

GIORGINA CANTALINI

Corpus multimodale annotato per lo studio della gestualità co-verbale nel «parlato-parlato» e nel «parlato-recitato»

Il corpus in oggetto studia le relazioni della gestualità co-verbale con i diversi livelli di organizzazione linguistica, in particolare la strutturazione prosodica. È composto da campioni comparabili di interviste fatte ad attori di prosa, parallelamente registrati nell'esecuzione di uno stesso brano teatrale. Offre un modello di metodologia di annotazione in cui i quadri teorici utilizzati, di approccio configurazionale, mostrano caratteristiche di compatibilità e sovrapponibilità funzionalmente efficaci; inoltre comparando due varietà (spontaneo e recitato) e due modalità (orale e gestuale) e sfruttando l'opportunità doppia di confronto restituisce risultati significativi sia riguardo alla sincronizzazione tra modalità che alla variazione di strategie di modellizzazione tra varietà. Il tipo di analisi proposta consente dunque di approfondire lo studio della relazione della gestualità con il parlato orale. Consente inoltre avanzamenti nello studio della semantica del gesto.

Parole chiave: annotazione multilivello, corpora multimodali, gesto, prosodia, parlato recitato, parlato spontaneo, variazione diamesica, variazione modale.

1. Introduzione

Il corpus multimodale di parlato spontaneo e parlato recitato, di cui al demo¹ presentato al LIV Convegno Internazionale di Studi della SLI e del quale questo contributo scritto costituisce un'integrazione, è dedicato allo studio della gestualità co-verbale (a partire da Kendon 1972, 1980, 2004; McNeill 1992, 2005) e alle sue relazioni con i diversi livelli di organizzazione linguistica, in particolare la strutturazione prosodica (Loehr 2004; Cresti 2000).

¹ <https://doi.org/10.48448/3k4n-hg49>

Alla base dello studio condotto vi sono state sia motivazioni individuali che di interesse più generale. In quanto didatta della recitazione e attrice io stessa, mi occupo da tempo di parlato-scritto (Nencioni 1976) con l'obiettivo di farlo risultare parlato-parlato: una pagina scritta, in cui il linguaggio sia già stato espresso e depositato su un supporto cartaceo, deve essere eseguita simulandone condizioni di pronunciamento affatto differenti e invece proprie del parlato spontaneo (su tutto: il processo di ideazione deve svolgersi contemporaneo a quello di locuzione sotto la spinta di una reale interazione comunicativa). Numerose osservazioni empiriche negli anni mi hanno portato a constatare come tale risultato fosse ottenibile solo grazie a uno specifico coinvolgimento fisico-dinamico di tutto il corpo, che è cominciato ad apparirmi, negli anni, sistematico e necessario non solo per l'*espressione*, con riferimento alle performance attoriali o oratorie, ma anche per la *comprensione* del testo in quanto pagina scritta. Dunque, se per far funzionare un testo scritto oralmente non bastava solo la voce, ma era necessario anche il movimento, con riguardo particolare ai movimenti gesticolatori, la questione diventava capire in che maniera il corpo fosse coinvolto nel linguaggio, strutturalmente, tipologicamente e funzionalmente (quali gesti, quali spinte dinamiche, quale rapporto con la gravità e in che maniera voce e gesto fossero in relazione). In quanto linguista invece, l'interesse era rivolto allo studio del linguaggio e alla crescente attenzione che la letteratura attuale sta dando alla costituzione di corpora multimodali annotati di linguaggio e alla formazione di modelli di annotazione del gesto. Da una parte è ormai assodato che all'interno del variegato e complesso excursus di manifestazioni corporali accomunate dalla definizione di comportamento non-verbale (Ekman & Friesen 1969), o di linguaggio del corpo (Maricchiuolo 2017), è possibile individuare una tipologia specifica di gestualità considerata costitutiva e complementare al linguaggio parlato (McNeill 1985), che occorre co-verbalmente, alternativamente o in sostituzione ad esso², e che come tale non può più essere ignorata nello studio di quest'ultimo. D'altro canto ci si trova in un momento storicamente favorevole all'osservazione della gestualità, contrariamente a quanto avvenuto per la gran parte del secondo scorso, grazie alla diffusione di strumentazione audiovideo di alto livello, facile utilizzo e co-

² Si veda la classificazione "Kendon's continuum" in McNeill 1992.

sti accessibili. In questo quadro di domanda individuale e opportunità generale è da collocarsi dunque il lavoro qui esposto.

