

HOW BAD IS IT?

Lavoro gruppo 6

Riferimenti teorici

Lo studio dal quale siamo partiti si basa sull'approccio teorico di Bachmann (1990) per il quale la competenza comunicativa di chi impara una lingua può essere divisa in tre tipologie:

- Competenza socio-linguistica
- Competenza socio-culturale
- Competenza pragmatica

In particolare ci siamo soffermati sull'aspetto pragmatico.

Riferimenti teorici

Bachmann definisce la competenza pragmatica come la conoscenza delle condizioni, norme e massime che regolano il discorso, gli atti del parlato e le conversazioni. Per esempio, il chiedere scusa degli inglesi non viene fatto sotto le stesse situazioni sociali come il chiedere scusa degli spagnoli.

Domanda di ricerca

Come la comprensione pragmatica della gravità delle situazioni è influenzata dalla lingua d'origine?

Nel nostro lavoro abbiamo voluto indagare se le stesse differenze fossero riscontrabili anche tra la lingua italiana e quella inglese.

Metodi e materiali

Siamo partiti dal questionario trovato in uno studio di una tesi magistrale, abbiamo poi costruito due set di item differenti per verificare la replicabilità di tale studio declinato ad hoc.

Come strumento abbiamo utilizzato Google Form.

Metodi e materiali

I questionari hanno l'intento di osservare se esistono differenze culturali nell'interpretare la gravità delle situazioni che vengono presentate attraverso gli item.

Per fare questo abbiamo utilizzato l'inglese come lingua comune in un campione di soggetti madrelingua inglese e italiana, e una scala a 5 livelli che variano su un continuum da *not severe* a *severe* .

Esempio di item

You are eating an icecream in a crowded place and it falls on a stranger's shoe

Level of offence

1. (not severe)

2.

3. (somewhat severe)

4.

5. (severe)

Metodi e materiali

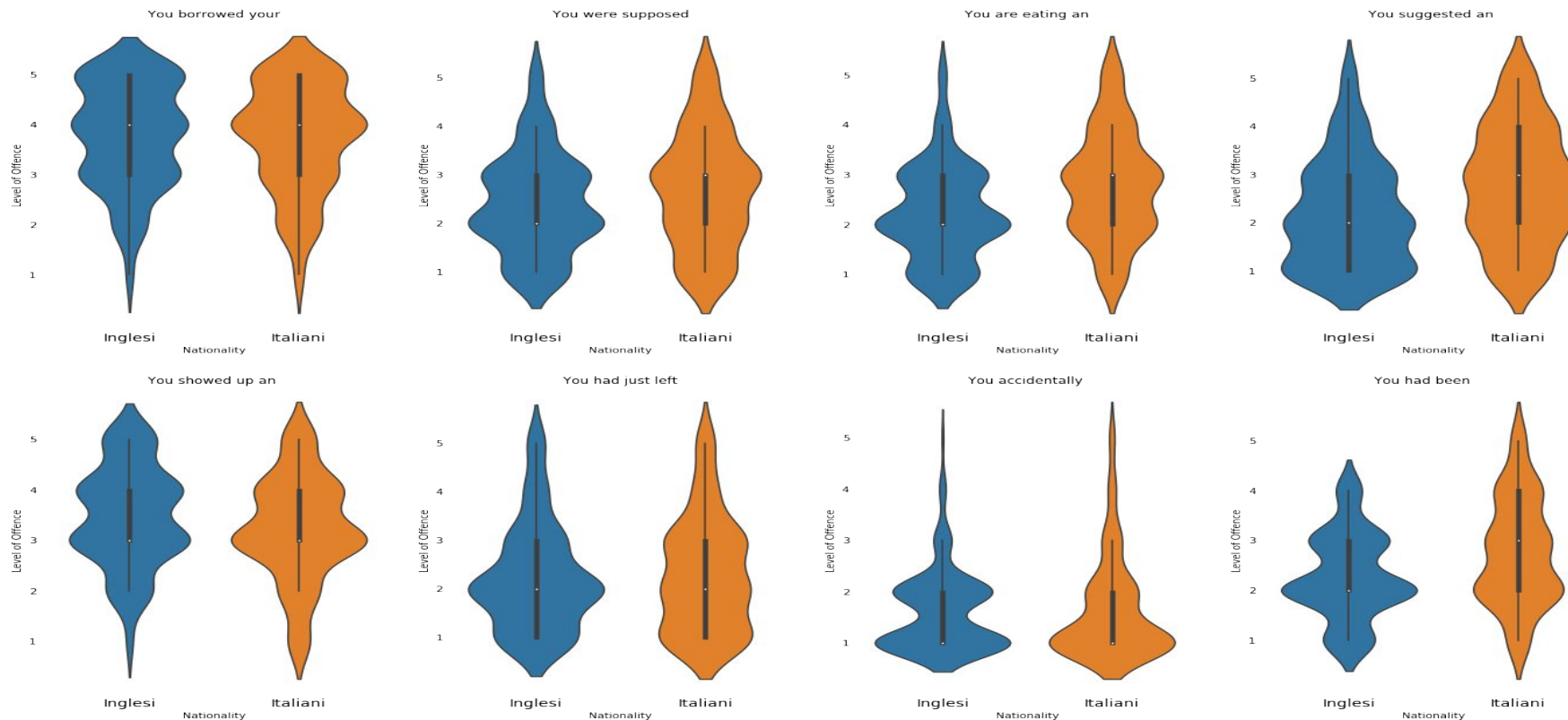
Siamo partiti dal questionario trovato in uno studio di una tesi magistrale, abbiamo poi costruito due set di item differenti per verificare la replicabilità del primo studio

