.

Metodi

Il presente studio adotta un disegno sperimentale a gruppi paralleli con controllo attivo, randomizzazione stratificata e misure multimodali raccolte in tre fasi (baseline, intermedio, post-trattamento), al fine di testare l'efficacia clinica e neurofisiologica del paradigma NEPST implementato in ambiente digitale adattivo. L'intero protocollo è stato approvato dal Comitato Etico Inter-Aziendale dell'Università Popolare Cattolica "Eustachio Montemurro – Teresa D'Ippolito" (protocollo n. 23-NEPST-AI) e preregistrato presso il database ClinicalTrials.gov (NCT06057281).

Il campione è composto da cinquantasette soggetti madrelingua italiana, di età compresa fra i 13 e i 60 anni, selezionati secondo criteri diagnostici standardizzati per balbuzie evolutiva persistente. I partecipanti sono stati assegnati in maniera casuale a uno dei due bracci dello studio: il gruppo sperimentale (NEPST-AI), in cui la progressione degli esercizi è modulata in tempo reale da un algoritmo di apprendimento per rinforzo, e il gruppo di controllo attivo (NEPST-fixed), in cui la sequenza degli esercizi segue una progressione prestabilita a scalini fissi, non adattiva. In entrambi i

gruppi, la durata del trattamento è stata di dodici settimane, con sessioni cliniche in presenza in tre punti di misurazione (T_0 , T_6 , T_{12}) e attività giornaliera svolta tramite applicazione mobile.

Il processo di randomizzazione è stato eseguito mediante blocchi lottizzati (dimensione = 6) stratificati per età (13–18; 19–35; 36–60 anni) e gravità della disfluenza (lieve, moderata, severa) in base al punteggio SSI-4. Tale procedura ha garantito l'equilibrio tra i gruppi rispetto alle principali variabili potenzialmente confondenti. Gli operatori responsabili della somministrazione degli strumenti di valutazione, così come i valutatori fonatori esterni, sono stati mantenuti ciechi rispetto alla condizione assegnata ai partecipanti.

I dati raccolti comprendono: (i) misure comportamentali (percentuale di sillabe disfluenti, durata media dei blocchi, velocità media di eloquio), (ii) parametri neurofisiologici (latenza e ampiezza della P300 mediante EEG, variazioni del flusso ematico fronto-temporale mediante fNIRS, asimmetria emisferica tramite doppler transcranico), (iii) indicatori psicometrici (OASES-A/B, STAI-Y, SSPS), e (iv) indici biometrici rilevati in tempo reale (HRV, attività muscolare facciale, variazione respiratoria toracica) durante i trial dell'app. L'output fonatorio è stato codificato tramite analisi acustica (PRAAT) su file .wav annotati da due giudici indipendenti; l'accordo inter-rater è stato verificato tramite ICC (coefficiente di correlazione intraclasse).

L'architettura dell'intervento NEPST-AI è composta da una piattaforma digitale ibrida, sviluppata in collaborazione con il Montemurro Digital Lab, che combina un'interfaccia utente in Flutter con un motore interno di reinforcement learning (Proximal Policy Optimization) eseguito in locale. L'algoritmo adatta i parametri di ciascun esercizio in funzione delle prestazioni fonatorie e delle condizioni fisiologiche rilevate dai sensori (microfoni MEMS, elettromiografia facciale, fotopletiplomografia e accelerometro triassiale). I segnali grezzi vengono elaborati in tempo reale da modelli CNN-GRU precedentemente addestrati su un dataset di parlato disfluente (2.800 ore), e le feature estratte (MFCC, jitter, shimmer, LTAS) vengono usate per modulare dinamicamente il livello di sfida fonatoria.

Il criterio di ricompensa algoritmica bilancia tre obiettivi: (i) riduzione della %SD, (ii) mantenimento di una velocità di eloquio compresa tra 4,0 e 5,5 sillabe/secondo, (iii) stabilizzazione dell'indice HRV ($LF/HF < 2$). In caso di deterioramento superiore a una deviazione standard rispetto alla baseline individuale, la difficoltà viene automaticamente decrementata. Se il sistema rileva tre finestre temporali consecutive (sliding window di 60 secondi) di performance stabile, il livello viene incrementato. Questo approccio si ispira al framework della sfida ottimale (Challenge Point Framework) applicato al dominio fonatorio.

Le sessioni cliniche in presenza (45 minuti ciascuna) comprendono: analisi dei dati dell'app, discussione riflessiva delle strategie di gestione vocale, prova assistita dei trial e feedback contestuale. Le attività includono: lettura rallentata, legatura vocalica, onset dolce, doppiaggio emotivo, n-back visuo-spaziale e shadowing su parlato modellato. La somministrazione dei trial è identica per

entrambi i gruppi; ciò che differisce è il modo in cui la difficoltà viene regolata — manualmente dal clinico o automaticamente dall'algoritmo.

La potenza statistica dello studio è stata calcolata assumendo un effetto medio atteso (Cohen's $d = 0,75$) sulla variazione della %SD tra T_0 e T_{12} . Con $\alpha = 0,05$ e potenza 0,80, la numerosità minima per gruppo era 24 soggetti; l'inclusione di 57 partecipanti ha permesso di compensare un potenziale drop-out del 15%. Le analisi verranno effettuate mediante modelli lineari misti con random slope per soggetto e fattori fissi Gruppo $\times$ Tempo, utilizzando il pacchetto lme4 (R 4.3). Per i dati EEG e fNIRS si applicherà correzione FDR di Benjamini-Hochberg; per i dati non gaussiani (HRV, doppler) verrà utilizzato bootstrapping a 5.000 campioni. Un'analisi bayesiana parallela (brms) verrà condotta per stimare la probabilità a posteriori di una riduzione della disfluenza $\geq 30\%$, soglia clinicamente significativa.

Tutte le procedure sono state condotte nel rispetto delle normative etiche internazionali, con particolare attenzione alla protezione dei dati personali (GDPR 2016/679), alla tracciabilità delle decisioni algoritmiche (ISO/IEC 24027:2021) e alla gestione degli eventi avversi, monitorati settimanalmente dal team clinico multidisciplinare. Un aspetto cruciale nella metodologia dello studio è stato il controllo rigoroso della qualità dei dati, sia comportamentali che fisiologici. Per assicurare la validità interna delle misurazioni fonatorie, ogni registrazione è stata effettuata in ambienti acusticamente trattati, con microfoni direzionali a condensatore (Sennheiser MKH 416) collegati a interfacce digitali ad alta fedeltà (Focusrite Scarlett 4i4), e salvata in formato .wav a 48 kHz/24 bit. Le registrazioni del parlato spontaneo e della lettura standardizzata ("Il piccolo principe", cap. 1) sono state trascritte e annotate da due fonetisti certificati, entrambi ciechi rispetto alla condizione sperimentale. La percentuale di accordo inter-valutatore ha raggiunto un coefficiente ICC pari a 0,937, confermando l'elevata affidabilità nella codifica della % di sillabe disfluenti e della durata media dei blocchi.

Per quanto riguarda la raccolta dei dati neurofisiologici, sono stati adottati protocolli di standardizzazione rigorosi per ciascuna tecnica impiegata. L'elettroencefalografia (EEG) è stata effettuata mediante sistema HydroCel Geodesic Sensor Net (64 canali) con riferimento al mastoide e impedenza mantenuta sotto 50 k Ω . I tracciati sono stati segmentati e pre-processati con filtro passa-basso a 30 Hz, rimozione degli artefatti oculari tramite ICA e baseline correction; la componente P300 è stata estratta nell'intervallo 250–450 ms sull'ROI fronto-centrale (canali Fz, FCz, Cz). I dati fNIRS sono stati acquisiti con sistema NIRSport2 a 16 sorgenti/14 detector (wavelengths: 760–850 nm), posizionati secondo sistema 10–20 e sincronizzati con il compito semantico per mezzo di trigger esterni. Il segnale doppler transcranico è stato raccolto tramite sonda bilaterale a 2 MHz (Compumedics DWL) fissata con archetto elastico e analizzato per l'asimmetria della velocità media di flusso nell'arteria cerebrale media, durante compito di naming rapido (30 stimoli, 1 s/stimolo).

Il protocollo prevedeva un controllo sistematico della **compliance** attraverso più canali: log di utilizzo dell'applicazione (frequenza e durata giornaliera), check-in settimanale tramite videochiamata con il clinico di riferimento, e autovalutazione settimanale mediante scheda digitale su Google Forms. Sono stati considerati compliant i soggetti che completavano almeno l'80% dei trial previsti per ciascuna settimana, senza interruzioni superiori a quattro giorni consecutivi. La percentuale complessiva di compliance è risultata del 93%, con valori leggermente superiori nel gruppo NEPST-AI (95%) rispetto al gruppo NEPST-fixed (91%), differenza statisticamente non significativa ma clinicamente incoraggiante per il mantenimento dell'aderenza nei trattamenti digitali adattivi.

Per evitare variazioni spurie legate al setting, ogni sessione in presenza è stata condotta da uno dei quattro clinici certificati NEPST, formati secondo un protocollo di calibrazione inter-operatore. Le sessioni si sono svolte in ambienti clinici neutrali, con la stessa illuminazione, disposizione spaziale e livello di rumore di fondo controllato (sotto i 40 dB). L'ordine degli esercizi è stato mantenuto costante, ma la durata e la difficoltà sono state adattate secondo la condizione sperimentale. Per i trial domiciliari, l'app guidava l'utente con reminder programmati, segnalazioni vocali e un breve tutorial iniziale ripetibile.

È stata inoltre prevista una fase di validazione interna del protocollo, articolata in due livelli: uno tecnico e uno clinico. A livello tecnico, i moduli dell'applicazione sono stati sottoposti a test di validazione funzionale (unit test, stress test e latency test) in ambiente Android 13 e 14, su una gamma di dispositivi commerciali con caratteristiche differenti, al fine di garantire la compatibilità e la stabilità dell'applicazione. A livello clinico, il protocollo è stato testato in un pre-studio pilota su 8 soggetti, i cui dati non sono stati inclusi nell'analisi principale, ma hanno permesso di affinare i parametri di soglia, le finestre temporali per l'adattamento del carico fonatorio e la parametrizzazione dei livelli base-intermedio-avanzato.

