

La variazione prosodica in italiano: l'utilizzo di un chatbot Telegram per la didattica assistita per apprendenti di italiano L2 e nella valutazione linguistica delle conoscenze disciplinari

Abstract

The comparison of foreign language learners' pronunciation with the utterances of native speakers is receiving increasing attention thanks to the new applications of computer-assisted teaching. The studies on prosodic variation between several native speakers reveal rhythmic-intonative patterns that should be part of the linguistic baggage of the native speaker; it seems problematic how to make explicit to the learner his degree of prosodic competence on the basis of the different acoustic indices (f0, intensity, pitch) and other parameters (number of syllables, lexical accent, speech rate, pauses). We present a chatbot designed as a proactive learning support for the improvement of oral skills in Italian L2. After describing its technical structure, the first results of a preliminary study are discussed. They provide a first starting point to reflect on some considerations that emerged.

1. Introduzione

Le caratteristiche prosodiche del parlato rivelano importanti informazioni come la modalità enunciativa, lo stile d'eloquio, l'attitudine e le connotazioni diatopiche o diastratiche del locutore; esse sono quindi essenziali per una buona riuscita della comunicazione interlocutiva (Hirst 1983, Cruttenden 1986). Tuttavia, in molti contesti relativi alla didattica delle lingue straniere (LS) la riflessione metaprosodica non è sempre soddisfacente per lo studente poiché non viene affrontata in modo esaustivo (Trouvain & Gut 2007). Riferirsi al ruolo di caratteristiche prosodiche come l'intonazione, il ritmo, la durata, la focalizzazione per spiegare differenze comunicative di significato può invece rivelarsi indispensabile per innescare nello studente una consapevolezza utile per riconoscere e riprodurre il parlato della LS (Cresti 1999, Chun 2002, Frost & Picavet 2014). Inoltre, ormai da diversi decenni, gli sviluppi tecnologici nell'analisi del parlato hanno aperto la strada a nuove prospettive di interazione comunicativa declinabili anche per la didattica delle lingue (Chun 1998, Cazade 1999), facendo emergere, sin dai primissimi studi condotti, anche varie criticità verso le possibili soluzioni di esplicitazione del grado di compe-

¹ Università degli Studi di Torino.

² Politecnico di Torino.

tenza raggiunto, ad esempio tramite l'utilizzo di rappresentazioni grafiche di supporto (James 1976, De Bot 1983, Martin 2010).

La didattica incentrata sull'oralità di una LS permette proprio quel tipo di riflessione metaprosodica grazie alla quale lo studente può sperimentare la varietà di combinazioni enunciative presenti nella lingua appresa e associarne il significato in termini di funzioni specifiche (modali, sintattiche, informative, emozionali) per affinare le proprie competenze comunicative orali (Delattre 1966, Cresti 1999). In italiano bisogna inoltre tener conto delle diverse scelte intonative anche sulla base della varietà diatopica utilizzata (Canepari 1985, Soriano 2006), una fonte di variabilità inclusa ormai da diverso tempo negli spazi della ricerca sull'intonazione della L2 (De Meo & Pettorino 2012, De Marco *et al.* 2014) che, discussa nei contesti didattici, porterebbe a una maggior sensibilizzazione da parte dell'apprendente, molto spesso esposto a un solo modello diatopico. Infine, ci sembra interessante soffermarci sulle diverse possibilità di realizzazione tipiche della lingua orale e su come occorra tenerne conto nel contesto della valutazione automatica e nella ricorrenza di andamenti intonativi.

Sulla base di queste considerazioni, questo studio si struttura nel seguente modo: dopo aver presentato l'architettura del chatbot, analizzeremo come avviene il confronto tra i dati di parlanti madrelingua italiana e apprendenti di italiano L2 a partire da un set di dieci frasi complesse costruite in modo da dare un'adeguata rappresentazione delle più frequenti soluzioni intonative di un parlato elicitato su contenuti di natura tecnico-scientifica, includendo però al tempo stesso anche alcune espressioni quotidiane. Discuteremo infine, sulla base dei risultati ottenuti da parte degli italofoeni e degli studenti, una serie di riflessioni nell'ottica futura di implementare successivi parametri per la valutazione del *chatbot*, nello specifico in che modo i valori dei parametri acustici rilevati dal chatbot rispecchiano una valutazione percettiva (Munro & Derwing 1999).

2. Architettura e funzionamento del chatbot

In questo paragrafo illustriamo in che modo è stato strutturato il *chatbot*³ come strumento di didattica assistita per la valutazione di tratti prosodici. L'adozione di un chatbot di questo tipo permette diversi vantaggi. Innanzitutto prevede una modalità di interazione, basata su testo, a cui il pubblico è già largamente abituato grazie alla popolarità di applicazioni di messaggistica istantanea quali *Whatsapp* o *Telegram*. Ciò rende questi strumenti più facilmente fruibili rispetto ad applicazioni dedicate o ai cosiddetti *Learning Management System* (LMS) come *Moodle* o *Blackboard*, il cui maggior ostacolo alla fruibilità è rappresentato proprio dalla necessità da parte degli utenti di adattarsi a un'interfaccia poco familiare. Un *chatbot* permette inoltre di interagire con gli apprendenti in maniera strutturata e asincrona: strutturata

