



UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Scienze Politiche e della Comunicazione

Corso di Dottorato in Scienze del Linguaggio, della Società,
della Politica e dell'Educazione

Tesi di Dottorato

Italian Multimodal Corpus:

Verbal and Non-Verbal Communication in Political Domain

Il Tutor:

Prof. *Annibale Elia*

Candidato:

Daniela Trotta

Matr. 8801400078

Il Coordinatore:

Prof. *Filippo Fimiani*

ANNO ACCADEMICO 2021/2022

Abstract

English - Speaker gestures are semantically co-expressive with speech and serve different pragmatic functions to accompany oral modality. Therefore, gestures are an inseparable part of the language system: they may add clarity to discourse, can be employed to facilitate lexical retrieval and retain a turn in conversations, assist in verbalizing semantic content and facilitate speakers in coming up with the words they intend to say. This aspect is particularly relevant in political discourse, where speakers try to apply communication strategies that are both clear and persuasive using verbal and non-verbal cues. This dissertation aims to analyze the co-speech gestures of several Italian politicians during face-to-face interviews using a multimodal linguistic approach. The work first introduces the corpus created: PoliModal corpus (Trotta et al., 2019, 2020), containing the transcripts of 56 TV face-to-face interviews of 14 hours, taken from the Italian political talk show “In mezz’ora in più” (for a total of 100,870 tokens) that has been manually annotated with information about metadata (i.e. tools used for the transcription, link to the interview etc.), pauses (used to mark a pause either between or within utterances), vocal expressions (marking non-lexical expressions such as coughs and semi-lexical expressions such as primary interjections), deletions (false starts, repetitions and truncated words), overlaps and facial displays, hand gestures and body posture. Then, the annotation scheme and the results of a series of statistical analyses aimed at understanding the relationship between annotated multimodal traits and language complexity are described in detail and testing the validity of existing studies on political orientation and language use. Finally, after the presentation of an additional semantic annotation layer related to the function assumed by hand movements, the relationship between them and other information layers such as a political party or non-lexical and semi-lexical tags is investigated. Concerning gesture speech relationship, the results obtained suggest that hand movements are mainly used with integrative and complementary functions. So, the information provided by such gestures adds precision and emphasis to spoken information. Its, also show that party affiliation does not significantly influence the gesture-speech relationship. Furthermore - testing the lexical retrieval hypothesis by calculating the association between the hand movements produced by each respondent and discourse disfluencies using weighted mutual information - it is shown that hand movements tend to co-occur with full pauses (i.e., repetition) and empty pauses (i.e., pause) and more frequently with interjections (semi-lexical tags), suggesting that gesticulation may represent an attempt at lexical retrieval.

Italian - I gesti dei parlanti sono semanticamente co-espressivi con il discorso e svolgono diverse funzioni pragmatiche per accompagnare la modalità orale. Pertanto, i gesti sono una parte inseparabile del sistema linguistico: possono aggiungere chiarezza al discorso, possono essere impiegati per facilitare il recupero lessicale e mantenere il turno nelle conversazioni; aiutano a verbalizzare il contenuto semantico e facilitano i parlanti nel recuperare le parole che intendono pronunciare. Questo aspetto è particolarmente rilevante nei discorsi politici, dove gli oratori cercano di applicare strategie di comunicazione chiare e persuasive utilizzando contemporaneamente elementi verbali e non verbali. Questa tesi si propone di analizzare i gesti co-occorrenti con il parlato di diversi politici italiani durante le interviste faccia a faccia, utilizzando un approccio linguistico multimodale. Il lavoro introduce innanzitutto il corpus creato: PoliModal corpus (Trotta et al., 2019, 2020) composto dalla trascrizione di 56 interviste televisive *face-to-face* tratte dal talk show politico “In mezz’ora in più” (14 ore per un totale di 100,870 tokens) annotato manualmente con informazioni sui metadati (ovvero gli strumenti utilizzati per la trascrizione, il link all’intervista, ecc.), le pause (utilizzate per contrassegnare un silenzio tra o all’interno degli enunciati), le espressioni vocali (contrassegnando espressioni non lessicali come colpi di tosse ed espressioni semi-lessicali come le interiezioni primarie), le cancellazioni (false partenze, ripetizioni e parole troncate), le sovrapposizioni e le espressioni del volto, i movimenti delle mani e la postura del corpo. Viene poi descritto in dettaglio lo schema di annotazione e i risultati di una serie di analisi statistiche volte a comprendere la relazione tra i tratti multimodali annotati e la complessità del linguaggio e a testare la validità degli studi

esistenti sull'orientamento politico e l'uso del linguaggio per la lingua inglese. Infine, dopo la presentazione di un ulteriore livello di annotazione semantica relativo alla funzione assunta dai movimenti delle mani, viene indagata la relazione tra questi e altri livelli informativi come il partito politico o i tag non lessicali e semi-lessicali. Per quanto riguarda la relazione gesto-discorso, i risultati ottenuti suggeriscono che i movimenti delle mani sono utilizzati principalmente con funzioni integrative e complementari; pertanto, le informazioni fornite da questi ultimi aggiungono precisione ed enfasi alle informazioni parlate. Essi, inoltre, mostrano che il partito politico di appartenenza non influenza in modo significativo la relazione gesto-parola. Inoltre - testando l'ipotesi del recupero lessicale attraverso il calcolo dell'associazione tra i movimenti delle mani prodotti da ciascun intervistato e le disfluenze del discorso, utilizzando come metrica la *mutual information* ponderata - viene dimostrato che i movimenti delle mani tendono a co-occorrere con le pause piene (ripetizioni) e le pause vuote (pause) e più frequentemente con le interiezioni (fenomeni semi-lessicali) suggerendo che la gesticolazione può rappresentare un tentativo di recupero lessicale.