Saranno presentati la tipologia di dato acquisito e i criteri operativi adottati per la sua acquisizione (2.); i modelli teorici presi a riferimento (3.); la metodologia di annotazione, nello specifico la strutturazione di quali e quanti livelli predisposti e le modalità di segmentazione del gesto e della prosodia utilizzate (4.); aspetti relativi alla validazione (5.); evidenze osservate e produttività del modello di annotazione approntato (6.); infine saranno riportate in breve alcune conclusioni (7.).³

2. Tipologia di dato acquisito e criteri operativi di acquisizione

Per ottenere materiale comparabile di due diverse varietà linguistiche si è scelto di far recitare a tre attori di prosa, uomini, italiani madrelingua, il medesimo pezzo teatrale prima e poi di sottoporli ad un'intervista strutturata. Il pezzo era tratto da "Il giuoco delle parti" di Luigi Pirandello⁴. Due degli attori sono stati registrati in uno studio di posa e hanno recitato il brano espressamente per la ricerca; il terzo è stato registrato durante lo spettacolo dal vivo dell'intera commedia di Pirandello in scena proprio nel periodo in cui è stata condotto il presente lavoro; e alle tre interpretazioni ottenute si è aggiunta quella di un quarto attore dalle Teche Rai. Si sono collezionati così circa 8 minuti di varietà recitata, in cui il personaggio principale, pur trattandosi di un dialogo, prende la parola e la mantiene quasi in esclusiva. I tre attori coinvolti sono stati poi sottoposti all'intervista, che ha riguardato il loro lavoro e il loro approccio alle battute di un copione, in particolare alla scrittura di Pirandello. Anche in questo caso, nonostante la natura dialogante e il contesto conversazionale, in linea con gli studi precedenti (p.es., Loehr 2004), dai circa 30 minuti complessivi registrati ne sono stati estrapolati 10 in totale in cui per ciascun parlante fossero compresi: una sequenza di flusso verbale di

³ Per un approfondimento di processi, metodi e risultati si rimanda a Cantalini 2018; Cantalini et al. 2020; Cantalini & Moneglia 2020; Cantalini & Moneglia 2022.

⁴ Luigi Pirandello: "Il giuoco delle parti" (1918. Mondadori: 1993 – Atto I, Scena Terza.): si tratta della scena in cui il protagonista, Leone Gala, discorre dei casi della vita utilizzando la metafora dell'uovo e del suo guscio: in breve il secondo (il guscio svuotato) è l'auspicabile per quanto cinicamente indifferente atteggiamento da assumere nella vita di contro alla forza distruttiva delle passioni (l'uovo dentro il guscio).

almeno 30 secondi consecutivi (di fatto ogni turno registrato è stato mediamente più lungo); la porzione analizzata non avesse tagli al suo interno ed esponesse almeno una risposta completa ad una delle domande dell'intervista strutturata (abbiamo estrapolato in realtà una media di ben due turni per parlante); ci fosse evidente coinvolgimento gestuale nelle selezioni estrapolate; e infine il materiale prodotto fosse prevalentemente di tipo monologico. In questo modo la comparabilità delle registrazioni è stata data dall'identità testuale nel recitato, da quella individuale tra spontaneo e recitato, dalla tipologia di contesto comunicativo per lo spontaneo; e per entrambe le varietà si sono selezionate porzioni testuali monologiche interne a dialoghi. Ai parlanti è stato detto che sarebbero stati osservati da un punto di vista 'linguistico' e 'comunicativo', con ciò intendendo che non avremmo in nessun modo affrontato o valutato aspetti di natura artistica, bensì aspetti definibili come 'grammaticali'. In questo modo si è evitato di menzionare la gestualità per evitare che ciò influenzasse il loro comportamento e al contempo li si è rassicurati sul fatto che in nessun modo la nostra ricerca avrebbe riguardato la loro qualità di attori, così da condizionarne il meno possibile l'interpretazione.