³ Il *chatbot* è un software progettato per simulare una conversazione con un utente umano (Fernoagă *et al.* 2018) che vede, tra le diverse applicazioni, anche quella educativa e didattica (Colace *et al.* 2018).

perché organizzata secondo schemi predefiniti, volti a ridurre le ambiguità; asincrona perché non vincolata a intervalli di tempo continuativi ma fruibile liberamente secondo le proprie disponibilità di tempo. Le potenzialità relative all'uso di *chatbot* come strumenti di ausilio alla didattica sono investigate già da alcuni anni (Pereira 2016, Fernoagà *et al.* 2018) e in diversi casi si osserva un orientamento verso la valutazione didattica della pronuncia (Cheng *et al.* 2020) e della prosodia (Lezhenin *et al.* 2017). Realizzato all'interno dell'applicazione di messaggistica istantanea *Telegram*, il *chatbot* (o bot) qui presentato prevede l'interazione con l'utente attraverso domande e risposte basate su una serie di domande a risposta chiusa (quiz) che possono avere carattere generale di comprensione linguistica o essere legate a un particolare ambito disciplinare (ad esempio tecnico-scientifico). Una volta individuata la risposta corretta, l'apprendente invia la registrazione della frase letta sotto forma di messaggio vocale. L'enunciato ricevuto viene così processato automaticamente dal bot al fine di ottenere una valutazione del livello intonativo del parlante. Tale processamento è riassumibile nei seguenti passi:

1. Preprocessing
2. Segmentazione fonetica
3. Estrazione dei valori di frequenza fondamentale
4. Analisi intonativa

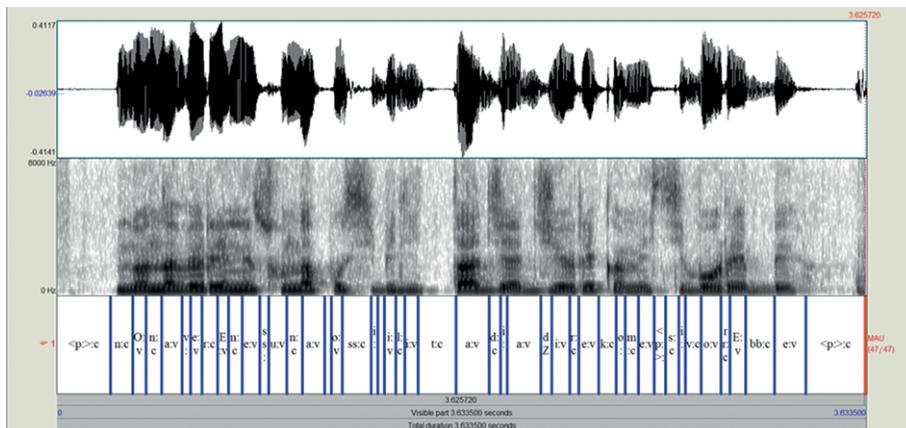
Il primo passo nella catena di processamento ha un duplice scopo: garantire un livello di qualità minimo per il segnale audio di input e preparare quest'ultimo alle elaborazioni successive. L'audio viene innanzitutto convertito in formato *wav* a singolo canale, ricampionato (se necessario) a 48 kHz e ripulito dal rumore di fondo costante⁴. Vengono poi stimati il valore medio di ampiezza e il rapporto segnale/rumore dell'audio risultante. Nel caso in cui i valori stimati risultino inferiori a determinate soglie predefinite, il *chatbot* richiede all'apprendente di registrare un nuovo enunciato in un ambiente meno rumoroso e/o parlando a un tono di voce più alto. Il segnale vocale viene successivamente sottoposto a segmentazione mediante il servizio web WebMAUS Basic (Kisler *et al.* 2018). Tale servizio, una volta inviato il segnale vocale come input e la trascrizione ortografica dell'enunciato fornita dal bot, ne restituisce una segmentazione in parole e fonemi sulla base del metodo descritto in Schiel (1999). La segmentazione fonetica dell'enunciato, fornita in formato *TextGrid* per agevolare la successiva elaborazione mediante il software Praat (Boersma & Weenink 2018), viene poi processata etichettando i singoli foni come vocali o consonanti (fig. 1). Uno script Praat⁵ viene poi richiamato per estrarre i va-

⁴ Questa procedura non pregiudica la valutazione poiché rimuove solo quelle componenti di rumore costante del segnale, ad esempio ronzii, fischi e sibili tipicamente dovuti all'apparecchiatura utilizzata per la registrazione o alla vicinanza del parlante a fonti di rumore continuativo.

⁵ Lo script è stato creato da Antonio Romano per agevolare l'applicazione del metodo AMPER nell'ambito di Matlab™ e successivamente implementato per Praat da Plinio Barbosa (Campinas, Brasile) e adattato da Albert Rilliard (Parigi). Lo script è stato successivamente automatizzato da Marco Gamba (UniTo) e, per questo lavoro, da Marco Palena.