Raccolta dati

La modalità utilizzata per raccogliere i dati è stata attraverso link mandato attraverso WhatsApp e Reddit. Abbiamo riscontrato sostanziali differenze tra gruppo A e B, in termini di numerosità del campione e composizione.

Abbiamo effettuato le analisi solo sui dati del gruppo A, campionando gli inglesi in modo da avere gruppi omogenei e rimuovendo i soggetti 'Altro'.

Distribuzione dei dati



Distribuzione dei dati

Effettuando dei t test abbiamo evidenziato una differenza significativa tra gli Italiani e gli Inglesi nella media della media delle risposte

t value: 3.0291

p value: 0.0026

Media della media delle risposte degli italiani: 2.7422

Media della media delle risposte degli inglesi: 2.5476

Distribuzione dei dati

Inoltre, confrontando le singole domande per media delle risposte degli inglesi e degli italiani, i t test hanno mostrato una differenza significativa per le domande :

- *You were supposed to go to a friend's graduation party but forgot to do so because you had a job interview.*
- *You are eating an ice cream in a crowded place and it falls on a stranger's shoe*
- *You had been invited to a dinner party and you forgot to take with you the bottle of wine that you have promised to bring*
- *You suggested an answer to a friend during an exam and you found out later that it was wrong*

Modello

Abbiamo definito un modello di regressione logistica addestrato attraverso una pipeline con 10-fold cross validation.

L'accuratezza media è risultata essere di 60% circa.

I coefficienti delle singole variabili sono maggiori per le variabili individuate come significative dal t-test

Abbiamo poi confrontato il modello con un algoritmo 'K-Nearest Neighbours', che ha mostrato un'accuratezza leggermente inferiore.

Modello

Colpiti dall'inaspettata incongruenza tra l'alta significatività dei t-test e la relativamente bassa accuratezza predittiva del modello, abbiamo costruito una rete neurale.

L'accuratezza del modello è in media 60%

La rete è costituita da 3 hidden layers completamente connesse con 4 nodi ciascuna, con regolarizzazione L2 e un tasso di Dropout di 0.5 per ogni layer.

Limiti

- Comprensione della lingua inglese
- Influenza del contesto
- Second language affects moral stance, la traduzione occupa tante energie cognitive modificando la risposta dei non nativi.

Conclusioni

Osserviamo quindi differenze etniche nelle risposte, ma gli algoritmi che abbiamo utilizzato non riescono a coglierla appieno.

È rassicurante che una linea non possa dividere così facilmente Italiani e Inglesi.

Siamo anche contenti di aver trovato differenze significative nelle domande create da noi.

GRAZIE PER
L'ATTENZIONE!