L'elicitazione di dati linguistici multimediali che offrano un flusso conversazionale il più naturale possibile⁵ è questione metodologica delicata che richiede particolari accorgimenti, come già testimoniato in letteratura. Molti e diversi sono stati gli espedienti messi in atto per ottenerlo: la mediazione di un secondo interlocutore che non fosse l'intervistatore nel caso di interviste a singoli⁶; predisposizione di set in cui i parlanti fossero composti da gruppi di amici⁷; riprese effettuate in luoghi pubblici e familiari per gli intervistati⁸. Anche gli argomenti scelti di cui parlare possono favorire un flusso non trattenuto: parlare del proprio lavoro di attori come nel nostro caso è stato sicuramente stimolante per gli intervistati; Loehr (2004) ha lasciato che la conversazione tra più parlanti coinvolti, tra di loro amici, prendesse piede da sola, estrapolandone i momenti di maggior coinvolgimento gestuale. Tutto ciò perché, è innegabile, la sola telecamera, così come la relazione con un interlocutore sconosciuto e investito del ruolo di

⁵ "conversation flow as natural as possible" in Loehr 2004: 72.

⁶ McNeill 2005.

⁷ Loehr 2004.

⁸ Kendon 1972.

‘osservatore esterno’, inibisce e condiziona il comportamento di un parlante, influenzando e bloccandone soprattutto la gestualità⁹. E’ perciò importante dare tempo alla conversazione di svilupparsi, così da permettere ai parlanti di dimenticarsi della presenza della telecamera, e di intervistarli assecondandone i momenti in cui appaiono particolarmente coinvolti in quanto stanno dicendo. In queste specifiche condizioni di acquisita familiarità rispetto al contesto estraneo e di investimento nel proprio pronunciamento, la gesticolazione non è mai davvero trattenuta ed è possibile osservarla in quanto naturalmente utilizzata. Inoltre è bene far sedere gli intervistati su sedie con spalliere, ma senza braccioli: in assenza di appoggi cui aggrapparsi per ‘tenere sotto controllo la gesticolazione’ le braccia inevitabilmente inizieranno ad accompagnare in maniera spontanea l’elocuzione¹⁰. Ed è bene che non siano fatti sedere su poltrone perché la loro schiena resti libera di muoversi, non affondando all’indietro. L’excursus gestuale, nella nostra osservazione, arriva alle mani partendo dal torso ed espandendosi lungo le braccia. Il modo di far star seduto il parlante, con i succitati accorgimenti, può favorire pertanto a nostro avviso sia l’ampiezza del gesticolare (e con l’ampiezza una più compiuta formalizzazione del gesto favorendone l’interpretazione), che il completo rilasciamento del movimento gestuale.

3. *Modelli teorici*

Per annotare il dato raccolto ci si è mossi all’interno di due quadri teorici, per la prosodia e per il gesto, teoricamente e proceduralmente comparabili. In realtà la scelta del modello prosodico da utilizzare in relazione all’analisi gestuale è questione cruciale e variamente affrontata in letteratura. Se il modello di struttura dell’excursus gestuale primariamente individuata da Kendon¹¹ è stata di fatto confermata negli studi successivi e non ha subito sostanziali modifiche, i quadri teorici prosodici invece non sono unitari né nell’approccio né nella descrizione

⁹ È tuttora culturalmente diffuso il concetto che gesticolare quando si parli sia sbagliato e diastraticamente di basso livello.

¹⁰ Purtroppo con il parlante “U”, attore di chiara fama intervistato nel suo camerino, è stato impossibile chiedergli di tenere completamente staccate le braccia dagli appoggi intorno a lui.

¹¹ Kendon 1972.

ne. D'altro canto gli studi sul gesto sono stati per lungo tempo ostacolati dall'assenza di strumentazione adeguata, mentre quelli prosodico fonetici sicuramente si sono avvalsi molto presto dei più diffusi strumenti di registrazione solo sonora. Così da una parte il gesto ha fornito da subito un modello standard di analisi, ma non le opportunità per la sua indagine; dall'altra il parlato è stato oggetto di approfondite indagini, ma in assenza di uno standard fonetico di annotazione condiviso e universale. Tutto ciò fa sì che ad oggi gli studi a disposizione sulla sincronizzazione gesto-parlato non siano realmente sovrapponibili per quanto riguarda le unità linguistiche prese a riferimento¹².

Stante l'eterogeneità dei lavori precedenti, la nostra scelta è stata quella di mettere insieme il modello gestuale di Kita et al.¹³, che sistematizza le fasi del movimento gestuale, già identificate in unità configurazionali, secondo un sistema di regole sintagmatiche all'interno di una gerarchia a inclusione, e affiancare ad esso un modello teorico prosodico analogo per approccio descrittivo e procedurale: il modello L-AcT¹⁴, anch'esso configurazionale, che postula unità linguistiche percettivamente riconoscibili e altrettanto gerarchicamente organizzate a inclusione. In particolare, in entrambi i modelli si osserva un segmento nucleare necessario e sufficiente costituito da un picco d'intensità (acustica o cinetica) intorno al quale si dispongono facoltativamente ulteriori unità strutturalmente differenti, nell'insieme delle quali, unità facoltative più picco, sono articolate le unità superiori.