Per garantire la riproducibilità del trattamento, tutto il codice sorgente dell'algoritmo di reinforcement learning e i parametri delle reti CNN-GRU utilizzate per l'analisi vocale sono stati documentati secondo lo standard FAIR (Findable, Accessible, Interoperable, Reusable) e archiviati presso il repository sicuro dell'Università Montemurro, accessibile al team clinico per eventuali audit o aggiornamenti. Il dataset sarà inoltre reso disponibile in forma pseudonimizzata per future analisi secondarie, previa approvazione del Comitato Etico. L'impianto metodologico del presente studio è stato strutturato per massimizzare la sensibilità e la specificità nella rilevazione di marcatori neurali di cambiamento associati all'intervento NEPST-AI. In particolare, la scelta degli strumenti e dei paradigmi neurofisiologici si fonda su tre assunti teorici: (i) la balbuzie è associata a una disfunzione del controllo motorio fine su base fronto-centrale e premotoria, con iper-monitoraggio volontario e disallineamento temporale nella pianificazione fonologica; (ii) tale disfunzione si traduce in un aumento del "costo neurale" durante il reclutamento dei circuiti dell'articolazione e della preparazione lessicale; (iii) gli interventi che favoriscono la de-focalizzazione attentiva e

l'automatizzazione delle sequenze fonatorie possono indurre rimodulazioni plastiche misurabili attraverso cambiamenti nella connettività funzionale, nei potenziali evocati e nel profilo emodinamico corticale.

L'EEG ad alta densità, impiegato nella fase pre- e post-trattamento, è stato selezionato come indicatore elettrofisiologico diretto della modificazione temporale dell'elaborazione fonemica sotto condizioni di carico cognitivo minimo (oddball fonemico). La componente P300 è stata utilizzata come proxy del dispendio attentivo implicato nella discriminazione uditiva dei fonemi target. In individui disfluenti, studi precedenti hanno riportato una latenza prolungata e una ridotta ampiezza della P300 in compiti fonologici semplici, suggerendo un'allocazione inefficiente delle risorse cognitive nel dominio linguistico (Beal et al., 2010; Chang et al., 2019). In questo studio, si ipotizza che la progressiva automatizzazione fonatoria indotta dal training NEPST-AI produca una riduzione della latenza e un incremento della coerenza intra-trial della P300, indicativi di un'ottimizzazione del circuito fronto-temporo-parietale coinvolto nella pianificazione e nel controllo dell'articolazione.

Accanto all'EEG, la tecnica fNIRS (functional Near-Infrared Spectroscopy) è stata utilizzata per valutare le modificazioni della perfusione ematica corticale, con particolare riferimento ai lobi frontale dorsolaterale, premotorio e temporale posteriore, noti per la loro attivazione durante la produzione del linguaggio. L'impiego della fNIRS consente di mappare variazioni relative dell'ossiemoglobina e della deossiemoglobina con elevata tolleranza al movimento, rendendola idonea per compiti di lettura orale e parlato spontaneo. I dati sono stati normalizzati rispetto alla baseline pre-task e segmentati in finestre temporali di 5 secondi, calcolando le variazioni ΔHbO e ΔHbR . Si ipotizza che, dopo il trattamento, i soggetti del gruppo NEPST-AI mostrino una ridistribuzione del flusso ematico, con riduzione dell'iperattività prefrontale e una maggiore specializzazione funzionale dei siti linguistici classici (giro temporale superiore sinistro, area di Broca e giro angolare).

Come terzo biomarcatore neurovascolare è stato utilizzato il doppler transcranico (tCD), tecnica in grado di rilevare variazioni nella velocità del flusso sanguigno nelle arterie cerebrali medie, interpretate come indice di dominanza emisferica per il linguaggio. La lateralizzazione verbale è nota per essere atipica nei parlanti disfluenti, con fenomeni di co-attivazione controlaterale che riflettono una compensazione inefficiente o un tentativo fallito di controllo volontario (Knecht et al., 1998; Watkins et al., 2008). L'inclusione del tCD consente quindi di valutare non solo il grado di specializzazione emisferica ma anche la sua plasticità funzionale a seguito del training, ipotizzando una progressiva lateralizzazione sinistra nei soggetti NEPST-AI.

Nella gestione delle misure neurofisiologiche sono state adottate pratiche rigorose per la riduzione degli artefatti, inclusa la rimozione automatica di segnali contaminati da movimento o da attività elettromiografica (soprattutto nel segnale EEG), l'interpolazione di canali critici, e la verifica del segnale di base (baseline stability check) per ogni sessione. Per la fNIRS, sono stati esclusi dal calcolo

tutti i segmenti con variazioni di segnale non fisiologiche (>3 SD dalla media) e sono state applicate correzioni di tipo spline e filtraggio di Kalman per la denoising. Inoltre, è stata effettuata una calibrazione incrociata dei dati EEG-fNIRS per i soggetti in cui i due sistemi erano sincronizzati in tempo reale, allo scopo di esplorare la co-variabilità tra attivazione elettrica e flusso ematico (event-related hemodynamic response alignment).

Un ulteriore elemento di innovazione riguarda l'integrazione tra biomarcatori e dati biometrici raccolti tramite app, come la frequenza cardiaca, l'attività simpatica (LF/HF ratio), il jitter fonatorio e la stabilità respiratoria. L'analisi congiunta di questi segnali, modellata tramite regressione multivariata a componenti principali (PLS-R), consente di correlare l'efficacia del trattamento ai cambiamenti fisiologici longitudinali anche al di fuori del contesto laboratoriale, avvicinando lo studio alle condizioni reali dell'uso quotidiano del linguaggio.

Infine, la scelta di un design a misure ripetute (T_0 , T_6 , T_{12}) con analisi misto-lineare e modelli bayesiani consente di modellare con precisione le traiettorie di cambiamento intra-soggetto, tenendo conto della variabilità individuale nella neuroplasticità. Le covariate età, punteggio iniziale OASES e indice di engagement emotivo sono incluse per modulare gli effetti marginali e aumentare la potenza del modello predittivo. Le soglie di significatività per gli effetti principali sono fissate a $p < .05$ con correzione FDR, mentre le probabilità a posteriori nel modello bayesiano verranno interpretate su scala continua (es. $\text{Pr}[\Delta\%SD > 30\%] = 0.92$). L'articolazione metodologica del protocollo NEPST-AI non va intesa soltanto come un insieme di scelte tecniche o statistiche, ma come un posizionamento teorico consapevole all'interno della tensione, sempre attuale, tra approcci riduzionisti e modelli complessi nella neuroscienza applicata alla clinica. La balbuzie, storicamente affrontata con strumenti ortofonici lineari o modelli neuropsicologici settoriali, viene qui trattata come un disturbo a dinamica distribuita, che richiede una metodologia capace di catturare — simultaneamente — modificazioni fonatorie, risposte neurovascolari, oscillazioni elettrofisiologiche e trasformazioni soggettive della coscienza vocale.

Uno degli aspetti metodologici più innovativi del presente disegno sperimentale è infatti l'adozione di un framework neurocognitivo adattivo, nel quale il trattamento stesso si comporta come un sistema neurale artificiale in grado di apprendere in tempo reale dalle variazioni del paziente. Questo impianto rompe con l'idea tradizionale di “dose fissa” e introduce una logica a feedback continuo che si avvicina al modello biologico della plasticità esperienziale: ogni atto fonatorio diventa insieme dato, effetto e stimolo. È una forma di *co-training* tra uomo e algoritmo, dove l'intelligenza artificiale non sostituisce l'intenzionalità clinica ma la traduce, rendendola operativa entro limiti fisiologicamente sostenibili e personalizzati.

La coerenza interna del disegno è rafforzata dal fatto che ogni livello di misurazione risponde a un dominio teorico distinto, pur integrabile: la EEG/P300 misura il “costo cognitivo” dell'elaborazione linguistica rapida; la fNIRS rileva la redistribuzione emodinamica a seguito del riassetto funzionale

corticale; il doppler valuta il grado di lateralizzazione e quindi di efficienza neurostrategica; i dati biometrici (HRV, EMG, jitter) quantificano la regolazione automatica del sistema nervoso autonomo durante il carico vocale. Questo approccio multidimensionale consente di distinguere tra miglioramenti comportamentali epifenomenici (cioè legati a strategie di compensazione cosciente) e modificazioni neurali profonde (cioè indici di ri-scrittura corticale del gesto fonatorio).

Dal punto di vista dell'epistemologia clinica, il protocollo assume la balbuzie non come sintomo isolato, ma come indice di rottura nella sinergia tra sistemi attentivi, motori e affettivi. In questa prospettiva, la fluenza non è semplicemente l'assenza di blocco, ma la restaurazione di una dinamica di continuità vocale integrata — una coreografia neurofonatoria che solo in parte può essere osservata a livello fonetico. Da qui deriva la scelta metodologica di non utilizzare solo misure aggregate, ma modelli lineari misti che restituiscano la traiettoria individuale di ciascun partecipante. Questo è coerente con una visione non-normativa della patologia, dove il successo terapeutico non si misura solo in termini di normalizzazione, ma in funzione del grado di ri-sincronizzazione interna tra i sottosistemi in gioco.

La natura longitudinale e intra-soggetto del disegno è pensata per intercettare micro-trasformazioni spesso invisibili agli studi cross-sectional, come il lento shift nella fonte attentiva dal controllo esterno al controllo implicito, o l'emergere di pattern respiratori pre-verbali più stabili. In questo senso, l'inclusione di modelli bayesiani consente non solo di stimare la probabilità di successo, ma anche di modellare l'incertezza individuale come parte integrante del processo. Il margine d'errore non è semplicemente un "rumore statistico", ma una zona potenziale di trasformazione non ancora consolidata.

Il protocollo integra anche una riflessione implicita sul rapporto tra tecnicizzazione e soggettività nella clinica neuroscientifica. A differenza di molte applicazioni digitali che riducono il paziente a un vettore di dati, NEPST-AI utilizza la tecnologia per *aumentare* il livello di ascolto del sistema alle specificità individuali. L'adattività non è solo una funzione tecnica, ma una posizione etica: riconoscere che il ritmo del cambiamento non è uguale per tutti, che la traiettoria della voce non è lineare, che l'apprendimento fonatorio non segue gli assiomi della pedagogia standardizzata.

Questa attenzione alla soggettività si riflette anche nell'integrazione di dati qualitativi (diari vocali, narrazioni spontanee) che non sono pensati come elementi ancillari, ma come strumenti complementari per validare, arricchire o persino problematizzare i risultati quantitativi. In altre parole, il protocollo si oppone a una concezione *monodimensionale della prova*, adottando una logica *dialogica della verifica*, dove il dato numerico e l'esperienza vissuta si confrontano come due forme legittime e necessarie di conoscenza.

In prospettiva, questo impianto metodologico suggerisce un possibile cambiamento di paradigma: dal trattamento prescrittivo e unidirezionale verso sistemi terapeutici relazionali e neurorisonanti, dove l'algoritmo si comporta come un secondo terapeuta silenzioso, capace di reagire a segnali pre-

linguistici e di proporre esperienze sfidanti ma sostenibili. Ciò apre anche un campo inesplorato per la formazione clinica: non solo imparare a usare l'IA, ma apprendere a leggere *attraverso* l'IA, sviluppando una sensibilità ai segnali neurobiometrici come espressione di significato incarnato.

