

## **LogosTextToSpeech Progetto Scienza Partecipata**

Pensiamo che la capacità di esprimersi debba essere gratuita e accessibile a tutti.

Questa è la storia di come Mario, uomo autistico non fonetico di 33 anni, è riuscito da solo ad interagire verbalmente con altre persone.

### **Presentazione di LogosTextToSpeech**

LogosTextToSpeech è un software web per la sintesi vocale nato dall'idea di un'educatrice che lavora con soggetti autistici e un fisico sviluppatore full stack di software web.

Logos vuole essere un software progettato e realizzato ad hoc sulle esigenze di soggetti con disabilità fonetica e vuole diventare l'alternativa gratuita a i software per la sintesi vocale attualmente in commercio.

### **Genesi dell'idea**

Tutto ha avuto inizio nel gennaio del 2021 quando un'educatrice specializzata nella sindrome dello spettro autistico lavorando con un ragazzo di 33 anni, che chiameremo Mario (nome di fantasia), soggetto autistico non fonetico vede in lui la potenzialità di esprimersi attraverso un computer tramite un software di conversione di un testo scritto in linguaggio parlato. Gli educatori tentarono con una app per tablet scaricabile gratuitamente ed ebbero subito ottimi risultati. Mario si dimostrò una persona con un vasto vocabolario, e con elevate capacità di espressione, cosa mai verificabile fino a quel momento data la sua caratteristica non fonetica e la sua difficoltà a scrivere con la penna. La app per tablet non è progettata ad hoc sulle esigenze di Mario e numerosi aspetti rallentarono i suoi progressi. Fu a quel punto che mi venne chiesto un aiuto per scegliere la soluzione tecnologica che più si adattasse alle necessità di Mario. Dopo una riunione decidemmo che la app per tablet era limitata e che avremmo dovuto sviluppare un software che soddisfacesse le specifiche esigenze di Mario. La app salva solo le ultime 10 frasi scritte, c'è poco spazio per scrivere (una ventina di caratteri per frase, dopodichè risulta impossibile vedere ciò che è stato scritto), scarsa intelligibilità e feedback della tastiera a schermo del tablet e nessuna possibilità di customizzare la voce (elemento da noi considerato fondamentale per un soggetto non fonetico). Per sviluppare la versione 0.1 di LogosTextToSpeech, un software che soddisfacesse tutte le nostre esigenze, ci vollero una decina di notti, nel frattempo fornii un computer portatile a Mario per permettergli di scrivere con una tastiera da pc, ma rilevammo subito alcune difficoltà. Il primo scoglio da superare fu la difficoltà di Mario nell'utilizzare una tastiera per pc standard: troppi tasti rispetto al necessario, troppe funzioni per ogni tasto e, nonostante il miglior feedback della tastiera fisica rispetto a

quella a schermo, la difficoltà per Mario nel premere i tasti autonomamente. Per aggirare i primi due ostacoli fu sufficiente costruire una tastiera semplificata con soli 40 tasti (26 lettere, 10 cifre, punto, punto interrogativo, spazio e invio) utilizzando un Arduino Leonardo, in modo da permettere a Mario di concentrarsi solo sui tasti strettamente necessari alla creazione della frase. Il costo complessivo della tastiera non supera le 50 euro. Per superare il terzo ostacolo l'educatrice e i suoi colleghi hanno inizialmente aiutato Mario a muovere la mano sulla tastiera reggendogli il polso. Con il passare delle settimane hanno iniziato a reggergli l'avambraccio, poi il gomito fino a quando per scrivere è stato sufficiente appoggiargli una mano sulla spalla. Da marzo 2021 a settembre 2021 Mario fece degli enormi passi avanti, prese dimestichezza con la tastiera custom e la disposizione dei tasti tanto che iniziò ad usare il software con la tastiera del pc. Utilizzando Logos Mario è riuscito ad interagire con altre persone in un modo nuovo, ha permesso alla famiglia e agli educatori di capire che ha un vastissimo vocabolario ed eccellenti capacità espressive.

### **La tastiera**

Il prototipo della tastiera custom è stato creato usando un Arduino Leonardo, quattro partitori di tensione realizzati ad hoc e 40 bottoni normalmente aperti montati su uno chassis in plastica autoprodotta. La mia proposta iniziale fu di disporre i bottoni in modo più ergonomico: Mario per premere i tasti usa solo il dito indice della mano destra, disponendo i bottoni "ad arco" avremmo potuto minimizzare i suoi movimenti, facendo perno sul gomito, e rendergli la scrittura più semplice. Ma su sua richiesta esplicita i bottoni sono stati disposti in modo simile ad una tastiera standard per pc. Feci molta attenzione a molti fattori: colore e dimensione dei bottoni, dimensione e font delle lettere scritte sui bottoni, distanza tra i bottoni, colore e texture della plastica ecc... per minimizzare il numero di elementi "disturbanti" ma Mario si dimostrò molto determinato nell'imparare ad usarla e tutti i miei dubbi vennero risolti nel momento in cui, riferendosi al pc, scrisse "lamiapossibilitaperlavorare".