Nel modello gestuale il nucleo realizza la fase semioticamente attiva del gesto ed è denominato *stroke*. Prima e dopo lo *stroke* è possibile trovare fasi di tenuta o *hold* (per esteso, *pre-stroke hold* se prima dello *stroke*; e *post-stroke hold* se dopo¹⁵) che ne estendono la portata espressiva, realizzando un segmento composito invece che singolo denominato *expressive phase*. Prima e dopo il solo *stroke* o la più composita *expressive-phase*, possono occorrere le fasi di preparazione (*preparation*) e ritrazione (*retraction*), le cui denominazioni indicano i segmenti di preparazione della mano a realizzare il picco o di successivo e progressivo rilassamento dopo averlo realizzato. Con *fase* si

¹² Kendon 1972, 1980; McClave 1981; Loehr 2004; Turk 2020.

¹³ Kita et al. 1998.

¹⁴ Language into Act Theory: Cresti 2000; Cresti & Moneglia 2018.

¹⁵ L'*independent hold* è invece alternativo allo *stroke* (Kita et al. 1998: 28).

indica quindi per il gesto l'unità discreta più piccola, potenzialmente separabile, della gerarchia¹⁶.

L'intera escursione del gesto intorno al picco d'intensità, sia se realizzato da un singolo stroke che articolato in più o tutte le possibili estensioni date dalle fasi menzionate, costituisce il livello superiore della gerarchia o sintagma gestuale (*gesture phrase*). Il sintagma a sua volta può essere contenuto in un'unità superiore variamente articolata, l'unità gestuale (*gesture unit*), che si realizza ogni qualvolta la mano lascia una superficie di appoggio (*home position*) per poi ritornarvi. Una volta appoggiata, ma non ancora completamente ferma, la mano è nella fase di riposo (*rest position*), considerata nel presente lavoro unità attiva e come tale inserita nella gerarchia. A livello percettivo quindi il *gesture phrase* è riconoscibile grazie al suo centro (il culmine del movimento), mentre la *gesture unit* lo è grazie alla chiara individuazione delle sue estremità (insorgenza del movimento e sua chiusura). I tre livelli possono articolarsi differenziati o coincidere (per esempio: uno stroke che parta dalla superficie d'appoggio e vi ritorni subito dopo)¹⁷.

A partire dal modello l'annotazione del gesto procede allora predisponendo sul foglio informatico del software Elan¹⁸ (Wittenburg *et al.* 2006) organizzato in griglia a partire dalle indicazioni date da LASG¹⁹ (Bressem *et al.* 2013), tre righe sovrapposte e concomitanti, in ciascuna delle quali segnare inizio e fine delle unità sovradescritte: al livello superiore iniziando dalle *gesture unit*, per poi dettagliare al secondo i *gesture phrase*, e infine riportando al terzo le *gesture phase*.

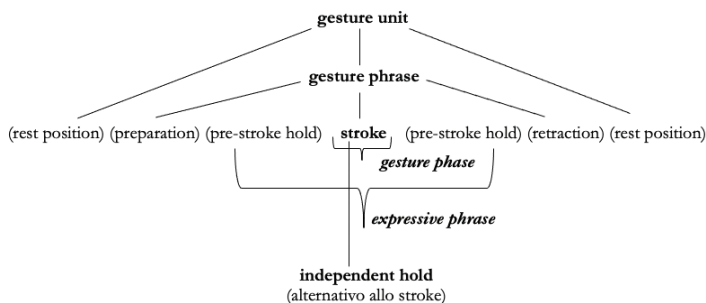
L'annotazione di inizio e fine unità (*break*) è alla base anche del modello linguistico-prosodico L-AcT. Postulando la corrispondenza tra unità prosodiche e unità linguistiche, L-AcT procede su due livelli individuando e annotando i *break prosodici terminali*, e successivamente al loro interno i *break prosodici* percepiti come non *terminali*.

¹⁶ Bressem *et al.* 2013.

¹⁷ Figura 1.