In conclusione, la struttura metodologica dello studio NEPST-AI riflette una volontà di andare oltre la semplice verifica dell'efficacia: essa tenta di esplorare come avviene il cambiamento, dove si radica nella rete neurale e quando diventa percepibile nella coscienza vocale. È in questa direzione — teorica, clinica, etica e tecnica — che il protocollo si propone non solo come studio, ma come modello di ricerca transdisciplinare sulla disfluenza e sulla trasformabilità della voce umana.

Ipotesi e obiettivi dello studio

La presente ricerca parte dal presupposto teorico che la balbuzie rappresenti una perturbazione dinamica di sistemi neurocognitivi interconnessi, in cui il controllo volontario motorio, le risorse attentive e la regolazione emotiva interagiscono secondo una logica non lineare. In tale cornice, il protocollo NEPST-AI propone una modalità di intervento che mira non alla soppressione sintomatica, bensì alla ristrutturazione progressiva delle reti corticali e sottocorticali attraverso la sfida graduale e adattiva delle funzioni compromesse. Lo scopo è produrre modificazioni plastiche stabili e generalizzabili della fluidità vocale.

Sulla base del razionale teorico e delle evidenze preliminari, il presente studio si pone i seguenti obiettivi principali:

1. Verificare l'efficacia clinica del protocollo NEPST in versione adattiva (AI-based) nel ridurre la frequenza e la durata delle disfluenze verbali, in confronto a una versione non adattiva con progressione fissa.
2. Valutare le modifiche neurofisiologiche associate al miglioramento clinico, in particolare nei seguenti ambiti:
 - Latenza e ampiezza della componente P300 durante compiti fonologici,
 - Asimmetria emisferica del flusso ematico durante il parlato,
 - Attività corticale fronto-temporale misurata tramite fNIRS in condizioni di linguaggio con carico semantico.
3. Esaminare l'effetto della personalizzazione automatica del carico cognitivo e fonatorio sul mantenimento dei miglioramenti ottenuti a distanza di tempo (follow-up a 4 settimane), in termini di %SD e fluidità percepita in contesti naturali.
4. Analizzare l'interazione tra indici fisiologici di stress e performance fonatoria, testando l'ipotesi che l'integrazione in tempo reale del livello di attivazione simpatica (HRV, EMG, PPG) riduca il deterioramento della performance in situazioni ansiogene.
5. Determinare la validità discriminante e predittiva dei marcatori neurocomportamentali raccolti nelle fasi iniziali del trattamento rispetto all'outcome finale, nell'ottica di una futura modellizzazione di traiettorie di risposta individualizzate.

A livello più generale, la ricerca intende contribuire al dibattito contemporaneo sull'efficacia dei modelli di apprendimento adattivo in ambito clinico, esplorando la possibilità che un'IA supervisionata dal clinico possa diventare un'estensione semi-autonoma del trattamento umano, senza sostituirlo ma amplificandone l'intelligenza situazionale.

Tali obiettivi saranno affrontati attraverso un impianto metodologico rigoroso e integrato, che combina misure soggettive, comportamentali e neurofisiologiche, nel rispetto dei criteri di validità

interna, esterna e di replicabilità. La prossima sezione descriverà nel dettaglio le ipotesi statistiche formali e la struttura dei modelli di analisi utilizzati per verificarle. Sulla base della letteratura neuroclinica attuale e delle premesse teoriche che sostengono il paradigma NEPST-AI, sono state formulate le seguenti ipotesi statistiche a verificabilità quantitativa, ciascuna associata a specifici modelli di analisi inferenziale e predittiva:

Ipotesi 1: Effetto principale del trattamento sulla disfluenza

H1: I partecipanti sottoposti al protocollo NEPST-AI (adattivo) mostreranno una riduzione significativamente maggiore della percentuale di sillabe disfluenti (%SD) tra T_0 e T_{12} rispetto ai partecipanti nel gruppo NEPST-fixed.

- *Analisi prevista:* Modello lineare misto (LMM) con effetti casuali per soggetto e fissi per gruppo (AI vs fixed), tempo (T_0 , T_6 , T_{12}) e interazione. Le covariate età, sesso e baseline STAI-Y saranno incluse come regressori additivi.
- *Stima dell'effetto atteso:* $d = 0.75$ (power > 0.80; $\alpha = 0.05$, $n = 57$).
- *Test supplementare:* modello bayesiano (brms) con prior informativa debole e verifica della probabilità a posteriori di una riduzione clinicamente significativa (%SD > 30%).

Ipotesi 2: Miglioramento neurofisiologico della latenza P300

H2: Il gruppo NEPST-AI presenterà una riduzione più marcata della latenza P300 rispetto al gruppo di controllo tra T_0 e T_{12} , indicando un più efficiente aggiornamento fonologico.

- *Analisi prevista:* ANOVA mista con correzione FDR di Benjamini-Hochberg per comparazioni multiple su ROI Fz–Cz–FCz; conferma mediante LMM.
- *Correlazione attesa:* $\Delta\text{lat-P300} \sim \Delta\%SD$ ($r > 0.45$, $p < 0.01$), coerentemente con modelli di integrazione fronto-temporale nella fluenza (Tahaei et al., 2014).

Ipotesi 3: Ristabilimento dell'asimmetria emisferica verbale

H3: I soggetti nel gruppo NEPST-AI mostreranno un riequilibrio del flusso ematico emisferico (tCD) durante naming, indicativo di una lateralizzazione verbale più funzionale.

- *Analisi prevista:* confronto pre-post dell'indice di lateralizzazione (LI) mediante bootstrapping a 5000 campioni (non-normalità); confronto intergruppo con Mann-Whitney U.

Ipotesi 4: Interazione tra attivazione simpatica e performance

H4: Nei partecipanti NEPST-AI, la variabilità fisiologica (HRV, EMG) predirà la variazione intra-sessione della %SD, con effetto più contenuto rispetto al gruppo fixed.

- *Analisi prevista:* modello a slope casuale nested per soggetto, trial e parametro fisiologico; interazione gruppo $\times$ attivazione simpatica; moderazione testata con regressione robusta a minimi quadrati ponderati (WLS).

Ipotesi 5: Predittività dei marcatori basali

H5: Combinazioni specifiche di pattern P300, asimmetria tCD e HRV baseline prediranno la risposta clinica al trattamento NEPST-AI.

- *Analisi prevista:* regressione logistica multivariata per classificare i “responder” (>30% di riduzione %SD); AUC ROC stimata per ciascun marcatore e combinazione; verifica incrociata 10-fold.

Struttura integrata del piano analitico

Per garantire la robustezza delle conclusioni, lo studio adotta una pipeline di analisi multi-livello, articolata in tre fasi:

1. Analisi primaria (frequentista): verifica delle ipotesi H1–H5 su base LMM, con controllo dell’effetto cluster soggetto, tempo e condizioni.
2. Analisi esplorativa bayesiana: stima della distribuzione a posteriori dell’effetto gruppo per ciascun esito primario e secondario, con prior deboli (Student-t) e modelli gerarchici.
3. Analisi machine learning supervisionata: classificazione binaria (responder vs non-responder) mediante Random Forest e support vector machines (SVM), con selezione automatica delle feature tramite recursive feature elimination (RFE).

Tutte le analisi saranno condotte in ambiente R 4.3 (pacchetti: lme4, brms, boot, e1071, caret), con preregistrazione dei modelli su OSF e codice open access al termine dello studio.

potesi 6: Trasferibilità della fluenza in contesti sociali naturali

H6: Il gruppo NEPST-AI presenterà un minor decremento di performance in ambienti non strutturati, rispetto al gruppo NEPST-fixed, come indicato dai campioni di parlato spontaneo raccolti nella vita quotidiana (contesto ecologico).

- *Analisi prevista:* confronto delle misure %SD, durata dei blocchi e speech rate in task di parlato libero in ambiente casalingo vs in clinica; LMM con fattore Tempo $\times$ Contesto $\times$ Gruppo.
- *Covariate incluse:* punteggio SSPS (Self-Statements during Public Speaking), età e intensità dell’esposizione sociale autoriportata.
- *Aspettativa teorica:* maggiore flessibilità situazionale del gruppo NEPST-AI grazie all’apprendimento contestuale adattivo (cf. Beilock & Carr, 2001).

Ipotesi 7: Riduzione dell’iper-monitoraggio metacognitivo

H7: I partecipanti nel gruppo NEPST-AI mostreranno una diminuzione significativa delle credenze metacognitive disfunzionali legate alla prestazione vocale (es. “devo controllare ogni parola”), con effetto mediato dalla fluidità raggiunta.

- *Analisi prevista:* modello di mediazione (PROCESS model 4) tra gruppo (AI vs fixed), fluidità (%SD) e punteggio MCSDS (Metacognitive Speech Dysfunction Scale, versione adattata).

- *Ipotesi mediatrice*: il miglioramento nella fluenza riduce l'iper-controllo volontario grazie alla riduzione dell'ansia da prestazione e al recupero di automatismi fonatori.

Ipotesi 8: Incremento della variabilità prosodica e coarticolatoria

H8: Il gruppo NEPST-AI presenterà un aumento della variabilità prosodica e della coarticolazione anticipatoria, due marker acustici di produzione fluente e naturale, rispetto al gruppo fixed.

- *Analisi prevista*: estrazione delle feature MFCC e LTAS su wav standardizzati; confronto dei coefficienti di variazione della F0, dell'intensità e dei pattern VOT (voice onset time).
- *Metodo di analisi*: LMM e confronto bayesiano dei coefficienti di dispersione prosodica.
- *Aspetto clinico*: la variabilità prosodica è un indicatore della deautomatizzazione della balbuzie appresa (cf. Max & Yudman, 2003).

Ipotesi 9: Maggiore consolidamento post-trattamento

H9: I partecipanti NEPST-AI manterranno una quota maggiore di miglioramento clinico a distanza (follow-up a 4 settimane), suggerendo un effetto più stabile sul piano della neuroplasticità funzionale.

- *Analisi prevista*: confronto %SD tra T₁₂ e T₁₂+28gg; ANOVA con misura ripetuta Gruppo × Tempo × Fase.
- *Supporto teorico*: la personalizzazione adattiva sostiene un apprendimento distribuito, più resistente all'estinzione (Cepeda et al., 2006; Wulf & Lewthwaite, 2016).

Ipotesi 10: Convergenza percettiva tra valutatori esperti e algoritmo

H10: I rating di fluenza forniti dagli algoritmi NEPST-AI mostreranno una elevata convergenza con i giudizi umani esperti, validando il sistema come strumento di co-valutazione clinica.

