

RISORSE LINGUISTICHE PER LO STUDIO DEI *PROMESSI SPOSI*



COSA E PERCHÉ

- Sviluppo in corso di nuove risorse linguistiche per l'analisi dei *Promessi Sposi*
- Finalità didattiche e di ricerca
- Creazione di *dataset* e strumenti per:
 - linguisti e storici della letteratura
 - ricercatori nel campo dell'NLP
 - docenti e studenti di scuole superiori

ANNOTAZIONE LINGUISTICA

Annotazione morfogrammaticale

- Annotazione manuale di 3 capitoli (I, VIII e XXIII) nelle 3 redazioni dei *Promessi Sposi* con lemmi, PoS, *features* (tratti morfologici e flessivi, cui si aggiungono tratti specifici sull'ultima colonna del formato CoNLL-U):
 - correzione su foglio di calcolo dell'*output* del modello ISDT di UDPipe
 - addestramento di nuovi modelli

1	«	«	PUNCT	FB	-	-	-	-
2	Ah	ah	INTJ	I	-	-	-	-
3	,	,	PUNCT	FF	-	-	-	-
4	non	non	ADV	BN	PronType=Neg	-	-	-
5	perdiam	perdere	VERB	V	Mood=Ind Number=Plur Person=1 Tense=Pres VerbForm=Fin	-	-	Variant=Apoc
6	tempo	tempo	NOUN	S	Gender=Masc Number=Sing	-	-	-
7	!	!	PUNCT	FS	-	-	-	-
8	»	»	PUNCT	FB	-	-	-	-
9	sciamò	esciamare	VERB	V	Mood=Ind Number=Sing Person=3 Tense=Past VerbForm=Fin	-	-	sciamare Style=Arch
10	Federigo	Federigo	PRON	SP	-	-	-	-

Annotazione delle emozioni

- *Dataset* di dialoghi e soliloqui tratti dai capitoli annotati nelle diverse redazioni del romanzo:
 - annotazione manuale delle emozioni di base di Plutchik assegnate a livello di frase
 - addestramento di nuovi modelli

ALLINEAMENTO TRADUZIONI

Chi le darà un parere?...»	Who will give you advice?"
	"Oh! peace, peace!
«Ohimè! tacete, e non apparecchiate altro: datemi un bicchiere del mio vino.»	Do not make matters worse.
	Give me a glass of my wine."

Sentence alignment

- Creazione di un corpus parallelo costituito dalla Ventisettana e dalla sua traduzione inglese del 1834:
 - allineamento semi-automatico con LF Aligner e correzione manuale dell'*output*
 - relazioni 1:1, 1:0, 1:N, N:1, *crossing*

Word alignment

- Allestimento di un *dataset* formato da alcuni capitoli della Ventisettana allineati a livello di parola con la corrispondente traduzione inglese del 1834
 - allineamento manuale con Ugarit

LAVORI FUTURI

- Estendere l'allineamento a livello di frase ad altre traduzioni del romanzo (sia Ventisettana che Quarantana)
- Valutare strumenti automatici per il *word alignment* come fast_align e AWESOME
- Annotare altri capitoli dei *Promessi Sposi*
- Estendere l'annotazione al livello sintattico
- Rendere disponibili le risorse sul portale *ManzoniOnline* e su ILC4CLARIN

Per informazioni:

{rachele.sprugnoli,arianna.redaelli, marco.sartor}@unipr.it



UNIVERSITÀ DI PARMA



LABORATORIO DI INFORMATICA UMANISTICA E CULTURA DIGITALE