¹⁸ Si tratta di uno strumento professionale di annotazione e trascrizione di registrazioni audio o video manuali o semiautomatiche,

¹⁹ Language Annotation System for Gestures.

Figura 1 - *La struttura del gesto*²⁰

da Kita, van Gijn & van der Hulst 1998

L'annotazione di inizio e fine unità (*break*) è alla base del modello linguistico-prosodico L-AcT. Postulando la corrispondenza tra unità prosodiche e unità linguistiche, L-AcT procede su due livelli individuando e annotando i *break prosodici terminali*, e successivamente al loro interno i *break prosodici* percepiti come non *terminali*. Una sequenza conclusa da un *break* terminale (*pattern prosodico*) corrisponde a una unità di riferimento autonoma dal punto di vista pragmatico (*enunciato*); le unità delimitate da *break* non terminali (*unità prosodiche*) individuano le *unità informative*, in cui è segmentato il flusso del pensiero²¹. Il parlato consiste quindi di due livelli gerarchicamente riconoscibili, di cui il *comment* costituisce l'unità fondante. Analogamente allo *stroke*, il *comment* è unità necessaria e sufficiente alla realizzazione di un atto linguistico e può occorrere non accompagnato da ulteriori unità informative, nel qual caso abbiamo allora una neutralizzazione dei due livelli e la loro coincidenza in un enunciato non modellizzato (*unpatterned*). Diversamente le unità informative intorno al *comment* possono essere variamente articolate realizzando enunciati modellizzati (*patterned*). Nel foglio di annotazione i due livelli saranno disposti su due righe predisposte che si aggiungeranno alle tre del gesto già menzionate.

²⁰ Il grafico è nostra raffigurazione del modello, le parentesi indicano opzionalità dell'elemento contenuto.

²¹ "flow of thought" in Chafe 1998.

4. Metodologia di annotazione

L'annotazione multimodale è stata condotta separatamente, audio (parlato) e video (gesto), da due annotatori differenti. Il parlato è stato osservato senza accesso alle immagini, il gesto ascoltando il parlato ma non accedendo alla sua trascrizione. In letteratura la metodologia relativamente alla separatezza o meno delle annotazioni non è univoca. McNeill trascrive prima il parlato e al suo interno successivamente inserisce gli indici gestuali allo scopo di segnarne con esattezza l'occorrenza in relazione a singole sillabe o vocali. L'autore ribadisce quindi l'assoluta necessità di trascrivere le unità gestuali non sono ascoltando il parlato ma anche guardandone la trascrizione²². D'altro canto Cornelia Müller (2018: comunicazione personale) consiglia di osservare i gesti senza alcun accesso all'audio. Nel nostro caso volendo verificare l'eventuale sincronia tra modalità, abbiamo optato per l'approccio di McNeill come però attuato da Loehr (2004): osservare il gesto ascoltandone il parlato co-occorrente senza tuttavia avere accesso alla sua trascrizione, quindi senza conoscerne preventivamente i break prosodici.

La trascrizione gestuale viene fatta sul software Elan, quella prosodica su WinPitch (Martin 2004) e su PRAAT (Boersma 2001) e da PRAAT importata successivamente in Elan e allineata ai dati gestuali.

In Figura 2 è riportato il totale delle unità annotate²³.

Figura 2 - Unità annotate (dati aggregati)

	Durata	Unità terminate (parlato)	Unità informative (parlato)	Unità comunicative (no parole)	Gesture Units (gesto)	Gesture Phrases (gesto)	Gesture Units no unità linguistiche	Gesture Phrases no unità linguistiche	Durata interruzione gestuale
Spontaneo	10:05,802	107	461	1	56	336	2	8	01:13,787
Recitato	07:49,345	113	271	7	53	177	0	8	02:22,136

Una volta riconciliate le due annotazioni in Elan si è proceduto a verificare inizio e fine delle unità gestuali in relazione ai break prosodici

²² In McNeill 2005: Appendix.