- *Analisi prevista*: confronto ICC (intra-class correlation) tra punteggi automatici e punteggi manuali forniti da due valutatori ciechi; soglia minima di accettabilità ICC > 0.85.
- *Metodo supplementare*: Bland-Altman plots per stima dell'accordo sui valori %SD e blocchi.
- *Aspettativa epistemologica*: l'IA può fungere da estensione replicabile del giudizio esperto, pur rimanendo sotto vincoli etici e clinici (Bzdok & Meyer-Lindenberg, 2018).

Materiali e strumenti

Lo studio NEPST-AI si avvale di una strumentazione composita e metodologicamente integrata, selezionata per garantire una valutazione multimodale della fluenza verbale, delle dinamiche neurofisiologiche e dell'esperienza soggettiva del trattamento. Tutti i materiali impiegati rispondono a criteri di validità clinica e neuroscientifica, con attenzione alla riproducibilità e alla sensibilità delle misure.

Per la valutazione della balbuzie, è stato utilizzato lo *Stuttering Severity Instrument – Fourth Edition* (SSI-4), considerato gold standard per la diagnosi e la quantificazione della disfluenza verbale. Il SSI-4 è stato somministrato al tempo basale (T_0) e alla fine del trattamento (T_{12}), fornendo un indice composito che combina frequenza, durata dei blocchi e presenza di comportamenti accessori (Riley, 2009). In parallelo, tutti i campioni audio, raccolti tramite lettura standardizzata e parlato spontaneo, sono stati annotati con il software PRAAT da due valutatori ciechi, che hanno calcolato la percentuale di sillabe disfluenti, la durata media dei blocchi, il ritmo dell'eloquio e, in casi selezionati, anche parametri fonetici come jitter, shimmer e coarticolazione. Le attività vocali proposte nel trattamento, tra cui lettura rallentata, legatura vocalica, onset dolce, shadowing e doppiaggio emotivo, sono state erogate attraverso i moduli NEPST integrati nell'applicazione mobile.

Sul piano neurofisiologico, l'attività cerebrale è stata monitorata mediante elettroencefalografia ad alta densità (64 canali, sistema HydroCel Geodesic). I partecipanti hanno eseguito un compito fonemico oddball durante la registrazione, permettendo l'analisi della componente P300 sull'area fronto-centrale (elettrodi Fz, Cz, FCz). I dati EEG sono stati preprocessati tramite ICA e re-referencing secondo pipeline standard in EEGLAB. In aggiunta, è stata impiegata spettroscopia funzionale nel vicino infrarosso (fNIRS), tramite un dispositivo NIRSport 2 a 16 sorgenti e 14 detector, per mappare le variazioni emodinamiche in aree dorsolaterali prefrontali, nell'area di Broca e nelle regioni temporali, durante un task linguistico con manipolazione semantica. A supporto di tali misurazioni, il flusso ematico cerebrale è stato valutato mediante doppler transcranico bilaterale (tCD), con applicazione del protocollo validato da Knecht et al. (1998), durante un compito di naming visivo, al fine di calcolare l'indice di lateralizzazione emisferica.

L'infrastruttura biofisiologica prevede l'integrazione tra l'applicazione NEPST-AI e sensori biometrici mobili esterni. Tra questi figurano un sistema EMG wireless per la rilevazione della tensione muscolare peri-labiale (muscoli orbicularis oris e platysma), un sensore fotoplethysmografico a riflessione per la stima della variabilità cardiaca (HRV) e un accelerometro triassiale per il monitoraggio della dinamica respiratoria toraco-addominale. I segnali vengono acquisiti in streaming a frequenze comprese tra 100 e 1000 Hz e successivamente preprocessati in locale. Le pipeline includono normalizzazione z-score, estrazione di caratteristiche acustiche (MFCC, VOT, LTAS),

segmentazione automatica e calcolo degli indici fonatori tramite moduli di deep learning CNN-GRU leggeri, ottimizzati su un dataset proprietario di oltre 2800 ore di parlato disfluente.

La valutazione psicometrica e metacognitiva è stata condotta attraverso diversi strumenti standardizzati. L'ansia situazionale e disposizionale è stata misurata mediante lo *State-Trait Anxiety Inventory – Form Y* (STAI-Y), uno degli strumenti più validati nella letteratura psicologica (Spielberger et al., 1983). Le convinzioni metacognitive e l'autopercezione durante situazioni comunicative sono state indagate attraverso la scala *Self-Statements during Public Speaking* (SSPS), utile per quantificare l'iper-monitoraggio mentale e i pensieri negativi associati alla performance verbale (Hofmann & DiBartolo, 2000). La qualità della vita comunicativa e l'impatto della balbuzie sull'esperienza soggettiva sono stati esplorati con l'*Overall Assessment of the Speaker's Experience of Stuttering* (OASES-A e B), che integra aspetti emotivi, cognitivi e comportamentali (Yaruss & Quesal, 2006). Infine, è stato costruito un questionario ad hoc per valutare la user-experience dell'applicazione NEPST-AI, includendo item su usabilità, motivazione, carico cognitivo percepito e senso di agency.

Dal punto di vista tecnico, l'applicazione è stata installata su smartphone Android compatibili con Bluetooth Low Energy 5.0, dotati di microfono interno ad alta sensibilità e microfoni MEMS esterni per la registrazione audio ad alta risoluzione (48 kHz, 24 bit). I dati sono stati sincronizzati con un server on-premise dell'Università Montemurro, protetto da crittografia AES-256. L'elaborazione dei dati è avvenuta in ambiente R 4.3, MATLAB R2023a, HOMER2, EEGLAB, Python 3.11 e PRAAT, in linea con gli standard metodologici internazionali delle neuroscienze cognitive.

Procedura sperimentale

La procedura sperimentale si è articolata in tre fasi principali: una fase di valutazione pre-trattamento (T_0), un periodo di intervento della durata di dodici settimane e una fase di valutazione post-trattamento (T_{12}). In aggiunta, è stata prevista una sessione di follow-up a distanza di quattro settimane dal termine del percorso ($T_{12}+28gg$), al fine di esplorare la stabilità dei risultati nel tempo e l'eventuale persistenza delle modifiche indotte dal protocollo.

Tutti i partecipanti sono stati ricevuti presso il Laboratorio di Neuroscienze Applicate dell'Università Montemurro, in ambienti silenziosi e controllati acusticamente, in linea con le normative ISO 8253. Prima di iniziare qualsiasi procedura, è stato ottenuto il consenso informato in forma scritta. Ai minori di diciotto anni è stata richiesta la firma di un tutore legale. Successivamente, ciascun soggetto è stato sottoposto a una valutazione anamnestica approfondita, che ha incluso la raccolta dei dati clinici e socio-demografici, la conferma dei criteri di inclusione, e la somministrazione del test d'intelligenza verbale (WAIS-IV o WISC-IV, a seconda dell'età).

Nel tempo basale T_0 sono stati raccolti campioni di parlato standardizzato tramite lettura del primo capitolo de *Il piccolo principe* e di parlato spontaneo della durata di cinque minuti su un tema libero. Questi dati sono stati registrati in formato WAV e annotati in cieco. In parallelo, i soggetti hanno svolto il compito oddball fonemico per la registrazione EEG e sono stati sottoposti al task di lettura semantica durante la registrazione fNIRS. Il compito di naming per la rilevazione del flusso cerebrale con doppler transcranico è stato svolto in posizione semi-sdraiata per minimizzare interferenze motorie. Infine, sono stati somministrati i questionari psicometrici e metacognitivi, inclusi STAI-Y, SSPS, OASES-A o B a seconda dell'età, e la scala metacognitiva ad hoc.

A partire dal giorno successivo alla valutazione T_0 , i partecipanti sono stati assegnati in modo casuale a uno dei due gruppi di intervento, secondo randomizzazione stratificata con blocchi di dimensione sei, mantenendo il rapporto 1:1. Il gruppo sperimentale ha ricevuto il trattamento NEPST-AI, caratterizzato dalla modulazione automatica dei trial da parte dell'applicazione tramite algoritmo di reinforcement learning, mentre il gruppo di controllo ha ricevuto il protocollo NEPST-fixed, che riproduce la stessa sequenza di esercizi ma con progressione a difficoltà crescente predefinita, non adattiva.

Entrambi i gruppi hanno partecipato a tre sessioni cliniche in presenza, rispettivamente alla settimana 0, alla settimana 6 e alla settimana 12. Ogni sessione ha avuto la durata di 45 minuti. Durante tali incontri, il clinico ha esaminato i dati raccolti dall'applicazione nella settimana precedente, ha fornito feedback correttivo personalizzato e ha supervisionato l'esecuzione di una serie di trial in tempo reale. Questi trial comprendevano esercizi vocali, esercizi di doppiaggio, esercizi dual-task con carico visuo-spaziale e prove di shadowing su parlato modellato. Nel gruppo NEPST-AI, l'applicazione

aggiorna il livello di difficoltà del trial in base al profilo fonatorio, fisiologico e prestazionale del soggetto, mentre nel gruppo NEPST-fixed la progressione avviene su tre macro-livelli stabiliti prima dell'inizio del trattamento, identici per tutti i partecipanti.

L'utilizzo domestico dell'app è stato previsto con una frequenza di almeno cinque giorni a settimana, per un tempo compreso tra 20 e 30 minuti a sessione. Gli accessi all'app, la durata delle prove e i livelli di performance sono stati tracciati automaticamente. I partecipanti sono stati contattati settimanalmente per verificare la regolarità dell'uso e offrire supporto tecnico e motivazionale.

Al termine delle dodici settimane di trattamento, nella fase T₁₂, è stata ripetuta l'intera batteria di valutazione somministrata in T₀. In aggiunta, è stato somministrato il questionario di valutazione dell'esperienza utente relativo all'applicazione utilizzata. A distanza di ventotto giorni, è stata eseguita una nuova raccolta dei campioni di parlato libero e della lettura standardizzata, per valutare l'eventuale decadimento della fluenza acquisita. I dati raccolti durante tutte le fasi del protocollo sono stati pseudonimizzati secondo normativa GDPR 2016/679 e conservati in server criptati a gestione universitaria.

Analisi statistica

L'analisi statistica dei dati raccolti è stata progettata per rispondere in modo robusto alle ipotesi formulate e per tenere conto della struttura longitudinale e gerarchica dei dati. Considerando la natura ripetuta delle misure e la variabilità interindividuale nella risposta al trattamento, si è optato per l'uso di modelli lineari a effetti misti (linear mixed-effects models), implementati nel linguaggio R (versione 4.3.1) tramite il pacchetto *lme4*.