### **Il software**

Il software è sviluppato in php (framework Laravel) e javascript (framework VueJs). Il motore database è MySQL. L'interfaccia è disegnata per essere fruibile su desktop o su tablet. Ogni utente del software ha un dizionario di parole personale sempre in aggiornamento in cui viene registrata la frequenza del loro utilizzo e un dizionario di relazioni tra parole consecutive, che viene utilizzato per fornire i suggerimenti. Ogni frase inserita viene sintetizzata alla pressione del tasto invio e viene salvata a database.

La schermata principale consiste in una grande textarea (fino a 5 righe di testo, richiesta esplicita di Mario). In modalità tablet, sotto alla textarea è presente una zona che contiene fino a 4 suggerimenti. Esistono suggerimenti di due tipi: mentre si sta scrivendo una parola vengono proposte fino a 4 parole che sono simili a quella che si sta scrivendo considerando anche gli errori di battitura, finito di scrivere una parola vengono suggerite le 4 parole che la seguono più frequentemente. Sempre quando si è in modalità tablet, sotto alla zona suggerimenti è presente una tastiera più semplice di quella del tablet a schermo. L'ultima zona al di sotto di tutto è l'elenco delle frasi in ordine decrescente di inserimento con il bottone "Leggi" per rileggerle nuovamente.

Nella sezione "Il mio profilo" è possibile customizzare la voce scegliendo tra 4 voci (2 maschili e 2 femminili) e viene data la possibilità di modificarne il tono e la velocità. Inoltre è possibile attivare la lettura ad alta voce dei tasti premuti (opzione disattivata per Mario, per lui lo stimolo acustico è disturbante).

Il motore di sintesi utilizzato è quello fornito da Google Cloud, a pagamento. Verranno dati maggiori dettagli nella sezione "Motori di sintesi vocale".

Per gli educatori è presente una sezione in cui è possibile vedere le statistiche sulle parole utilizzate e le frasi sintetizzate.

### **Changelog di LogosTextToSpeech**

Versione 0.1: singolo utente, sintesi vocale tramite svariati software free su S.O. Linux.

Versione 1.0: multi utente, sintesi vocale tramite api di Google Cloud, implementazione sezione statistiche.

Versione 1.1: customizzazione della voce, grafica responsive per tablet, aggiunta tastiera virtuale customizzata per tablet.

### **Future implementazioni**

Il software è attualmente in versione beta, stiamo ancora attendendo gli esiti dei test da parte di Mario e degli educatori. Siamo intenzionati ad implementare altre soluzioni come una tastiera programmabile a video per la CAA, potenziare il sistema di statistiche per gli educatori in modo da poter visualizzare i progressi ecc...

### **Motori di sintesi vocale**

Tra gli anni 90 e gli anni 2000 sono stati sviluppati numerosi motori di sintesi vocale gratuiti per piattaforme linux e sono facilmente integrabili in un software. Il loro principale problema è la qualità della voce sintetizzata, non è per niente naturale, molto robotica, e alcuni fonemi non

vengono pronunciati correttamente (con rotacismo o accenti sulla sillaba sbagliata). Negli ultimi 20 anni sono stati messi a punto motori di sintesi vocale molto più performanti e che suonano in modo più “naturale”, il loro unico difetto è che sono a pagamento. Google Cloud, ad esempio, mette a disposizione il suo motore di sintesi in modo gratuito fino a 1 milione di caratteri al mese e a pagamento dal milione in poi. Questo è un problema se il software deve essere distribuito in modo gratuito. Abbiamo pensato ad un paio di soluzioni: La prima, efficace ma poco trasparente, è quella di aprire una decina di account Google e avere così a disposizione molti milioni di caratteri al mese, la seconda invece intende fare leva sul fatto che Google fornisce già alcuni servizi gratuiti per collaborazioni con associazioni no-profit o altro, ma questo servizio ad oggi non è incluso. La nostra speranza è che, anche grazie alla partecipazione a questo bando, Google ci rilasci delle chiavi per l’attivazione del servizio illimitato.

### **Conclusioni**

Questo software è nato per permettere ad un soggetto autistico non fonetico di “parlare”. Speriamo che voi vediate tutto il potenziale che vediamo noi in Logos.