

Panoramica delle convenzioni di trascrizione HIAT

Thomas Schmidt, 20 novembre 2008

Peter M. Fischer, 18 gennaio 2010

Questa è una breve panoramica delle più importanti convenzioni di trascrizione HIAT come sono descritti in

Rehbein, J.; Schmidt, T.; Meyer, B.; Watzke, F. & Herkenrath, A. (2004) Handbuch für das computergestützte Transkribieren nach HIAT. In: Arbeiten zur Mehrsprachigkeit, Folge B (56). [http://www1.uni-hamburg.de/exmaralda/files/azm_56.pdf]

La panoramica vuole rappresentare un aiuto per coloro che non leggono il tedesco oppure coloro che vogliono comprendere il supporto HIAT in EXMARaLDA senza dover approfondire la teoria che sta dietro HIAT. Non è intesa come un sostituto equivalente della pubblicazione di cui sopra. La pubblicazione più recente su HIAT in lingua inglese è (per quanto mi consta) questo:

Ehlich, K. (1992). HIAT - a Transcription System for Discourse Data. Edwards, Jane / Lampert, Martin (eds.): Talking Data – Transcription and Coding in Discourse Research. Hillsdale: Erlbaum, 123-148.

1. Parole

L'ortografia standard è il modello base per la trascrizione delle parole, vale a dire: i confini di parola e la trascrizione delle parole seguono le convenzioni ortografiche vigenti, a meno che la loro pronuncia reale si discosti fortemente dalla loro pronuncia standard. In caso di forte deviazione, la corrispondenza fonema-grafema della rispettiva ortografia è usata per fare una cosiddetta 'trascrizione letterale'.

Esempi:

Italiano: 'sciamo' (per 'ci siamo')
Inglese: 'luvly' (per 'loveley')
Tedesco: 'zwohunnerfuffzich' (per 'zweihundertfünfzig', il numero 250)

La trascrizione letterale è da usare con parsimonia. La maggior parte dei fenomeni di fonetica standard della lingua parlata (come assimilazioni, elisioni in posizione finale) dovrebbe essere trascritta con la grafia standard piuttosto che con la trascrizione letterale.

Anche le lettere maiuscole si scrivono in base alle convenzioni ortografiche. In tedesco, per esempio, tutti i sostantivi e nomi propri sono scritti in maiuscolo e altrettanto vale per le lettere iniziali delle prime parole delle enunciazioni. Se s'impiega HIAT per altre lingue, occorre applicare le relative regole per l'uso delle maiuscole.

2. Parole speciali

Le parole speciali sono parole che o non sono prese in considerazione nelle regole di ortografia (in prima linea unità semi-lessicali, come le interiezioni) o la cui grafia standard non segue il principio della corrispondenza fonema-grafema (come abbreviazioni, numeri, ecc.). La regola di base per le parole speciali è descrivere la loro forma fonetica secondo gli stessi principi che si seguirebbero per le parole normali. Per la precisione:

- Interiezioni: Utilizza la grafia comune nella tua lingua per descrivere interiezioni (ad es. 'ehm', 'eh', 'ehi' in Italiano, 'ähm', 'äh', 'öhm' in Tedesco, 'eh', 'ehm', 'erm' in Inglese, 'euh' in Francese, ecc.). Se non c'è un'ortografia stabilita per un'interiezione data, usa la trascrizione letterale per riprodurre la forma fonetica il più fedelmente possibile.
- Numeri e date: Non scrivere i numeri in cifre, ma rendine la forma fonetica (trascrivi 'duecento' invece di '200'). Lo stesso vale per le date (trascrivi 'il due agosto millenovecentosettanta' invece di '2/08/1970').
- Abbreviazioni: Anche le parole per le quali esiste una grafia abbreviata ortograficamente corretta sono trascritte per intero, a meno che non vengano pronunciate appunto come abbreviazioni (ad es. 'Chiesa dei Santi Pietro e Paolo' invece di 'Chiesa SS. Pietro e Paolo', 'Professoressa Gianvittorio' invece di 'Prof.ssa Gianvittorio', 'eccetera' invece di 'ecc.'). Le abbreviazioni pronunciate in forma abbreviata sono trascritte nella loro forma standard (ad es. 'ONU'). Se la forma standard contiene dei puntini, essi vengono omessi (ad esempio 'SPQR', non 'S.P.Q.R.').
- Ortografia: Sequenze di spelling sono trascritte con degli spazi intermedi tra le lettere (ad es. 'Zimbabwe si scrive Z i m b a b w e').

3. Parole incomprensibili e incertezza nella trascrizione

Se il trascrittore è incerto su una parola o una sequenza di parole, la metta tra parentesi tonde. La parentesi di chiusura deve essere sempre collocata all'interno dell'enunciazione (cioè non oltre il simbolo di fine enunciazione). Inoltre, i passi incerti non devono superare i confini delle enunciazioni.

Giusto: Non sono (sicuro) che sia così.
 Non sono sicuro che (sia così).
 Non sono sicuro che (sia così). (Tu lo) sei?

Sbagliato: Non sono sicuro (che sia così).
 Non sono sicuro che (sia così. Tu lo) sei?

Se un tratto del discorso è del tutto incomprensibile, esso è indicato con la stringa '((incomprensibile))', possibilmente segnalando la lunghezza del tratto, ad es. '((incomprensibile, 2.5s))' per un tratto incomprensibile di 2,5 secondi. In lingue diverse dall'italiano si dovrebbe ricorrere alle traduzioni corrispondenti (ad es. '((unintelligible))' in inglese o '((unverständlich))' in tedesco). Se il tratto incomprensibile è molto breve, anche la notazione può essere abbreviata (ad es. '((incompr.))').