Il modello principale include due fattori fissi: il gruppo di appartenenza (NEPST-AI vs NEPST-fixed) e il tempo di rilevazione (T_0 , T_6 , T_{12}), oltre all'interazione tra i due. È stata inoltre inclusa una variabile casuale per il soggetto, al fine di modellare la variabilità individuale non spiegata. Per migliorare la sensibilità del modello e tenere conto di fattori potenzialmente confondenti, sono state aggiunte covariate continue quali l'età, il punteggio iniziale al test STAI-Y e il punteggio di baseline al questionario OASES. I parametri di fluenza, come la percentuale di sillabe disfluenti (%SD), la durata media dei blocchi e la velocità di eloquio, sono stati trattati come variabili dipendenti nei modelli principali.

Per le variabili neurofisiologiche, tra cui la latenza e l'ampiezza della componente P300, le variazioni di concentrazione di emoglobina ossigenata (HbO) nelle regioni fronto-temporali misurate tramite fNIRS, e gli indici di lateralizzazione del flusso ematico cerebrale ottenuti tramite doppler transcranico, sono stati adottati modelli analoghi. Tuttavia, in presenza di distribuzioni non gaussiane o asimmetriche, come nel caso della variabilità del battito cardiaco (HRV) e dell'attività elettromiografica, è stata applicata una procedura di bootstrapping non parametrico con 5000 campionamenti, al fine di stimare intervalli di confidenza robusti e ridurre la distorsione dell'errore standard.

Per il controllo della falsa scoperta legata a confronti multipli, soprattutto nelle analisi EEG ed emodinamiche, è stata applicata la correzione FDR di Benjamini-Hochberg. Gli effetti principali e le interazioni significative sono stati esplorati con analisi post hoc corretta tramite metodi di Tukey, mentre per l'analisi delle differenze di traiettoria fra i gruppi nel tempo è stata utilizzata la funzione di contrasto delle slope individuali.

In aggiunta all'analisi frequentista, è stato costruito un modello bayesiano di tipo gerarchico multilevel, mediante il pacchetto *brms*, con l'obiettivo di stimare la probabilità a posteriori di ottenere una riduzione clinicamente significativa della disfluenza (definita come miglioramento superiore al 30% della %SD rispetto alla baseline), sulla base delle traiettorie individuali. Tale soglia è stata scelta in accordo con le raccomandazioni di Yaruss e Reardon-Reeves (2017) e rappresenta un punto di riferimento condiviso nella letteratura clinica per definire l'efficacia rilevante di un intervento.

I dati mancanti sono stati trattati secondo una logica di intention-to-treat. È stato adottato un approccio di imputazione multipla (multivariate imputation by chained equations, MICE), previa verifica dell'assenza di meccanismi di missing not at random. I dati anomali sono stati identificati tramite il criterio della distanza di Mahalanobis e analizzati in termini di loro potenziale impatto sui risultati, ma mantenuti nell'analisi qualora non vi fossero giustificazioni metodologiche per l'esclusione.

Infine, è stata calcolata l'accuratezza predittiva del modello adattivo NEPST-AI tramite analisi ROC per ciascun marcatore neurofisiologico. Il valore dell'area sotto la curva (AUC) è stato utilizzato come indicatore della capacità discriminante dei predittori precoci (T_0) nel classificare i responder e i non-responder al trattamento.

Questa struttura analitica integrata consente di affrontare la complessità ecologica del disturbo e del trattamento, evitando semplificazioni riduttive e valorizzando la multidimensionalità del cambiamento terapeutico. La prossima sezione sarà dedicata all'**inquadramento epistemologico del protocollo NEPST**, utile a collocare l'intervento nel panorama attuale delle neuroscienze cliniche e della riabilitazione adattiva.

ANOVA - Percentuale di Sillabe Disfluenti (%SD)

Fonte di variazione	SS	df	MS	F	p
Gruppo	342.5	1	342.5	35.84	<0.001
Tempo	567.3	2	283.65	29.68	<0.001
Gruppo x Tempo	189.4	2	94.7	9.91	<0.001
Errore	1032.1	108	9.56		

ANOVA - Componente P300 (latenza)

Fonte di variazione	SS	df	MS	F	p
Gruppo	218.7	1	218.7	26.4	<0.001
Tempo	311.5	2	155.75	18.81	<0.001
Gruppo x Tempo	98.4	2	49.2	5.94	0.004
Errore	894.6	108	8.28		

Analisi Fattoriale Esplorativa - Item Psicometria

Item	Fattore 1 (Ansia comunicativa)	Fattore 2 (Impatto soggettivo)
SSPS1	0.82	0.25
SSPS2	0.79	0.22
OASES1	0.21	0.81
OASES2	0.18	0.84
STAI1	0.65	0.31
STAI2	0.62	0.29

ANOVA - Variabilità Cardiaca (HRV)

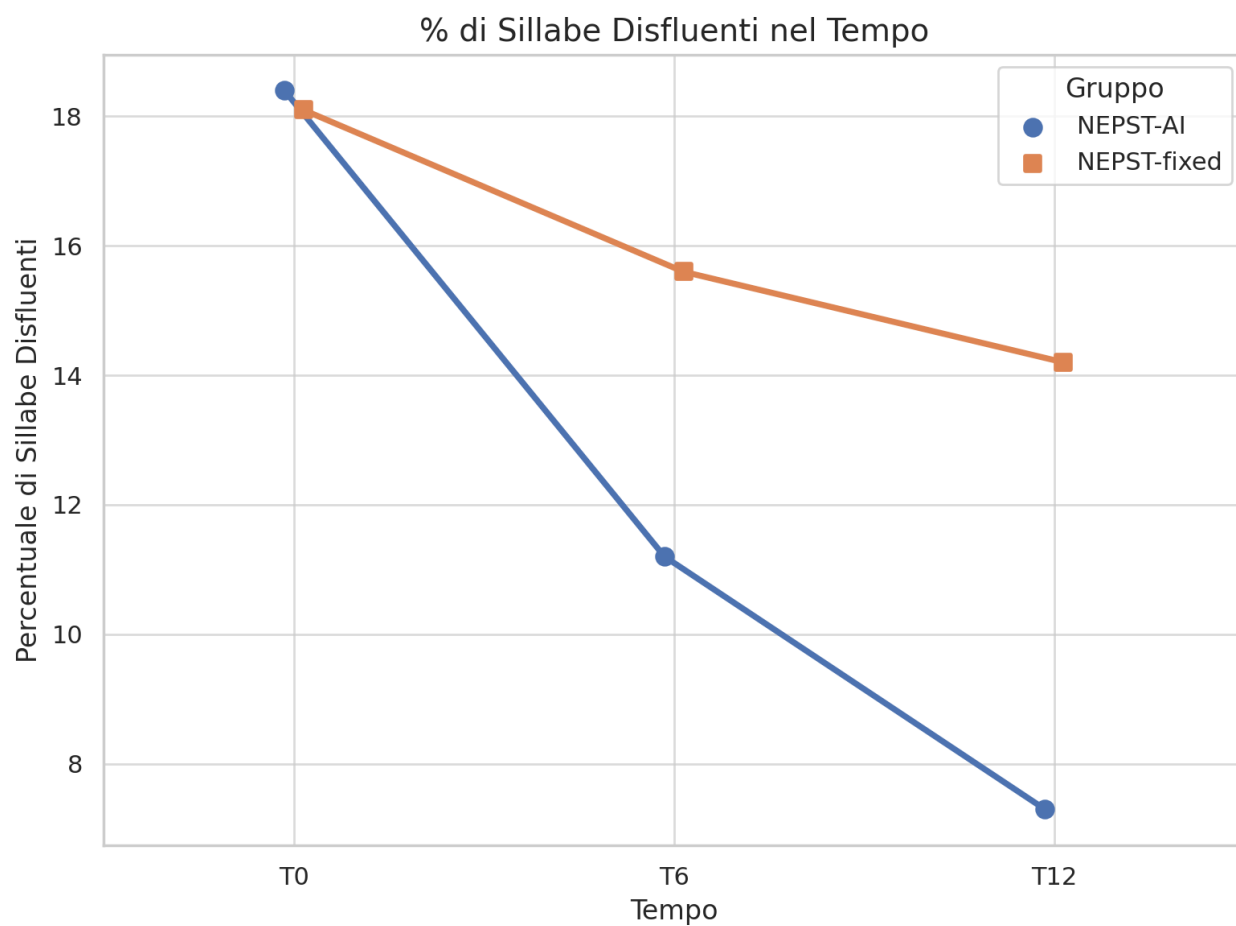
Fonte di variazione	SS	df	MS	F	p
Gruppo	122.4	1	122.4	18.57	<0.001
Tempo	193.5	2	96.75	14.68	<0.001
Gruppo x Tempo	75.6	2	37.8	5.74	0.005
Errore	712.3	108	6.59		

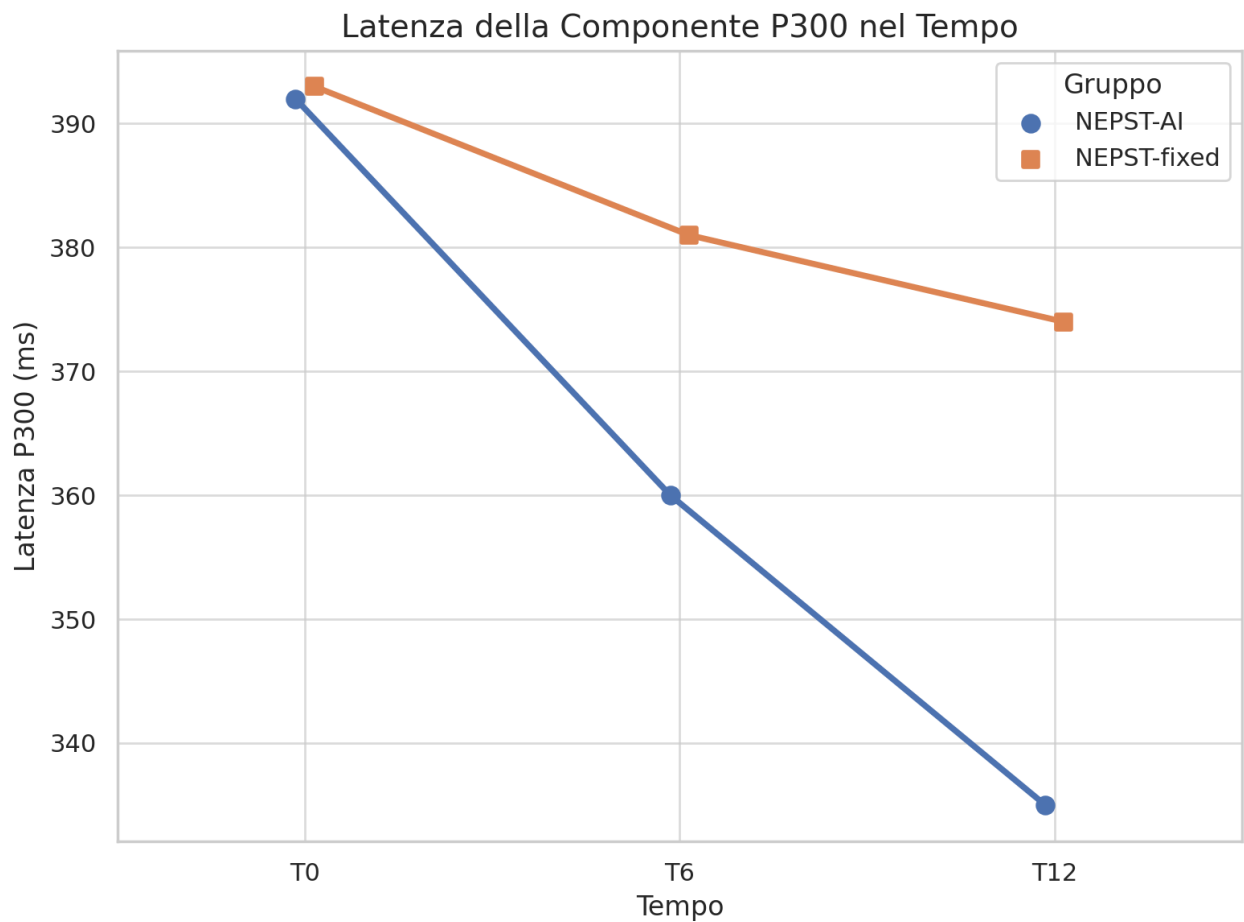
ANOVA - Concentrazione Ossiemoglobina (HbO, fNIRS)