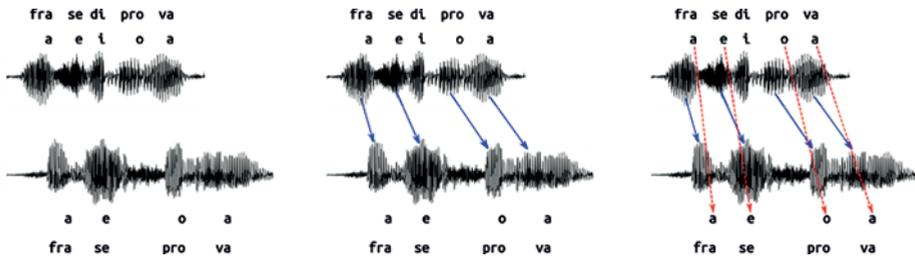
lori di frequenza fondamentale (f_0) dei foni vocalici⁶ precedentemente identificati, ottenendo così la curva intonativa dell'enunciato (per il modello di segmentazione ed estrazione dei valori, cfr. Romano *et al.* 2014).

Figura 1 - Finestra di Praat (wav e Textgrid) con la segmentazione fonetica effettuata da WebMaus della frase "Non avere nessuna possibilità di agire come si vorrebbe"



L'analisi dell'intonazione avviene mediante il confronto tra la curva intonativa dell'enunciato dell'apprendente e i tracciati f_0 dei corrispettivi enunciati di parlanti madrelingua, precedentemente raccolti e valutati utilizzando la medesima procedura automatica. Il confronto avviene mediante il calcolo di una misura di correlazione (Castro Moutinho *et al.* 2011) che confronta per ogni frase tre punti (iniziale, centrale, finale) di f_0 di ogni segmento vocalico individuato da Maus. Data la natura estemporanea dei segnali vocali registrati e la modalità automatica della procedura di segmentazione operata da Maus, i segmenti vocalici rilevati per gli enunciati degli apprendenti e quelli rilevati per i corrispettivi enunciati di parlanti madrelingua potrebbero differire. Pertanto, prima di calcolare la correlazione, i segmenti dei due parlanti vengono allineati sulla base sia dell'informazione fonetica in essi contenuta sia della loro posizione temporale (fig. 2). Questa procedura di allineamento serve per mettere in corrispondenza i segmenti dei due parlanti in modo da sapere quali confrontare tra loro per il calcolo della correlazione, quindi solo i segmenti per cui si è trovata corrispondenza sono inclusi nel calcolo della correlazione. Sono esclusi invece tutti i segmenti vocalici per cui non si trova corrispondenza fonetica tra i due enunciati. Ad esempio, se l'enunciato di un parlante include in una certa posizione un segmento etichettato come [a] mentre l'enunciato dell'altro parlante non lo include, quel segmento non viene preso in considerazione come punto ai fini del calcolo della correlazione. Questo perché per calcolare la correlazione abbiamo bisogno di due serie di dati della stessa numerosità.

⁶ Sono stati considerati solo i segmenti vocalici che MAUS etichetta con i seguenti fonemi: aææœœçø øiyiyzæz-ëäöë-ihəöïüäwλwçuπuσy.

Figura 2 - *Audio segmentato, allineamento temporale, corrispondenza dei segmenti*Figura 3 - *Schermate esemplificative dei passaggi principali dell'interazione chatbot-utente*

Lo svolgimento del task (fig. 3) prevede una prima fase di registrazione in cui l'utente fornisce alcuni dati sociolinguistici (il livello linguistico in base al QCER, se e dove ha vissuto in Italia, in che contesti è abituato a parlare italiano). Successivamente,

dopo aver effettuato una prova microfono che garantisca una sufficiente qualità dell'audio, l'utente può cominciare il task che consiste nel rispondere a 10 domande in italiano di discipline varie (grammatica, storia, matematica) registrando mano a mano la lettura individuale di ogni risposta: ciascuna domanda è presentata in forma scritta e orale⁷ mentre la scelta della risposta è compresa tra quattro opzioni (cfr. appendice per l'elenco completo delle frasi).

Se l'utente sbaglia risposta, può riprovare e, una volta selezionata la risposta corretta, il *chatbot* propone di registrare la risposta tramite un messaggio vocale. A questo punto dell'interazione, la frase elicitata viene comparata con quelle presenti nel database prodotte da parlanti italofofoni (si veda § 3 per la descrizione dettagliata della base dati) e dopo aver individuato quella con la vicinanza intonativa maggiore, il *chatbot* restituisce all'apprendente il valore di correlazione in percentuale⁸. Al completamento del task, viene infine restituito un punteggio complessivo ottenuto dalla media dei valori percentuali per ciascuna risposta.

3. *La base dati di riferimento di parlanti italofofoni*

Il confronto con dati di parlanti madrelingua è impiegato da diversi anni (Delmonte 2009)⁹ e si sposa perfettamente con le nuove modalità di valutazione dei sistemi automatici di riconoscimento del parlato. In una prima fase del progetto abbiamo quindi creato un set di 10 domande e risposte da utilizzare come dataset di riferimento da somministrare successivamente agli studenti. Nelle risposte, tra cui è presente la frase target, sono state inserite date, formule matematiche¹⁰ e strutture sintattiche mediamente complesse (l'elenco delle frasi è consultabile in appendice) per avere degli enunciati di difficoltà di lettura medio-alta, rivolte a un pubblico di studenti con un livello minimo B1. Il set è stato quindi sottoposto alla lettura da parte di 250 persone che hanno registrato domanda e risposta tramite un dispositivo cellulare portando alla creazione di una base di dati di 2500 frasi. Gli audio sono stati quindi ricampionati a 16kHz e, tramite una segmentazione manuale, sono state segmentate ed etichettate soltanto le risposte. I campioni registrati provengono da diverse parti d'Italia (la metà dal Piemonte)¹¹ e si dividono in 164 femmine e 86 maschi con un'età compresa tra i 15 e i 70 anni. La comparazione permette di trovare la frase all'interno della base di dati dei campioni italofofoni che più si avvicina a quella

⁷ La domanda è letta da uno speaker professionista senza inflessioni regionali.