²³ I lavori precedenti presentano campioni osservati di lunghezza sensibilmente inferiore. L'annotazione gestuale richiede circa un'ora di lavoro per secondo di dato ed è quindi estremamente dispendiosa, pur con la strumentazione attuale. Loehr (2004) descrive 164 secondi e 147 gesti; Kendon (1972) 90 secondi; McClave (1991) 147. Il nostro lavoro cerca di offrire uno sguardo insolitamente ampio sul comportamento gestuale offrendo spezzoni di discorso per un totale di circa 17 minuti.

terminali e non-terminali, rilevando tendenze e modalità di sincronizzazione: inizio e fine delle gesture unit in relazione sia agli enunciati che alle unità informative²⁴; ugualmente con i gesture phrase; di entrambi si è contato se e quante volte gli arrivi gestuali oltrepassassero i break sia terminali che non; infine rispetto agli stroke si sono condotte osservazioni tipologiche e qualitative che evidenziassero le corrispondenze e la frequenza di occorrenze rispetto alle differenti occorrenze di unità informative (queste ultime date da un set chiuso di circa 20 tipi, funzionalmente divisi tra testuali, dialogici, e privi di valore informativo). Figura 3 mostra i livelli di annotazione di una gesture unit, da posizione di riposo a posizione di riposo con due stroke, e come il gesto si allinea alla segmentazione del parlato in unità informative.

Figura 3 - *Sincronizzazione gesto prosodia in ELAN*

The screenshot displays the ELAN software interface. At the top, there are four video frames showing a person sitting and gesturing with their hands. Below the video is a control bar with various playback and editing icons. The main area is a timeline with multiple tracks. The top track shows the transcription of the spoken text: "spire / che in quel momento / il personaggio / eh ha un dubbio / oppure che / c'era un significato / che a prima lettura non appariva / del [1] del [1] del testo / proprio". Below this, there are tracks for gesture units (GE 1-GU, GE 2-GPHR, GE 3-GPHA) and prosodic annotations (SCA, PAR). The SCA track shows annotations like "dove /SCA", "seguendo semplicemente la logica /SCA", "comune /SCA", and "non la metteresti mai /PAR". The PAR track shows annotations like "SCAPAR" and "PAR". The GE 2-GPHR track shows annotations like "5" and "11". The GE 3-GPHA track shows annotations like "11", "12", "preparation", "stroke", and "retraction".

5. Validazione

Il processo di validazione prosodica è attestato in Moneglia et al. 2010.

Il processo di validazione gestuale sviluppato nella presente ricerca ha visto la replica del 20% dell'annotazione da parte di un secondo annotatore esperto, lo studio qualitativo dei disaccordi, e la riconduzione di questi a consenso producendo criteri condivisi di annotazione (Cantalini et al. 2020)²⁵.

²⁴ L'avvenuta corrispondenza è stata definita all'interno di un margine di 200 ms prima e dopo la fine dell'unità linguistica.

²⁵ Per maggiori dettagli sugli aspetti qualitativi discussi si veda il volume indicato.

Si è quindi proceduto alla validazione quantitativa. Il tasso di sovrapposizione delle unità e quello di categorizzazione sono stati calcolati attraverso le seguenti metriche standardizzate: “overlap /extent value”, per *gesture unit* e *gesture phrase*, che ha mostrato un valore di 0.83 per le *gesture unit* e 0.61 per i *gesture phrase*; e “modified Cohen’s Kappa”, per le *gesture phrase*, calcolata secondo le modalità descritte in Holle & Rein (2013), che negli *stroke* è risultata di circa 0.85. I risultati ottenuti sono paragonabili con gli unici altri dati pubblicati sulla validazione dell’annotazione gestuale (Dvoretzka et al. 2013; per le metriche: Holle & Rein 2013 e Helmich & Lausberg 2014).

7. Evidenze osservate e produttività dell’approccio adottato

Il corpus raccolto e la metodologia attuata per osservarlo hanno rilevato le seguenti evidenze di sincronizzazione tra modalità e di variazione tra varietà, mostrando rilevante efficacia di analisi.

Per quanto riguarda la sincronizzazione tra gesto e unità linguistiche nel parlato spontaneo, sia le unità linguistiche terminate che quelle non terminate *segnano confini significativi* per l’allineamento delle unità gestuali: le *gesture unit* possono abbracciare più *gesture phrase* ed estendersi su più unità linguistiche; le *gesture unit* tendono a chiudere sui confini delle unità prosodiche terminali, mentre i *gesture phrase* su quelle non terminali; le *gesture unit* attraversano frequentemente i confini delle unità prosodiche terminali, mentre i *gesture phrase* quasi mai; lo *stroke* di conseguenza si posiziona sempre all’interno dell’unità di informazione. Tutto ciò oltre a mostrare corrispondenze sistematiche tra gesto e parlato evidenzia differenze di comportamento statisticamente significative tra i due livelli gestuali *unit* e *phrase*.