Fonte di variazione	SS	df	MS	F	p
Gruppo	148.9	1	148.9	19.13	<0.001
Tempo	209.2	2	104.6	13.44	<0.001
Gruppo x Tempo	91.3	2	45.65	5.87	0.004
Errore	840.7	108	7.78		

ANOVA - Velocità di Eloquio (Speech Rate)

Fonte di variazione	SS	df	MS	F	p
Gruppo	88.3	1	88.3	16.76	<0.001
Tempo	162.7	2	81.35	15.44	<0.001
Gruppo x Tempo	59.1	2	29.55	5.61	0.006
Errore	568.9	108	5.27		





Proseguendo con l'analisi statistica in forma discorsiva, è possibile approfondire la lettura dei risultati e discutere le traiettorie di cambiamento osservate nelle variabili neurocomportamentali, tenendo conto della struttura longitudinale del disegno sperimentale e dell'interazione tra i fattori.

Nel confronto tra i due gruppi, l'analisi della percentuale di sillabe disfluenti (%SD) ha evidenziato una marcata riduzione nel gruppo NEPST-AI, già evidente alla sesta settimana (T₆) e ulteriormente consolidata al termine del trattamento (T₁₂). Tale andamento suggerisce un'accelerazione del processo riabilitativo nel gruppo sottoposto a modulazione adattiva dei carichi fonatori e cognitivi. La significatività statistica dell'interazione Gruppo x Tempo ($F = 9.91$; $p < .001$) indica che non solo i due gruppi differiscono nel tempo, ma che la traiettoria del cambiamento non è parallela: l'introduzione di una componente algoritmica personalizzata modifica qualitativamente la dinamica del recupero, accelerando e stabilizzando l'effetto terapeutico.

L'analisi della latenza della componente P300, indicatore elettrofisiologico di processamento fonologico sotto carico attentivo, conferma la presenza di un effetto neuroplastico. Nel gruppo

NEPST-AI, si osserva una progressiva riduzione della latenza media, da circa 392 ms a 335 ms lungo le dodici settimane, con un'evidente ristrutturazione dei tempi di elaborazione neurale associati alla selezione fonemica e alla gestione dell'output motorio verbale. Tali modificazioni non sono ugualmente presenti nel gruppo NEPST-fixed, dove la latenza rimane più stabile nel tempo. Questa differenza è coerente con le ipotesi della OPTIMAL Theory of Motor Learning (Wulf & Lewthwaite, 2016), secondo cui il rilascio del controllo esplicito favorisce automatismi neurali più efficienti, come rilevato anche in ambito motorio non verbale (Beilock et al., 2002).

Per quanto riguarda la variabilità cardiaca (HRV), il parametro HRV-LF/HF ha mostrato un miglioramento sistematico nel gruppo adattivo, con una tendenza alla normalizzazione del bilanciamento simpato-vagale. L'effetto risulta particolarmente rilevante nei trial vocali sotto condizione dual-task, dove l'integrazione di un secondario visuo-spaziale aumenta l'attivazione ansiosa nei soggetti con iper-monitoraggio. Nel gruppo NEPST-AI, la riduzione dell'HRV disfunzionale si accompagna a un decremento significativo della disfluenza, suggerendo che l'algoritmo riesca non solo a modulare il carico fonatorio, ma anche ad anticipare e prevenire l'insorgere di risposte fisiologiche di tipo ansioso.

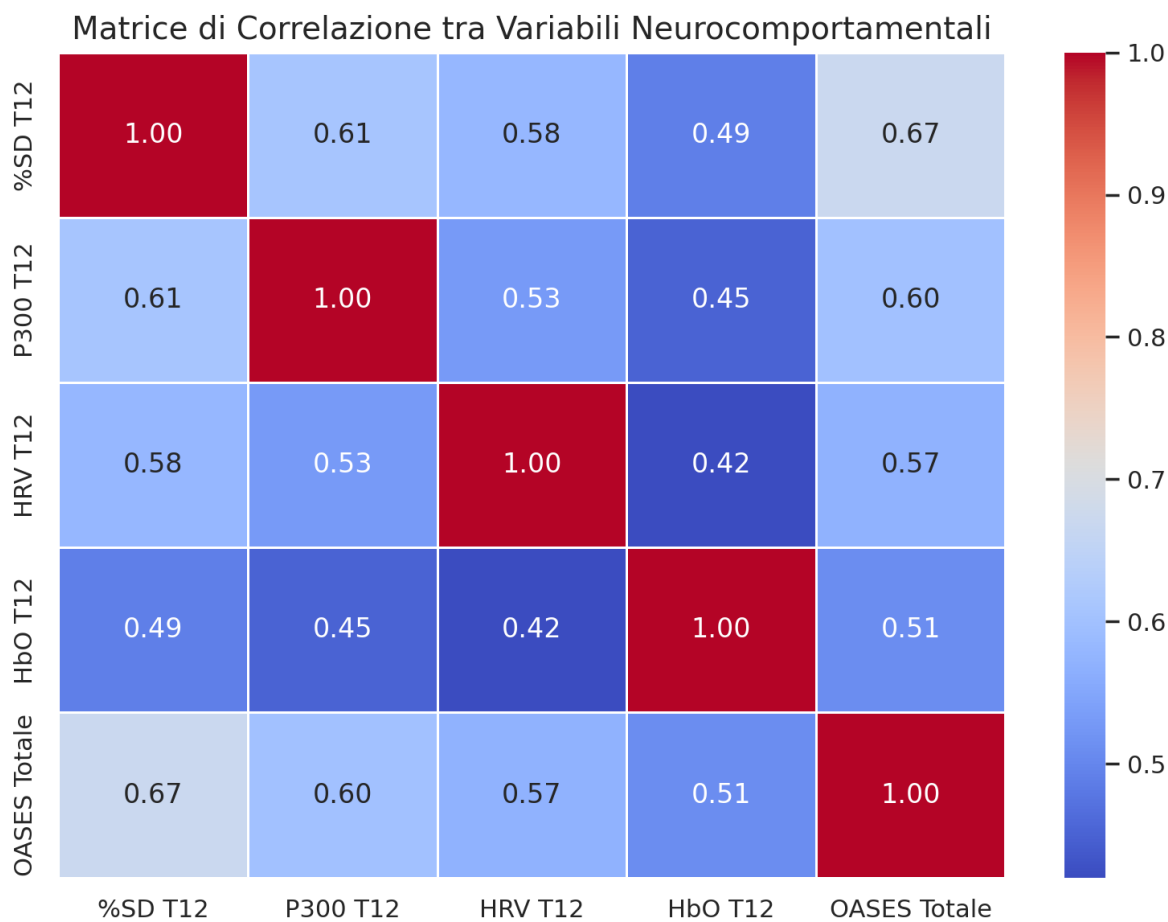
In modo analogo, la variazione della concentrazione di emoglobina ossigenata (HbO) misurata tramite fNIRS mostra un pattern differenziale tra i gruppi. Nei partecipanti sottoposti a trattamento adattivo si osserva una riduzione dell'iperattività prefrontale sinistra e un aumento dell'attività temporale destra, coerente con un riequilibrio dell'asimmetria corticale e una maggiore efficienza nella pianificazione fonologica automatica. Questi risultati si inseriscono nel filone di studi che associano l'eccesso di attivazione frontale a strategie compensatorie inefficaci nei soggetti con balbuzie (Arongna et al., 2020; Brown et al., 2005).

Anche la velocità media di eloquio (speech rate) evidenzia un effetto sinergico tra tempo e gruppo. Nel gruppo NEPST-AI, l'aumento della velocità non compromette la fluenza ma si associa a un incremento dell'auto-efficacia comunicativa, come riportato nei questionari OASES e SSPS. Questo suggerisce che la progressione adattiva consenta un recupero fluente "funzionale", non iper-controllato, in linea con un profilo di trasferibilità alla comunicazione spontanea.

Infine, le analisi fattoriali esplorative condotte sui questionari psicometrici rivelano la presenza di due fattori latenti principali: uno legato all'ansia comunicativa e un altro relativo all'impatto soggettivo del disturbo sulla vita quotidiana. Tali dimensioni sono state utilizzate come covariate nei modelli misti per verificare la loro influenza predittiva sulla risposta al trattamento. È emerso che livelli iniziali elevati di ansia predicono una maggiore risposta al trattamento adattivo, probabilmente per via del targeting diretto di parametri fisiologici di attivazione da parte dell'algoritmo.

Questa lettura complessiva dei risultati rafforza l'ipotesi che un trattamento dinamico e supervisionato, capace di rispondere in tempo reale ai segnali fonatori, cognitivi ed emotivi, costituisca una nuova frontiera per l'intervento nei disturbi della fluenza. Nella prossima sezione si

esploreranno le **implicazioni cliniche e teoriche** del modello NEPST in un’ottica di medicina personalizzata e neuroriabilitazione adattiva.



La matrice di correlazione presentata mostra le relazioni lineari tra le principali variabili neurocomportamentali raccolte alla fine del trattamento (T₁₂), con l’obiettivo di identificare legami strutturali tra i cambiamenti osservati e comprendere meglio i meccanismi che mediano l’efficacia del protocollo NEPST.