⁸ Il confronto tra utente e parlante di riferimento non è legato al genere poiché il confronto si basa sulle differenze di intonazione e non sulla frequenza media del parlante.

⁹ Una rassegna bibliografica dettagliata è disponibile in Romano & Giordano (2017).

¹⁰ La lettura di formule matematiche e, in generale, di numeri può essere complessa per un apprendente L2, non a caso l'ultima risposta del set (" $1/6 + 3/2$ ") ha ricevuto da parte degli utenti una lettura per lo più erronea (si veda tabella in appendice).

¹¹ Sebbene in questa prima fase gli italofofoni registrati provengano principalmente dal Piemonte, uno degli obiettivi futuri consiste nell'ampliare la base di dati così da avere una proporzione più significativa tra le varie aree regionali d'Italia e poter quindi eventualmente correlare anche le caratteristiche prosodiche regionali.

pronunciata dallo studente e di restituirne la percentuale di correlazione massima. È importante sottolineare che la percentuale restituita alla fine di ogni risposta data esprime esclusivamente il parametro acustico responsabile dell'intonazione, ovvero la frequenza fondamentale (f_0): di conseguenza, per offrire una valutazione più accurata, le frasi di riferimento prodotte dagli italofoeni sono state valutate da parte di due operatori umani per poter successivamente correlare alcuni indici acustici (come la velocità d'eloquio, la fluenza o la scansione pausale) con la valutazione percettiva. Nello specifico, per ogni frase si segnalano¹²:

1. la velocità di lettura (lenta, normale o veloce);
2. l'accento regionale (punteggio da 1 – accento non riconducibile a una specifica varietà regionale – a 3 – accento marcato regionalmente);
3. la scansione delle parole (lenta, normale o veloce);
4. la fluenza intonativa (punteggio da 1 – min spontaneità – a 3 – max spontaneità);
5. l'età del locutore (bambino, ragazzo, adulto o anziano).

4. *Primi risultati su un campione di studenti francofoni*

In un primo tempo¹³, per testare il chatbot su un gruppo linguistico omogeneo, il task è stato sottoposto a 15 studenti universitari francofoni (12 femmine e 3 maschi) di un'età compresa tra i 20 e i 22 anni che studiano l'italiano da 2-5 anni¹⁴; dal totale di 150 frasi sono state escluse 15 frasi perché non sono state registrate dall'utente o perché non sono state lette correttamente¹⁵. Le frasi prodotte sono state quindi confrontate con quelle della base di dati che, per questo set, hanno ricevuto generalmente una valutazione percettiva massima (97% dei casi) e media (3%). Gli utenti francofoni invece hanno ricevuto un punteggio alto (3) nel 47% dei casi (63 frasi), medio (2) per il 47% (63 frasi) e basso (1) nel restante 6% (9 frasi). Mediamente per ciascuna frase gli utenti hanno ricevuto una valutazione percettiva dell'intonazione di 2 nella metà dei casi e 3 nel restante, tranne la frase 9 per la quale soltanto 5 utenti hanno ricevuto una valutazione media (2). Se osserviamo i confronti tra le frasi degli utenti francofoni con i riferimenti italofoeni con un punteggio alto, notiamo che la media della vicinanza intonativa delle frasi si attesta tra il 37% e il 94%. Più precisamente 6 frasi hanno una vicinanza compresa tra il 25 e il 50%, 32 tra il 50 e il 75% e 24 tra il 75 e il 100%. Nel confronto grafico di due curve intona-

¹² In questa sede viene presa in esame soltanto la fluenza intonativa, gli altri parametri percettivi sono in corso di analisi e saranno oggetto di un altro lavoro.

¹³ Il task è stato somministrato nel tempo a più gruppi di studenti con provenienza mista (anglofoni, ispanofoni, ecc.) ma per questa prima presentazione si è scelto di focalizzare l'analisi sul campione di francofoni.

¹⁴ In questo lavoro non abbiamo preso in considerazione e correlato le caratteristiche sociolinguistiche degli studenti limitandoci a una presentazione dei dati qualitativa.

¹⁵ In questo caso, ci riferiamo all'ultima frase ("1/6 + 3/2") che è stata letta "un sesto più tre mezzi" soltanto da 4 utenti.

tive utente-campione riferimento (fig. 4) con un'alta percentuale di correlazione del 92%, si nota ad esempio un andamento di f_0 complessivamente simile.

