

PoliModalCorpus

Verso la costruzione del primo corpus multimodale di dominio politico in
italiano

DANIELA TROTTA, TERESA ALBANESE, ANNIBALE ELIA*

ABSTRACT: This work introduces the PoliModalCorpus, the first multimodal political domain corpus in Italian. The corpus was constructed to fill the lack of Italian linguistic resources for political–institutional communication. The data includes the transcripts of 59 face-to-face interviews in the political talk show “In mezz’ora in più”. This paper illustrates the methodology employed for data collection, the corpus construction, and the annotation scheme proposed to structure the data. A new level of analysis is proposed, which consists in a linguistic and terminological analysis, not only on a quantitative level through textual statistics based on a morpho–syntactic analysis, but by inserting a level of annotation focusing on non-verbal aspects — pauses, non-lexical backchannels — that occur during the political interviews. Non-verbal expressions do not simply accompany attempts to interrupt the speakers but are indicators of the success of their intentions, including persuasive strategies.

Keywords: Political communication, Corpus Linguistics, Multimodal corpora, XML–TEI annotation, Natural Language Processing.

1. Introduzione

Le ricerche italiane di area linguistico–semiologica condotte negli ultimi decenni hanno dedicato un’attenzione sempre maggiore al “linguaggio politico”, espressione con la quale — in accordo con Gensini¹ — si può intendere: a) lo studio degli usi politici della lingua; b) le tecniche discorsive e argomentative finalizzate alla politica (propaganda elettorale, dibattiti fra esponenti di partiti diversi, comizi ecc.); c) i meccanismi mediali della prassi politica, venuti nel massimo risalto con l’avvento della “neotelevisione”² e la crescente integrazione fra “sfera pubblica” e mondo dei media. Da

* Laboratorio di Linguistica Computazionale “M. Gross”, Dipartimento di Scienze Politiche, Sociali e della Comunicazione, Università degli Studi di Salerno, Fisciano (SA), Italia.
d.trotta9@studenti.unisa.it; t.albanese@studenti.unisa.it; elia@unisa.it

1. S. GENSINI, *Comunicazione politica*, in «Linguaggio e comunicazione politica negli studi linguistici italiani: bilancio e temi attuali.» n. 1, 2009, pp. 85–98.

2. U. ECO, *Tv: la trasparenza perduta*, in *Sette anni di desiderio*, 1983, pp. 163–79.