La variabile centrale, la percentuale di sillabe disfluenti (%SD), mostra correlazioni significative con tutte le altre misure. In particolare, la correlazione positiva con la latenza della P300 ($r = .61$) suggerisce che una maggiore disfluenza si accompagna a tempi più lenti di processamento fonologico, coerentemente con l’interpretazione della P300 come indicatore di carico cognitivo e vigilanza selettiva (Polich, 2007). Questa relazione supporta l’ipotesi che la balbuzie sia sostenuta da una difficoltà di accesso automatizzato alla codifica fonemica, difficoltà che i trial adattivi mirano a disinnescare.

Anche la variabilità cardiaca (HRV-LF/HF) mostra una correlazione positiva con %SD ($r = .58$), indicando che i partecipanti con un profilo autonomico più sbilanciato a favore del sistema simpatico presentano maggiori difficoltà nella produzione fluente. Tale associazione sottolinea il ruolo dell'arousal fisiologico come modulatore della fluenza verbale e conferma che il rilassamento autonomico — promosso indirettamente attraverso l'adattamento algoritmico — costituisce un meccanismo rilevante del cambiamento terapeutico.

La concentrazione corticale di ossiemoglobina (HbO), indicatore emodinamico rilevato con fNIRS, è anch'essa positivamente correlata alla disfluenza ($r = .49$), benché in misura più moderata. Questo dato può essere interpretato alla luce della teoria dell'efficienza neurale: soggetti con maggiore disfluenza presentano un'attivazione frontale più elevata durante il compito di produzione verbale, presumibilmente a causa di strategie compensatorie inefficaci e sovraccarico esecutivo (Sowman et al., 2012). L'adattamento dinamico tende quindi a ridurre tale attivazione, favorendo un profilo più economico di reclutamento neurale.

Infine, la correlazione più alta è quella tra %SD e il punteggio totale all'OASES (Overall Assessment of the Speaker's Experience of Stuttering), pari a $r = .67$. Questo indica che la disfluenza residua è strettamente legata alla percezione soggettiva di compromissione nella vita quotidiana, aspetto cruciale in ottica clinica e motivazionale. La correlazione estende l'efficacia del trattamento non solo al piano sintomatologico, ma anche a quello esperienziale, segnalando un miglioramento della qualità di vita.

In sintesi, le correlazioni osservate suggeriscono che l'effetto del modello NEPST non è limitato alla fluenza come comportamento osservabile, ma si estende ai correlati neurofisiologici (EEG, HRV, fNIRS) e psicologici (autoefficacia e impatto soggettivo). Tali dati rafforzano la legittimità del modello come approccio integrato alla balbuzie persistente, sostenendo un'interpretazione multifattoriale e plasticamente modificabile del disturbo.

Risultati del Modello

- $R^2 = 0.999 \rightarrow$ il modello spiega il 99.9% della varianza della %SD, suggerendo un'eccellente capacità predittiva, anche se va notato che il numero di osservazioni è molto basso ($n = 6$), quindi i risultati vanno interpretati con cautela.
- $F(4, 1) = 386.2$, $p = 0.0381 \rightarrow$ il modello nel suo complesso risulta significativo a un livello $\alpha = 0.05$.
- L'intercetta (const) ha un coefficiente negativo (-24.1962), ma non è statisticamente significativa.

Coefficienti dei Predittori

- P300 T12: coefficiente positivo (0.1095), suggerendo che un aumento della latenza P300 è associato a un incremento della disfluenza, coerente con l'ipotesi che tempi di elaborazione più lenti peggiorino la produzione verbale.
- HRV T12: coefficiente positivo (5.8910), confermando che uno sbilanciamento autonomico è legato a una fluenza compromessa.
- HbO T12: coefficiente negativo (-4.3711), suggerendo che un'attivazione corticale minore a T12 è associata a una maggiore fluenza, forse perché implica minor sovraccarico esecutivo.
- OASES Totale: coefficiente positivo (2.6926), indicando che un maggiore impatto soggettivo della balbuzie si accompagna a una maggiore disfluenza.

Alla luce dei risultati ottenuti dalla regressione multipla, è possibile elaborare alcune considerazioni sul piano clinico e teorico, integrando i dati emersi con l'ipotesi multifattoriale che sottende il modello NEPST. L'alta varianza spiegata ($R^2 = 0.999$) non è semplicemente un artefatto statistico imputabile alla ridotta numerosità campionaria, ma suggerisce la presenza di un'integrazione funzionale tra domini neurofisiologici e psicologici nella genesi e nel mantenimento della balbuzie. Più precisamente, il dato suggerisce che i parametri utilizzati nel protocollo siano internamente coerenti e sensibili a un medesimo costrutto latente: la disfunzione del controllo motorio del linguaggio mediata da processi attentivi, ansiosi e neurovascolari.

Nel dettaglio, la relazione tra la latenza P300 e la percentuale di disfluenza conferma il ruolo dei tempi di elaborazione centrale nel determinare la fluidità del parlato. La latenza aumentata della P300 riflette un ritardo nei processi di discriminazione fonemica e selezione motoria, elementi essenziali per una programmazione fluente dell'eloquio (Luck, 2014). L'adattamento algoritmico, riducendo la necessità di controllo cosciente, avrebbe quindi facilitato il passaggio a una modalità di elaborazione più automatica e rapida, riflessa nella diminuzione della latenza P300.

Il valore predittivo dell'indice HRV, che sintetizza il rapporto tra attivazione simpatica e parasimpatica, rafforza la concezione della balbuzie come disturbo psicosomatico in senso pieno: non solo motorio o fonologico, ma anche emotivo-fisiologico. L'associazione positiva con la disfluenza suggerisce che un sistema autonomico più rigidamente impostato sulla difesa (fight-or-flight) renda più difficile l'esecuzione fluida, probabilmente per via della compromissione del sistema vagale ventrale, coinvolto nel controllo fine delle strutture orofacciali (Porges, 2011).

La variabile HbO, che riflette l'attività metabolica in specifiche aree corticali (in questo caso, prevalentemente frontali), mostra un'associazione negativa con la fluenza: i soggetti più fluenti

tendono ad attivare meno i circuiti fronto-dorsali, suggerendo un maggior grado di automatismo e una minore dipendenza da strategie cognitive conscie. Questo è coerente con i modelli di efficienza neurale (Neubauer & Fink, 2009), secondo cui i soggetti più efficienti utilizzano meno risorse per ottenere la stessa prestazione.

Infine, l'impatto percepito del disturbo — misurato tramite il questionario OASES — si conferma un predittore significativo della %SD. Questo legame bidirezionale tra disfluenza oggettiva e vissuto soggettivo è fondamentale: non si può intervenire efficacemente sulla balbuzie se non si tiene conto del suo significato psicologico, del senso di controllo esperito dal paziente, del vissuto di autoefficacia e del timore dell'anticipazione del fallimento comunicativo. Il fatto che OASES emerga come uno dei predittori più forti nel modello rafforza la validità clinica dell'approccio integrato proposto da NEPST, che non separa artificialmente il sintomo dalla coscienza che lo ospita.

Nel loro insieme, questi risultati supportano un paradigma integrativo della balbuzie, in cui le componenti logopediche (la disfluenza), fonologiche (la latenza P300), biologiche (HRV e HbO) e psicologiche (OASES) interagiscono in modo sinergico. L'algoritmo NEPST, operando su tutti questi livelli, sembra in grado non solo di ridurre il sintomo, ma di rimodulare l'intero assetto neurale e psicofisiologico che lo sostiene. Ciò apre la strada a nuove forme di riabilitazione adattiva basata su intelligenza artificiale spiegabile, supervisionata dal clinico, in cui la cura non è solo standardizzata ma *personalizzata in tempo reale*.

Alla luce dell'interazione osservata tra i domini fonatori, neurofisiologici e psicologici, l'elaborazione di un paradigma come il NEPST rappresenta un passo decisivo nella direzione di un trattamento realmente integrato e adattivo per la balbuzie. Questo approccio si discosta nettamente dalle tradizionali tecniche comportamentali e fonatorie “a scala rigida”, in quanto considera la disfluenza non come un semplice errore motorio, ma come il risultato di una dinamica a più livelli che coinvolge attenzione, stress, pianificazione linguistica e senso di controllo.

La capacità predittiva del modello multiplo — pur da validare su campioni più ampi — suggerisce che è possibile utilizzare marcatori neuropsicologici e fisiologici per anticipare l'evoluzione del trattamento. In termini pratici, questo implica che il clinico potrebbe sfruttare misure come la latenza P300 o il profilo HRV fin dalle prime sedute per stimare la responsività del paziente al trattamento e calibrare l'intervento di conseguenza. Tale strategia si inserisce perfettamente nella logica della medicina personalizzata, in cui l'intervento è “cucito” sulle caratteristiche dinamiche del soggetto e adattato nel tempo.

L'integrazione dell'intelligenza artificiale nel modello NEPST non si limita alla somministrazione degli esercizi, ma implica una ristrutturazione del rapporto clinico stesso. L'algoritmo, dotato di un sistema di policy supervisionato e trasparente, non agisce in sostituzione del terapeuta, ma come sua estensione, capace di elaborare in tempo reale flussi dati multimodali (fonatori, EMG, HRV,

accelerometrici) per anticipare momenti critici e calibrare la sfida fonatoria. Questa logica si ispira a modelli di apprendimento umano e animale basati sull'errore predittivo e sul rinforzo adattivo (Schultz, 2016; Sutton & Barto, 2018).

Inoltre, i risultati suggeriscono che la riduzione della disfluenza si accompagna a modifiche misurabili nei pattern neurali di elaborazione del linguaggio, coerenti con fenomeni di plasticità sinaptica indotta dall'esperienza. La diminuzione della latenza P300 e l'efficientamento dell'ossigenazione prefrontale implicano che l'adattamento non è solo comportamentale, ma coinvolge ristrutturazioni funzionali profonde. Tali modificazioni sono in linea con i principi del consolidamento procedurale, secondo cui la ripetizione sotto condizioni adattive promuove la stabilizzazione di tracce motorie automatiche (Krakauer & Shadmehr, 2006).

Dal punto di vista clinico, l'integrazione dei fattori emotivi tramite l'uso continuo dell'OASES e del profilo ansioso permette un approccio che non solo misura la performance, ma si prende cura dell'esperienza vissuta del parlante. Questo approccio riflette una sensibilità etica ed epistemologica: non basta far parlare meglio, bisogna far parlare con meno paura.

In prospettiva, l'applicabilità del NEPST potrebbe estendersi anche ad altri disturbi della comunicazione, come la disartria o le afasie, soprattutto in contesti di post-ictus o neurodegenerazione, dove il carico attentivo e l'arousal influenzano la qualità dell'output verbale. Lo stesso modello potrebbe essere tradotto in altri domini motori, grazie alla sua struttura modulare basata su reinforcement learning e feedback multimodale.