Figura 4 - Confronto delle curve intonative tra parlante di riferimento (segmenti continui) e utente (segmenti tratteggiati) della frase "Il 14 luglio 1789" a partire dai valori dei segmenti vocalici estratti (righe verticali)

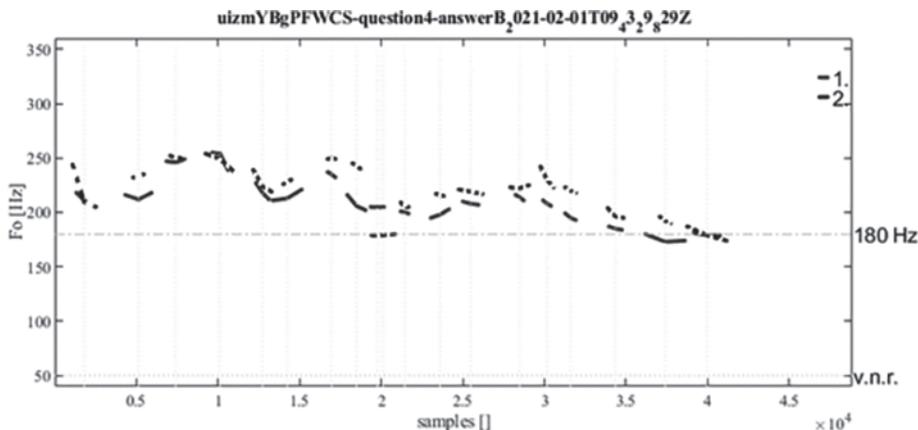
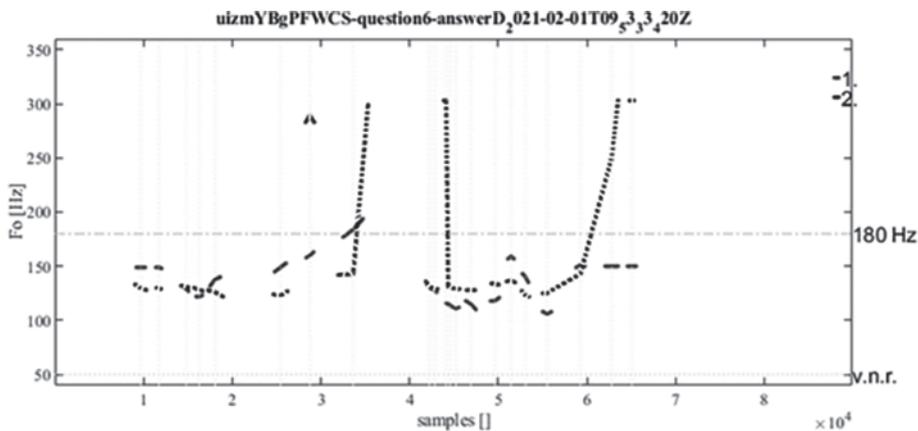


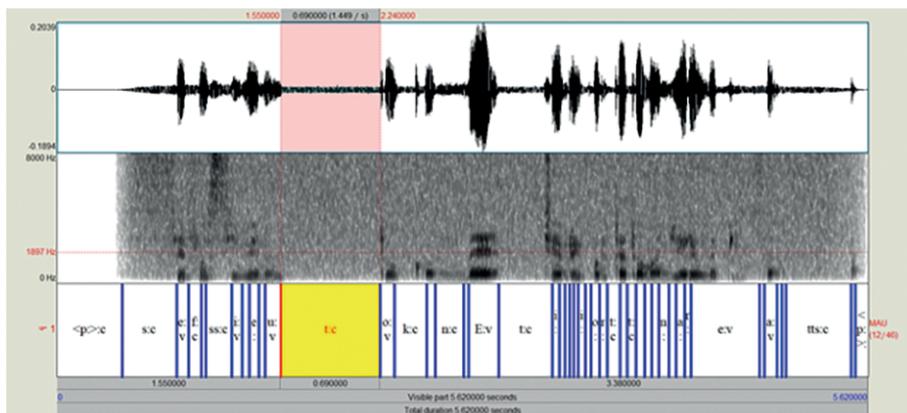
Figura 5 - Confronto delle curve intonative tra parlante di riferimento (segmenti continui) e utente (segmenti tratteggiati) della frase "Se fossi venuto con me, ti avrei portato a mangiare una pizza" a partire dai valori dei segmenti vocalici estratti (righe verticali)



Sono stati anche analizzati quei casi in cui sia il locutore italofono che l'utente francofono hanno un punteggio percettivo alto (3) ma la percentuale di correlazione intonativa è bassa: poiché ci aspettiamo che laddove la valutazione percettiva sia in entrambi i casi alta, anche il punteggio sia alto, occorre capire se un'eventuale percentuale bassa sia dovuta ad altri fattori (segmentazione di Maus non corretta, qualità scarsa della registrazione, bassa presenza di segmenti allineabili). La figura 5 mostra ad esempio che, nonostante la valutazione percettiva dell'intonazione

sia alta per l'utente e l'italofono di riferimento, la correlazione è comunque bassa (27%), forse a causa della segmentazione (e quindi dei valori estratti di f_0) fornita da Maus. In questa stessa frase, inoltre, subentrano altri fattori come un'esitazione all'interno di parola (fig. 6), che andranno segnalati in una prospettiva di valutazione della produzione orale complessiva.

