



*Ministero dell'Università e della Ricerca*

Segretariato Generale  
Direzione generale della ricerca  
Ufficio III

**Relazione scientifica integrativa – Prima annualità**

**PRIN 2017 – codice progetto 2017STJCE9**

**Settore ERC SH4 – Linea A**

Principal Investigator

**LOMBARDI VALLAURI Edoardo**

Università degli Studi ROMA TRE

**Aspetti da chiarire sulla base degli evidenti disallineamenti riscontrati dal CDS in sede di valutazione della relazione scientifica intermedia**

*(Descrizione dettagliata e approfondita che chiarisca i punti di debolezza rilevati dal CDS ed evidenzi le motivazioni dei disallineamenti rispetto agli obiettivi scientifici previsti dal progetto.*

*Descrizione attività che si intendono svolgere per la prosecuzione del progetto al fine di recuperare gli scostamenti segnalati).*

La presente relazione integrativa risponde prima ai rilievi sollevati dal CDS riguardo allo stato di avanzamento del progetto, e poi elenca per intero, integrando quanto fatto nella precedente relazione intermedia, le pubblicazioni scientifiche e gli interventi in sedi scientifiche, inerenti all'oggetto del progetto e prodotti da membri del progetto stesso.

27 giugno 2021

---

**STATO DI AVANZAMENTO DEL PROGETTO**

Senza dubbio la mia relazione scientifica intermedia era redatta in modo fuorviante. Avevo male interpretato ciò che doveva contenere, credendo che fosse chiesto di riferire sugli effettivi risultati progressi alla data di gennaio (quindi anche di giustificare gli effetti della pandemia sul progetto), e non di esporre le misure adottate per il futuro. Di conseguenza l'avevo formulata in modo eccessivamente sommario, e anche proprio ignorando che alcune ulteriori informazioni erano richieste (specialmente contromisure alla pandemia e pubblicazioni scientifiche e non solo divulgative in corso, oltre a quelle già pubblicate). Mi dispiace molto di avere generato legittime preoccupazioni che, come spero si vedrà nel seguito di questa relazione integrativa, non sono necessarie.

## RACCOLTA, ALLESTIMENTO E ANNOTAZIONE DEL CORPUS

Nella relazione intermedia si specifica che la consistente riduzione del finanziamento per il personale e i contratti operata dal Ministero rispetto alla richiesta ha comportato una (meno significativa) rimodulazione del numero di discorsi del corpus. Non so se questo sia stato considerato dal CDS un disallineamento, ma già la relazione intermedia specifica che lo si è compensato con una maggiore lunghezza media dei discorsi analizzati, per cui il corpus, nonostante la riduzione di budget, finirà per constare di un numero di token sostanzialmente pari alle previsioni.

Quanto ai tempi di realizzazione dell'annotazione, a scanso di eventuali equivoci si precisa anzitutto che nella relazione intermedia è indicato che dopo il primo anno era stato annotato *poco meno di un terzo del corpus*, cioè che nel primo anno era stato annotato poco meno di un terzo del totale da annotare *in tre anni*, e non un terzo di ciò che era da annotare nel primo anno. Questo significa che, come esposto nella relazione intermedia, nonostante le difficoltà iniziali la raccolta e annotazione del corpus non ha mai sofferto di un grave ritardo. Anzi, sta procedendo con il ritmo previsto, ed è ormai vicina a raggiungere la metà del totale da annotare.

La Relazione Periodica del CDS osserva giustamente che, stando alla relazione intermedia, "l'attività permane poco visibile". Il lavoro centrale del progetto, cioè l'annotazione del corpus di discorsi politici per le categorie pragmatiche degli impliciti, non può essere in sé e per sé "visibile" prima che sia terminato, trattandosi di un grande corpus da trascrivere e annotare prima, e pubblicare poi (come web resource); ma procede regolarmente e secondo le previsioni, con un ritardo modesto. Questo ritardo era dovuto alla iniziale difficoltà di realizzare alcuni passaggi nelle fasi di insorgenza della pandemia: essenzialmente i contatti con istanze esterne fornitrici di materiale audiovisuale, l'attività di coordinamento interno dei vari gruppi del progetto, e quella di analisi congiunta di ciascun discorso politico da parte degli annotatori in parallelo. Tali problemi sono risolti da tempo mediante la messa a punto di procedure online, e anche se il modesto ritardo iniziale da pandemia non è facile da recuperare interamente in tempi brevi (e per questo la relazione intermedia lo segnalava), con la proroga espressamente prevista e concessa dal Ministero non esiste alcun motivo di temere sui tempi di conclusione del progetto.