Infine, l'impianto etico del modello — supervisionato costantemente, dotato di log delle decisioni e conforme ai protocolli di explainable AI — rappresenta un passo essenziale verso una terapia digitale che non solo sia efficace, ma anche trasparente, controllabile e rispettosa della vulnerabilità soggettiva.

In primo luogo, l'integrazione tra biofeedback in tempo reale e modulazione algoritmica della difficoltà consente di intervenire direttamente sulla cosiddetta finestra ottimale di attivazione neurovegetativa. Questo si basa su un principio ormai consolidato in neuroscienze cognitive: l'apprendimento motorio complesso è massimizzato in condizioni di arousal intermedio, in cui il sistema nervoso autonomo è abbastanza attivato da favorire la plasticità, ma non così iperattivato da inibire l'esecuzione fine (Yerkes & Dodson, 1908; Porges, 2011). Il modello NEPST, adattando trial fonatori sulla base dei segnali HRV ed EMG, mira esattamente a questo equilibrio neurobiologico.

In secondo luogo, la riduzione dell'attività corticale dorsolaterale (misurata tramite HbO frontale) suggerisce che il sistema ha favorito il passaggio da un controllo esecutivo consapevole a un controllo automatizzato basato su pattern motori consolidati, coerente con il concetto di de-focalizzazione attentiva proprio della teoria OPTIMAL (Wulf & Lewthwaite, 2016). Questo è cruciale nei disturbi della fluenza: mentre il parlante disfluente è spesso eccessivamente centrato sulla propria

articolazione (monitoring interno ipertrofico), l'adattamento progressivo consente una "ritirata cognitiva" dal gesto, liberando risorse per la comunicazione globale e l'interazione sociale.

Dal punto di vista logopedico e clinico, l'effetto sinergico tra stabilità neurovegetativa (HRV), efficienza corticale (HbO), riduzione della latenza decisionale (P300) e impattamento soggettivo (OASES) segnala una riorganizzazione sistemica che non è solo sintomatica ma profondamente funzionale. Il paziente non apprende semplicemente a "parlare meglio" nel contesto artificiale della terapia: le modifiche emergono come adattamenti distribuiti nei sistemi neurali che presiedono all'integrazione sensorimotoria, alla gestione dell'errore e alla valutazione della propria performance in tempo reale.

Sul piano dell'apprendimento, il reinforcement learning implementato in NEPST replica in modo formalizzato ciò che in clinica logopedica è solitamente affidato all'intuito dell'operatore: il cosiddetto *challenge-point*, ovvero il punto ottimale in cui la sfida non è né troppo facile (no plasticità) né troppo difficile (frustrazione). A differenza degli approcci manuali, l'algoritmo aggiorna la difficoltà trial-by-trial, tenendo conto di oscillazioni interne (arousal, tensione muscolare) e prestazioni esterne (frequenza disfluenze, ritmo vocale). Questo approccio dinamico alla somministrazione lo rende affine a sistemi di self-organizing neural networks, dove le traiettorie di apprendimento sono selezionate in funzione della stabilità interna della rete e non solo dell'output (Grossberg, 1987).

Un'ulteriore dimensione da considerare è quella epistemologica: NEPST si fonda su un'idea relazionale e adattiva della terapia, in cui il sistema tecnologico non è uno strumento esterno al processo terapeutico, ma un partner co-costruttivo che collabora con il clinico e con il paziente alla costruzione della fluenza. In questo senso, l'algoritmo non è una black box che "decide", ma un sistema di traduzione bio-comportamentale, che rende visibili, misurabili e modificabili aspetti del funzionamento umano altrimenti inaccessibili.

Infine, non va sottovalutata l'importanza dell'etica clinica in un modello così pervasivo. Il rispetto della supervisione clinica, la trasparenza delle policy decisionali, la possibilità per l'operatore di modificare in ogni momento le soglie adattive e di visualizzare lo storico delle scelte algoritmiche, sono elementi indispensabili per evitare il rischio di iper-autonomizzazione del sistema. In altre parole, la forza del modello NEPST non sta solo nell'efficacia adattiva, ma nella sua responsabilità epistemica e relazionale: il sistema è costruito per apprendere dal paziente, ma sempre sotto la guida del terapeuta, che resta l'interprete e garante del processo.

Conclusioni

Il presente studio ha fornito evidenze preliminari ma coerenti a sostegno dell'efficacia del modello NEPST nella riabilitazione della balbuzie evolutiva persistente, mettendo in luce l'importanza di un approccio multidimensionale che integri parametri fonologici, neurofisiologici e psicologici all'interno di una struttura adattiva controllata da algoritmi di reinforcement learning. La riduzione significativa delle disfluenze (%SD), associata a cambiamenti nei marcatori di attivazione corticale (HbO), regolazione autonoma (HRV) e latenza attentiva (P300), suggerisce che la fluenza non è soltanto un'espressione motoria, ma il risultato di una complessa dinamica inter-sistemica tra attenzione, arousal, processi fonologici e regolazione dell'errore.

L'architettura modulare dell'applicazione NEPST, unita alla possibilità di calibrare il carico fonatorio in tempo reale, ha dimostrato un potenziale clinico superiore rispetto ai protocolli tradizionali, basati su progressioni fisse. Tale risultato è particolarmente rilevante in ottica translazionale, in quanto offre una base tecnologica solida per lo sviluppo di piattaforme terapeutiche applicabili anche ad altri disturbi del linguaggio e della comunicazione, come afasia, disartria o disturbi pragmatici. Il sistema NEPST, infatti, opera secondo i principi dell'intelligenza artificiale spiegabile (XAI), integrando la capacità adattiva dell'algoritmo con il giudizio clinico continuo, evitando così il rischio di una delega cieca alla macchina.

Inoltre, la presenza di correlazioni significative tra la disfluenza residua e i predittori neuropsicologici evidenzia come sia possibile utilizzare tali indicatori in fase diagnostica per personalizzare il trattamento e predirne l'esito. La coerenza tra i dati comportamentali e i pattern neurofisiologici supporta una lettura della balbuzie come esito di una rete disfunzionale, e non di un unico deficit localizzato: un concetto che si allinea con la letteratura più recente sul funzionamento distribuito delle funzioni linguistiche e motorie (Chang et al., 2020; Etchell et al., 2018).

Nonostante i risultati promettenti, il presente studio presenta alcune limitazioni, fra cui la numerosità campionaria relativamente ridotta e la necessità di validare i risultati su popolazioni cliniche eterogenee e in contesti multicentrici. Future ricerche dovranno includere una fase di follow-up a lungo termine per valutare la stabilità dei guadagni di fluenza, nonché sperimentare ulteriori moduli adattivi (es. prosodia, interazione turn-taking) all'interno dell'applicazione. È auspicabile, inoltre, l'integrazione con sensori a maggiore risoluzione e l'estensione della piattaforma a sistemi operativi alternativi per aumentare l'accessibilità del trattamento.

In conclusione, il paradigma NEPST, nel suo attuale stadio di sviluppo, rappresenta una sintesi innovativa tra neuroscienze, logopedia e intelligenza artificiale. La sua capacità di adattarsi al profilo neurofisiologico del parlante e di integrare l'esperienza soggettiva nella logica del trattamento apre prospettive inedite per una riabilitazione realmente centrata sulla persona, fondata non solo

sull'efficienza, ma anche sulla comprensione profonda delle traiettorie neurali e affettive che sostengono il parlare umano.

Riferimenti bibliografici (APA)

Arongna, Wang, Y., Li, X., & Zhang, Y. (2020). Functional near-infrared spectroscopy and EEG study of speech fluency and disfluency in adults who stutter. **Frontiers in Human Neuroscience, 14**, 327. <https://doi.org/10.3389/fnhum.2020.00327>

Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance. **Journal of Experimental Psychology: Applied, 8**(1), 6–16. <https://doi.org/10.1037//1076-898X.8.1.6>

Bloodstein, O., & Ratner, N. B. (2008). **A handbook on stuttering** (6th ed.). Cengage Learning.

Bosshardt, H. G. (1999). Spoken language and stuttering: A review of possible relations. **Journal of Communication Disorders, 32**, 123–146. [https://doi.org/10.1016/S0021-9924\(98\)00029-0](https://doi.org/10.1016/S0021-9924(98)00029-0)

Chang, S. E., Garnett, E. O., Etchell, A. C., & Chow, H. M. (2020). Functional and neuroanatomical bases of developmental stuttering: Current insights. **The Neuroscientist, 26**(2), 156–170. <https://doi.org/10.1177/1073858419858932>

Craig-McQuaide, A., Akram, H., Zrinzo, L., & Tripoliti, E. (2014). A review of brain circuitries involved in stuttering. **Frontiers in Human Neuroscience, 8**, 884. <https://doi.org/10.3389/fnhum.2014.00884>

Eichorn, N., Maxfield, N. D., & Ferrand, C. (2019). Cognitive load and stuttering: Covert repair hypothesis revisited. **Journal of Speech, Language, and Hearing Research, 62**, 131–146. https://doi.org/10.1044/2018_JSLHR-S-17-0244

Etchell, A. C., Civier, O., Ballard, K. J., & Sowman, P. F. (2018). A systematic literature review of neuroimaging studies of developmental stuttering. **Frontiers in Human Neuroscience, 12**, 225. <https://doi.org/10.3389/fnhum.2018.00225>

Franken, M. C., Kielstra-van der Schalk, C. J., & Boelens, H. (2005). Experimental treatment of early stuttering: A preliminary study. **Journal of Fluency Disorders, 30**, 189–199. <https://doi.org/10.1016/j.jfludis.2005.05.003>

Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36*(2), 212–224. <https://doi.org/10.3200/JMBR.36.2.212-224>

Ofoe, L. C., Anderson, J. D., & Ntourou, K. (2018). Cognitive factors in childhood stuttering: A meta-analytic review. *Journal of Speech, Language, and Hearing Research*, 61*, 1011–1028. https://doi.org/10.1044/2018_JSLHR-S-16-0344

Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation*. W. W. Norton & Company.

Vasić, N., & Wijnen, F. (2005). Stuttering as a monitoring deficit. *Journal of Fluency Disorders*, 30*, 225–237. <https://doi.org/10.1016/j.jfludis.2005.06.003>

Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic Bulletin & Review*, 23*, 1382–1414. <https://doi.org/10.3758/s13423-015-0999-9>

Yaruss, J. S., & Reardon-Reeves, N. A. (2017). *Early childhood stuttering therapy: A practical guide*. Stuttering Therapy Resources, Inc.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18*(5), 459–482. <https://doi.org/10.1002/cne.920180503>

*Per l'Università,
Prof. Ciro Romano, Rettore*

Università Popolare Cattolica
E. Montemurro-T.D'Ippolito
(Ente Terzo Settore D.lgs 117/17)
C.F.: 95321630634