Figura 6 - Finestra di Praat (wav e Textgrid) con la segmentazione fonetica effettuata da WebMaus per la frase letta dall'utente "Se fossi venuto con me, ti avrei portato a mangiare una pizza"



5. Discussione e conclusioni

La valutazione delle componenti prosodiche rappresenta una sfida importante nell'insegnamento delle lingue perché porta lo studente a riflettere su come la variazione degli indici acustici nel parlato rifletta dei cambiamenti di significato sotto diversi aspetti. In questa prima fase del progetto che abbiamo presentato ci siamo concentrati su come rendere effettivo un chatbot per la valutazione linguistica e su quali indici acustici basarsi per aiutare lo studente a prendere coscienza delle proprie competenze orali in italiano. Partendo dalla struttura del chatbot, abbiamo voluto in questa sede concentrarci sull'esposizione delle fasi per il completamento del task. Una prima valutazione percettiva delle frasi prodotte dai campioni di riferimento e dagli studenti di italiano L2 ci ha portato a fare una serie di considerazioni: in generale, le frasi prodotte dagli studenti hanno ottenuto un punteggio medio-alto e laddove il punteggio degli studenti e dei campioni fosse alto, le percentuali di vicinanza intonativa si attestano per la maggior parte delle frasi oltre il 50%. Aspettandoci che a un alto punteggio da parte del campione e dell'utente corrispondesse anche una percentuale di vicinanza intonativa alta, riteniamo utile analizzare prossimamente se, tra le motivazioni di una bassa percentuale, possano essere intervenuti fattori esterni legati alla procedura automatica di segmentazione ed estrazione o altre componenti extra-intonative. Restano altri passaggi di analisi che prevediamo di approfondire nella prossima fase del progetto. Avendo ottenuto risultati rassicuranti su

un largo campione, è in corso la fase di quantificazione e soluzione dei casi in cui gli enunciati si presentino difforni da quelli attesi a causa di un ordine di parole diverso e/o della presenza di frammenti involontari di parlato (balbettii, esitazioni, false partenze) che causano differenze tali da impedire il conseguimento di una minima distanza prosodica. Dopo aver concluso l'etichettatura della base dati per gli altri parametri descritti, vorremmo concentrarci su quelli prodotti dagli studenti per una valutazione fonetico-percettiva da comparare in un secondo tempo con gli indici acustici principali. In un'ultima fase, vorremmo incrementare ed equilibrare il corpus di riferimento dei parlanti madrelingua italiani per ampliare le varietà enunciate in termini diatopici e diafasici (Crocco 2017). In questa direzione, un ulteriore passo riguarda l'eventuale classificazione dalla quale partire per addestrare un algoritmo in grado di sostituire l'operatore umano lungo una linea che distingue un parlato più artefatto da uno sciolto e spontaneo (Nencioni 1976, Voghera 1989, Zmarich *et al.* 1996).

Infine, un aspetto ancora poco approfondito riguarda la lettura di enunciati di lettura più complessa, che prevedano la presenza di semplici formule matematiche, sigle, acronimi e forestierismi, senza trascurare le espressioni richieste dalle soluzioni enunciate tipiche di alcuni linguaggi specialistici. Su questi aspetti nessuna formazione specifica è prevista nei curricula di avvicinamento alle materie tecnico-scientifiche e nessuna pubblicazione nazionale dà indicazioni esaustive. Ad esempio, per quanto tutti riconoscano espressioni grafiche come “3/2”, “12,1%”, “011 6709718”, “FBI”, “report” etc., non esiste una fonte pubblicamente disponibile in grado di chiarire a un apprendente di italiano la pronuncia (o le pronunce) più tipiche di questi oggetti linguistici, nelle preferenze dei parlanti nativi (Fry 1989).

Ringraziamenti

Questo progetto ha beneficiato di un finanziamento della Fondazione CRT – Bando Erogazioni Ordinarie 2020 - CALL-UniTO. Si ringrazia la prof.ssa Donatella Bisconti e tutti coloro che hanno preso parte alla registrazione delle frasi e gli studenti che si sono sottoposti al task.

Bibliografia

- Bailey, D. (2019). Chatbots as conversational agents in the context of language learning. In *Proceedings of the Fourth Industrial Revolution and Education*, 32-41.
- Berruto, Gaetano. 2012. *Sociolinguistica dell'Italiano contemporaneo*. Roma: Carocci.
- Boersma, Paul & Weenink, David. 2018. *Praat: doing phonetics by computer* [Computer program]. Version 6.0.37, retrieved 14 March 2018 from <http://www.praat.org/>.
- Busà, Maria Grazia. 2012. The role of prosody in pronunciation teaching: a growing appreciation. In Busà, Maria Grazia & Stella Antonio (eds.), *Methodological Perspectives on Second Language Prosody*, 101-105. Padova: Cleup.