D’altra parte *la presenza di un dato normativo nello spontaneo consente di evidenziare le peculiarità del recitato* sia per quanto riguarda il parlato che per il gesto. È riscontrabile infatti una diversa incidenza, nelle due varietà, di enunciati semplici, cioè di realizzazioni in cui i due livelli linguistici coincidono e il livello informativo è schiacciato in quello illocutivo: nel parlato recitato vi sono un maggior numero di enunciati semplici, sia scanditi che non (cioè accompagnati da unità prosodiche non terminate prive di valore informativo), mostrando una preferenza dei parlanti per l’interpretazione illocutiva rispetto a quella

informativa. Analogamente a livello del gesto le gesture unit tendono a coincidere con i gesture phrase e quindi a non abbracciare più porzioni di discorso bensì ad allinearsi con enunciati quando non addirittura con unità informative. Come con il parlato, anche per il gesto le strategie modellizzanti risultano sensibilmente ridotte.²⁶

8. Conclusioni

Il modello di annotazione gesto – unità prosodiche (terminate e non terminate) approntato permette di studiare le modalità di sincronizzazione significative tra parlato e gesto. L'annotazione permette altresì di rilevare le differenze significative nelle modalità di sincronizzazione tra varietà (in questo caso tra spontaneo e recitato) e può restituire dati rilevanti di corrispondenze tra occorrenze gestuali e linguistiche. Consente inoltre avanzamenti nello studio della semantica del gesto.

Ci preme sottolineare come, con riguardo all'analisi linguistica, in molti casi di ambivalenza di interpretazione, sia proprio il gesto che può aiutare a disambiguare l'informazione linguistica e a contribuire ad una sua più accurata identificazione.

Riferimenti bibliografici

- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341–345.
- Bressem, Jana & Ladwig, Silva H. & Müller, Cornelia. 2013. Linguistic Annotation System for Gestures (LASG). In Müller, Cornelia, Cienki, Alan, Fricke, Ellen, Ladewig, Silva H., McNeill, David & Teßendorf, Sedinha (a cura di), *Body – Language – Communication: An International Hand-book on Multimodality in Human Interaction*. Handbooks of Linguistics and Communication Science 38(1), 1098–1124. Berlin: De Gruyter Mouton.
- Cantalini, Giorgina. 2018. *La gestualità co-verbale nel parlato spontaneo e nel recitato. Annotazione del gesto e correlati prosodici in campioni comparabili di attori italiani*. Roma: Università Roma Tre. (Tesi di dottorato.)
- Cantalini, Giorgina & Moneglia, Massimo & Gagliardi, Gloria & Proietti, Morgana. 2020. La correlazione gesto/prosodia e la sua variabilità: il par-

²⁶ Per una disamina più approfondita dei risultati si veda Cantalini & Moneglia 2022.

- lato spontaneo di contro alla performance attorale. In De Meo, Anna & Dovetto, Francesca M. (a cura di), *La Comunicazione Parlata / Spoken Communication. Napoli 2018*, 63–88. Roma: Aracne.
- Cantalini, Giorgina & Moneglia, Massimo. 2020. The annotation of gesture and gesture / prosody synchronization in multimodal speech corpora. *Journal of Speech Sciences* (9): 07-30.
- Cantalini, Giorgina & Moneglia, Massimo. 2022. Reduction of gesticulation and information patterning strategies in acted speech. In Fernandes, Carla & Evola, Vito & Ribeiro, Cláudia (a cura di), *Dance Data, Cognition, and Multimodal Communication*, 346–362. London–New York: Routledge.
- Chafe, Wallace. 1998. Language and the Flow of Thought. In Tomasello, Michael (a cura di), *The New Psychology of Language. Cognitive and Functional Approaches to Language Structure*, 87–104. New York: Routledge.
- Cresti, Emanuela. 2000. *Corpus di italiano parlato*. Firenze: Accademia della Crusca, Firenze.
- Cresti, Emanuela & Moneglia, Massimo. 2018. The illocutionary basis of Information Structure. Language into Act Theory (L-Act). In Adamou, Evangelia & Haude, Katharina & Vanhove, Martine (a cura di), *Information structure in lesser-described languages: Studies in prosody and syntax*, 359–401. Amsterdam: Benjamins.
- Dvoretzka, Daniela & Petermann, Kerstin & Skomroch, Harald. 2013. Calculating Temporal Interrater Agreement for Binary Movement Categories. In Lausberg, Hedda (a cura di), *Understanding Body Movement. A Guide to Empirical Research on Nonverbal Behaviour (With an Introduction to the NEUROGES Coding System)*, 253–260. Frankfurt am Main, Peter Lang Verlag.
- Ekman, Paul, & Friesen, Wallace V. 1969. The repertoire of nonverbal behavior. *Semiotica* 1. 49–98.
- Helmich, Ingo & Lausberg, Hedda. 2014. Hand Movements with a Phase Structure and Gestures that Depict Action Stem from a Left Hemispheric System of Conceptualization. *Experimental Brain Research* 232(10). 3159–3173.
- Holle, Henning & Rein, Robert. 2013.. The Modified Cohen’s Kappa: Calculating Interrater Agreement for Segmentation and Annotation. In Lausberg, Hedda (a cura di), *Understanding Body Movement. A Guide to Empirical Research on Nonverbal Behaviour (With an Introduction to*