Se la pandemia dovesse continuare, per quanto riguarda specificamente l'annotazione coordinata del corpus si sono ormai messe a punto le seguenti contromisure, che sono del tutto operative e quindi possono rimanerle per tutto il tempo necessario al completamento del lavoro:

- Gli accordi con i soggetti da cui vengono ottenuti i materiali audiovisivi dei discorsi politici sono da tempo stabilizzati, e si sono sviluppate le procedure per ricevere da essi i materiali interamente online. Queste procedure si applicano senza intoppi e possono continuare così, sia che perduri la pandemia, sia che essa cessi.
- Le procedure di coordinamento degli e fra gli annotatori del corpus si svolgono interamente online.
- L'utilizzo della piattaforma WebAnno permette l'annotazione simultanea dei testi in rete da parte di un numero a piacere di annotatori, e consente la collazione dei risultati delle annotazioni parallele da parte del revisore finale, anche questa in rete. Lo stesso vale del calcolo dell'inter-rate agreement.

Tutto questo fa sì che la pandemia, anche se perdurasse fino alla fine del progetto, non possa più causare ritardi.

Anche la messa a punto degli aspetti informatici della web resource procede senza problemi, e anzi - non avendo avuto a soffrire della pandemia - è in anticipo (si veda anche la stipula del contratto di cui alla relazione intermedia), ma non ha potuto dare luogo a risultati "visibili" perché, come previsto da tabella di marcia / workflow allegato alla domanda di finanziamento, solo in questi mesi (successivi alla presentazione della relazione intermedia) si sono cominciate a

elaborare mediante lo strumento informatico in via sperimentale le prime porzioni di corpus annotato. I risultati saranno disponibili, in senso proprio, solo a conclusione del progetto, ma nel frattempo verranno resi noti nelle pubblicazioni in preparazione e in comunicazioni a congressi (vedi sotto).

Non si dubita dunque di portare a termine il progetto come originariamente indicato nella domanda di finanziamento del progetto stesso, anche se utilizzando la proroga concessa; al netto della quale la tabella di marcia allegata alla domanda iniziale è pienamente rispettata.

#### ESPERIMENTI NEUROFISIOLOGICI

Rispetto alla raccolta del corpus, maggiore è stato l'ostacolo iniziale incontrato dal primo esperimento neurofisiologico a causa della pandemia. Però a questo segmento del progetto si è nel frattempo (nel 2021) impresso un ritmo accelerato per ridurre il ritardo, tanto che l'esperimento si è concluso, e i risultati saranno resi noti, per cominciare, già nel congresso *Xprag.it 2021* (online, 8-9 luglio 2021) e in quello *ESPP2021* (Leipzig, 30 agosto - 2 settembre 2021) (si veda la lista delle comunicazioni in congressi), cui seguirà la pubblicazione su rivista appropriata (inizialmente e prevedibilmente, *Neurocomputing o Language and Cognition*, dove abbiamo già pubblicato studi su esperimenti analoghi). Il secondo esperimento è già iniziato, con un ritardo ormai ben minore, e la proroga concessa a causa della pandemia rimedia pienamente a tale residuo ritardo. Sarebbe possibile ridurlo ulteriormente, ma procedendo in un modo frettoloso che non gioverebbe alla piena accuratezza delle procedure sperimentali di cui si tratta. Per questo motivo si è chiesta la proroga giustamente concessa dal Ministero in ragione della pandemia, e si ritiene più serio valorizzarne l'assai opportuna disponibilità. In ogni caso, per proseguire senza ulteriori ostacoli anche in caso di perdurare delle limitazioni da corona virus, si sono ormai prese le seguenti contromisure, alcune delle quali rese possibili dalle nuove e più efficaci norme anti-covid adottate nelle sedi di Roma Tre a partire dalla fine del 2020:

- Il reperimento dei soggetti sperimentali (prevalentemente studenti) è stato spostato prevalentemente sulle piattaforme informatiche con cui si svolgono i corsi universitari online;
- Le procedure di acquisizione dei dati neurofisiologici sono state riformulate secondo i protocolli divenuti vigenti dopo la prima fase di incertezza dell'ateneo (Roma Tre), che prevedono:
  - distanziamento in locali di sufficienti dimensioni;
  - sanizzazione dei locali e degli strumenti dopo ogni acquisizione;
  - turnazioni degli sperimentatori e dei soggetti sperimentali, sempre presenti solo individualmente e su un calendario di appuntamenti, ma in numeri giornalieri paragonabili a quelli possibili prima della pandemia.

Con queste procedure, ormai implementate e quotidianamente applicate da diversi mesi, anche in eventuale costante e perdurante regime di pandemia sarà possibile procedere nei tempi previsti dal progetto.

#### DIVULGAZIONE SCIENTIFICA DEL PROGETTO

Anche se il corpus finché è in corso di realizzazione non consente di pubblicare i veri e propri risultati dell'intero progetto, la visibilità scientifica di questa parte del lavoro consiste in questa fase nella pubblicazione di lavori che illustrano il quadro teorico e i criteri operativi sviluppati dalla direzione e dai membri del progetto, e applicati alla realizzazione del corpus stesso; oltre che, in prospettiva ancora provvisoria, alcuni risultati parziali e alcune fasi della sua stessa realizzazione. Questo è avvenuto in numerose pubblicazioni scientifiche, significativamente ulteriori rispetto a