- Canepari, Luciano. 1983. *Italiano standard e pronuncia regionale*. Padova: CLEUP.
- Canepari, Luciano. 1985. *L'intonazione. Linguistica e paralinguistica*. Napoli: Liguori.
- de Castro Moutinho, Lurdes & Coimbra, Rosa Lídia & Rilliard, Albert & Romano, Antonio. 2011. Mesure de la variation prosodique diatopique en portugais européen. *Estudios de fonética experimental* 20. 33-55.
- Cazade, Alain. 1999. De l'usage des courbes sonores et autres supports graphiques pour aider l'apprenant en langues. *Apprentissage des Langues et Systèmes d'Information et de Communication, online* 2(2). 3-32.
- Cheng, Victor Chi-Wa & Lau, Vincent King-Tang & Lam, Ringo Wai-Kit & Zhan, Tian-Jie & Chan, Pak-Kin. 2020. Improving English phoneme pronunciation with automatic speech recognition using voice chatbot. In Lee, Lap-Kei & U, Leong Hou & Wang, Fu Lee & Cheung, Simon K. S. & Au, Oliver & Li Kam Cheong (eds.), *International Conference on Technology in Education*, 88-99. Singapore: Springer.
- Chun, Dorothy. 1998. Signal analysis software for teaching discourse intonation. *Language Learning & Technology* 2. 61-77.
- Chun, Dorothy. 2002. *Discourse Intonation in L2: From theory and research to practice*. Amsterdam: Benjamins.
- Colace, Francesco & De Santo, Massimo & Lombardi, Marco & Pascale, Francesco & Pietrosanto, Antonio & Lemma, Saverio. 2018. Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research* 7(5). 528-533.
- Cresti, Emanuela. 1999. Force illocutoire, articulation topic/comment et contour prosodique en italien parlé. *Faits de langue* 13. 168-181.
- Crocco, Claudia. 2017. Everyone has an accent: standard Italian and regional pronunciation. In Cerruti, Massimo & Crocco, Claudia & Marzo Stefania (eds.), *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, 89-117. Berlin/New York: Mouton de Gruyter.
- Cruttenden, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press.
- De Bot, Kees. 1983. Visual feedback of intonation I: Effectiveness and Induced Practice Behavior. *Language and Speech* 26(4). 331-350.
- De Iacovo, Valentina & Romano, Antonio. 2019. DataDriven intonation teaching: an overview and new perspectives. *EL.LE*, 8(2). 393-408.
- De Marco, Anna & Sorianello, Patrizia & Mascherpa, Eugenia. 2014. L'acquisizione dei profili intonativi in apprendenti di italiano L2 attraverso un'unità di apprendimento in modalità blended learning. In De Meo, Anna & D'Agostino, Mari (a cura di), *Varietà dei contesti di apprendimento linguistico*, 189-210. Milano: AIELEA
- De Meo, Anna & Pettorino Massimo (eds.), 2012. *Prosodic and Rhythmic Aspects of L2 Acquisition: The Case of Italian*. Cambridge: Cambridge Scholar Publishing.
- Delattre, Pierre. 1966. Les dix intonations de base du français, *French Review* 40. 1-14.
- Delmonte, Rodolfo. 2009. Prosodic tools for language learning. *International Journal of Speech Technology* 12(4). 161-184.
- Fernoagă, Vlad & Stelea George Alex & Gavrilă, Cristinel & Sandu, Florin. 2018. Intelligent education assistant powered by chatbots. *The International Scientific Conference eLearning and Software for Education* 2. 376-383.

- Frost, Dan & Picavet, Francis. 2014. Putting prosody first – Some practical solutions to a perennial problem: The Innovalangues Project. *Research in Language* 12(3). 233-243.
- Fry, Edward B. 1989. Reading formulas: maligned but valid. *Journal of Reading*, 32(4). 292-297.
- James, E.F. 1976. The acquisition of prosodic features of speech using a speech visualizer. *International Review of Applied Linguistics in Language Teaching* 14(3). 227-244.
- Kisler, Thomas & Reichel, Uwe & Schiel, Florian. 2017. Multilingual processing of speech via web services. *Computer Speech & Language* 45. 326-347.
- Hirst, Daniel. 1983. Structures and categories in prosodic representations. In Cutler, Anne & Ladd, Robert (eds.), *Prosody: models & measurement*, 93-109. Berlin: Springer.
- Lacheret-Dujour, Anne. 2001. Modéliser l'intonation d'une langue. Où commence et où s'arrête l'autonomie du modèle? L'exemple du français parlé. *Colloque international: Journées Prosodie 2001*. 57-60.
- Levy, Michael. 1997. *CALL: context and conceptualisation*. Oxford: Oxford University Press.
- Lezhenin, Yuriy & Lamtev, Anton & Dyachkov, Vadim & Boitsova, Elena & Vylegzhanina, Karina & Bogach, Natalia. 2017. Study intonation: mobile environment for prosody teaching. *IEEE International Conference on Cybernetics (CYBCONF)* 3. 1-2.
- Martin, Philippe. 2010. Learning the prosodic structure of a foreign language with a pitch visualizer. *Speech Prosody 2010*, paper 980.
- Munro, Murray J. & Derwing, Tracey M. 1999. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 49. 285-310.
- Nencioni, Giovanni. 1976. Parlato-parlato, parlato-scritto, parlato-recitato. *Strumenti Critici* X(1). 1-56.
- Pereira, Juanan. 2016. Leveraging chatbots to improve self-guided learning through conversational quizzes. In *Proceedings of the fourth international conference on technological ecosystems for enhancing multiculturalism*, TEEM '16, 911-918. ACM Press, New York.
- Romano, Antonio & Giordano, Gelsomina. 2017. Esperienze e riflessioni sulla didattica assistita dell'intonazione in italiano, inglese e francese. In Damascelli Adriana (ed.), *Digital Resources, Creativity, Innovative Methodologies and Plurilingualism: New Approaches to Language Teaching and Learning*, 176-200. Newcastle upon Tyne: Cambridge Scholar Publishing.
- Romano, Antonio & Contini, Michel & Lai, Jean Pierre. 2014. L'Atlas Multimédia Prosodique de l'Espace Roman: uno strumento per lo studio della variazione geoprosoica. In Tosques, Fabio (Hrsg.), *20 Jahre digitale Sprachgeographie*, 27-51. Berlin: Humboldt-Universität-Institut für Romanistik.
- Schiel, Florian. 1999. Automatic phonetic transcription of non-prompted speech. In *Proceedings of the ICPhS 1999*, 607-610.
- Sorianello, Patrizia. 2006. *Prosodia: modelli e ricerca empirica*. Roma: Carocci.
- Trouvain, Jürgen & Gut, Ulrike. 2007. *Non-native prosody. Phonetic description and teaching practice*. Berlin: Mouton de Gruyter.
- Voghera, Miriam. 1989. L'intonazione dell'italiano parlato spontaneo, *The Italianist* 9(1). 116-141.
- Zmarich, Claudio & Magno Caldognetto, Emanuela & Ferrero, Franco. 1996. Analisi confrontativa di parlato spontaneo e letto: fenomeni macroprosodici e indici di fluenza. In Cutugno Francesco (a cura di), *Fonetica e fonologia degli stili dell'italiano parlato. Atti delle VII Giornate di Studio del "Gruppo di Fonetica Sperimentale" dell'Associazione Italiana di Acustica*, 111-139. Roma: Esagrafica.