- the NEUROGES Coding System*), 261–275. Frankfurt am Main: Peter Lang Verlag.
- Kendon, Adam. 1972. Some relation between body motion and speech: An analysis of an example. In Siegman, Aron W. & Pope, Benjamin (a cura di), *Studies in Dyadic Communication*, 177–210. New York: Elsevier.
- Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In Ritchie Key, Marie (a cura di), *The Relationship of Verbal and Nonverbal Communication and Language*, 207–227. The Hague: Mouton.
- Kendon, Adam. 2004. *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kita, Sotaro & van Gijn, Ingeborg & van der Hulst, Harry. 1998. Movement phases in signs and co-speech gestures, and their transcription by human coders. In Wachsmuth, Ipke & Fröhlich, Martin (a cura di), *Gesture and Sign Language in Human-Computer Interaction*, 23–35. Berlin: Springer-Verlag.
- Loehr, Dan. 2004. *Gesture and Intonation*. Ph.D dissertation. Washington, DC: Georgetown University.
- Maricchiuolo, Fridanna. 2017. La comunicazione non verbale. Caratteristiche e funzioni. *The Inquisitive Mind* 12.
- Martin, Philippe. 2004. WinPitch Corpus, a text to speech alignment tool for multimodal corpora. In Lino, Maria Teresa & Xavier, Maria Francisca & Ferreira, Fatima & Costa, Rute & Silva, Raquel (a cura di), *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*, 537–540. ELRA.
- McClave, Evelyn Z. 1991. *Intonation and Gesture*. Ph.D. dissertation. Washington, DC: Georgetown University.
- McNeill, David. 1985. So you think gestures are nonverbal? *Psychological Review* 92(3). 350–371.
- McNeill, David. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- McNeill, David. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.
- Moneglia, Massimo & Raso, Tommaso & Malvessi-Mittmann, Maryualê & Mello, Heliana. 2010. Challenging the perceptual relevance of prosodic breaks in multilingual spontaneous speech corpora: C-ORAL-BRASIL / C-ORAL-ROM. In *5th International Conference on Speech Prosody 2010*. International Speech Communications Association.

- Müller, Cornelia. 2018. Personal communication. In *8th Conference of the International Society for Gesture Studies: Gesture and Diversity. Cape Town.*
- Nencioni, Giovanni. 1976. Parlato-parlato, parlato-scritto, parlato-recitato. *Strumenti critici LX*. 1-56.
- Pirandello, Luigi. 1993[1918]. *Il giuoco delle parti*. Milano, Mondadori.
- Turk, Olcay. 2020. *Gesture, Prosody and Information Structure Synchronisation in Turkish*. Victoria University of Wellington. (Tesi di dottorato.)
- Wittenburg, Peter & Brugman, Hennie & Russel, Albert & Klassman, Alex & Sloetjes, Han. 2006. ELAN: A Professional Framework for Multimodality Research. In Calzolari, Nicoletta & Choukri, Khalid & Gangemi, Aldo & Meegaard, Bente & Mariani, Joseph & Odijk, Jan & Tapias, Daniel (a cura di), *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, 1556-1559. ELRA.

Web Sites

- ELAN <https://www.mpi.nl/corpus/html/elan/index.html>
- NEUROGES <https://neuroges.neuroges-bast.info/>
- PRAAT <http://www.fon.hum.uva.nl/praat/>
- WinPitch <https://www.winpitch.com/>