Appendice

Elenco delle domande e risposte utilizzate:

- 1) Quale fra questi gruppi di opere è formato di sole opere italiane del Novecento?
 - A. L'Orlando furioso, L'Iliade e L'Odissea.
 - B. La divina commedia, Il fu Mattia Pascal e Lo Zibaldone.
 - C. I promessi sposi, La Vita Nuova e La Gerusalemme liberata.
 - D. Uno, nessuno e centomila, La coscienza di Zeno e Se questo è un uomo.
- 2) Chiara ha bisogno di sapere l'ora, come fa?
 - A. Scusi, ho lasciato l'orologio a casa, sa dirmi che ore sono?
 - B. Scusa, ho lasciato l'orologio a casa, per caso ha l'ora?
 - C. Sa dirmi che ora sono per favore che ho lasciato l'orologio a casa?
 - D. Mi scusi, ho lasciato l'orologio a casa, qual è il tempo oggi?
- 3) Marco e Salvatore si sono messi d'accordo per andare a mangiarsi una pizza e Marco deve dare appuntamento a Salvatore, cosa NON gli dirà?
 - A. Ciao Salvatore, allora ci troviamo alle 8 lì davanti?
 - B. Ciao Salvatore, allora ci becchiamo alle 8 davanti alla pizzeria?
 - C. Ciao Salvatore, allora ci incontriamo alle 8 lì?
 - D. Ciao Salvatore, allora ci andiamo a fare una partita a calcio uno di questi giorni?
- 4) Gianna vuole chiedere a Marta di comprare della frutta, cosa le dirà?
 - A. Senti Marta, una domanda: ma a te piace la cassata?
 - B. Senti Marta, stai uscendo? Se riesci a comprare della frutta, mi fai un favore.
 - C. Senti Marta, mi sa che non ci sono più pesche.
 - D. Senti Marta, che tu sappia, c'è ancora frutta in casa?
- 5) In quale anno l'Italia è diventata una Repubblica?
 - A. Nel giugno del 1950.
 - B. Durante l'autunno del 1939.
 - C. Tra il 1945 e il 1946.
 - D. Il 18 marzo 1861.
- 6) In quale giorno è scoppiata la Rivoluzione francese?
 - A. Nell'anno 1789.
 - B. Il 14 luglio 1789.
 - C. Nel marzo 1938.
 - D. Nel 1914.
- 7) L'espressione "avere le mani legate" vuol dire:
 - A. Non riuscire a prendere qualcosa che si trova in alto in uno scaffale.
 - B. Non avere nessuna possibilità di agire come si vorrebbe.
 - C. Avere tempi ristretti e non riuscire ad arrivare in orario a un appuntamento.
 - D. Fare di tutto senza essere all'altezza di una situazione.
- 8) In quale di queste frasi il periodo ipotetico è utilizzato correttamente?
 - A. Se vieni con me, ti portassi a mangiare una pizza.
 - B. Se venissi con me, ti avrei portato a mangiare una pizza.
 - C. Se verresti con me, ti porto a mangiare una pizza.
 - D. Se fossi venuto con me, ti avrei portato a mangiare una pizza.

- 9) Dov'è morto Giuseppe Garibaldi?
- A. Giuseppe Garibaldi è morto nell'isola di Caprera.
 - B. Nel gennaio del 1882.
 - C. Garibaldi è morto all'isola d'Elba.
 - D. Quando aveva 75 anni.
- 10) $1 + \frac{2}{3}$ equivale a:
- A. $\frac{4}{4} + \frac{4}{5}$
 - B. $\frac{1}{6} + \frac{3}{2}$
 - C. $\frac{1}{4} + \frac{4}{9}$
 - D. $\frac{1}{3} - 0,75$