4. Fenomeni non fonologici

Fenomeni non fonologici (ad es. tossire, ridere, schiarirsi la gola), se alternativi (dunque non simultanei) al parlato, vengono descritti tra parentesi tonde doppie, possibilmente notando la durata del fenomeno. Questi fenomeni possono verificarsi all'interno di un'enunciazione, all'inizio di un'enunciazione o isolatamente.

Ho proprio ((tossisce)) un bel mal di gola.
 Guarda, ((ride, 1.2s)) è divertentissimo!
 ((tossisce, 2s)) Ho fumato troppo.
 ((si schiarisce la gola))

5. Pause

Tutte le pause devono essere trascritte. Le pause sono descritte in base alla loro lunghezza. Se la lunghezza della pausa viene solo stimata, per indicarla si usano dei punti elenco, e cioè:

- un punto per una breve interruzione del flusso del discorso (micropausa)
- due punti per una pausa stimata fino a un mezzo secondo
- tre punti per una pausa stimata fino a un secondo

Le pause misurate, e inoltre quelle stimate superiori a un secondo vengono quantificate in numeri, come segue:

((2.3s)) per una pausa di 2,3 secondi

Le pause sono attribuite ai parlanti applicando le seguenti regole:

- Le pause tra un cambio di parlanti sono attribuite all'enunciazione successiva:

X [v]	Dico qualcosa.
B [v]	•• E poi tocca a me.

- Le pause all'interno di un'enunciazione sono attribuite al parlante che è di turno:

X [v]	Dico •• qualcosa.
-------	-------------------

- Questa regola si applica anche quando l'enunciazione di un parlante è interrotta da una (di solito breve) enunciazione di un altro parlante:

X [v]	Dico •• qualcosa.
B [v]	Sì!

- Le pause tra due enunciazioni dello stesso parlante sono attribuite all'enunciazione successiva:

X [v]	Dico qualcosa. •• E poi aggiungo un'altra cosa.
-------	---

6. Eventi, enunciazioni e catene di segmenti

Quando si trascrive in una partitura musicale, il trascrittore suddivide la trascrizione in eventi. Sia un nuovo evento per lo meno ogni volta che si verifica un cambio di parlante, ossia ogni volta che un nuovo parlante inizia o finisce un intervento entro il discorso. Ad esempio nel brano seguente ci sono in tutto tredici eventi – 8 per il parlante X e 5 per il parlante B.

	0	1	2	3	4	5	6	7	8	9	10
X [v]	Ciao.	Un attimino di pazienza	((tossisce, 1s)) Eccomi.	Allora il primo ottobre abbiamo detto.		Ott	ob	re.	Si per me va bene.		
B [v]					Si il primo di otto	bre.		Non	dalle dieci alle undici	di sera, eh.	

Nota: Gli eventi sono fenomeni esclusivamente diacronici. Quindi, in linea di principio è possibile avviare un nuovo evento in qualsiasi punto nell'asse del tempo.

Le catene di segmenti sono definite sulla base degli eventi: una catena di segmenti è una sequenza di eventi temporalmente contigui all'interno di una traccia. Così, nell'esempio di sopra ci sono 4 catene di segmenti:

- (1) **X**: Ciao. Un attimino di pazienza. ((tossisce, 1s)) Eccomi. Allora il primo ottobre abbiamo detto.
- (2) **B**: Sì il primo di ottobre.
- (3) **X**: Ottobre. Sì per me va bene.
- (4) **B**: Non dalle dieci alle undici di sera, eh.

Così come gli eventi, allo stesso modo anche le catene di segmenti non si riferiscono ad alcuna specifica teoria circa le entità nella lingua parlata.

Al contrario, il concetto di **enunciazione** è motivato linguisticamente, come spiegato in dettaglio da Rehbein et al. (2004: 19 ss.). In questa sede ci limitiamo a riassumere le caratteristiche salienti dell'enunciazione:

1. Ogni catena di segmenti può essere suddivisa in una o più enunciazioni. Le enunciazioni non superano mai i confini di una catena di segmenti.
2. I confini delle enunciazioni sono determinati in base a criteri morfo-sintattici (ad es. l'ordine delle parole) e criteri di intonazione (ad es. aumenti o diminuzioni di livello).
3. Il criterio fondamentale per individuare e classificare un'enunciazione è la sua modalità.
4. HIAT distingue cinque tipi di enunciazione. Per ogni tipo vi è un simbolo corrispondente che marca la fine dell'enunciazione.
 - a. Affermazioni (punto fermo, ad es. "Mi piace.")
 - b. Domande (punto interrogativo, ad es. "Ti piace?")
 - c. Esclamazioni, interpellazioni, richieste, ecc. (punto esclamativo, ad es. "Buongiorno!")
 - d. Enunciazioni interrotte (punti di sospensione, ad es. "Volevo solo...")
 - e. Enunciazioni non modali (punto in apice, ad esempio "Mmm'")

Per tanto, le 4 catene di segmenti del esempio di sopra possono essere suddivise in 8 enunciazioni (8 affermazioni).

- (1) **X**: Ciao.
- (2) **X**: Un attimino di pazienza.
- (3) **X**: ((tossisce, 1s)) Eccomi.
- (4) **X**: Allora il primo ottobre abbiamo detto.
- (5) **B**: Sì il primo di ottobre.
- (6) **X**: Ottobre.
- (7) **B**: Non dalle dieci alle undici di sera, eh.
- (8) **X**: Sì per me va bene.

Tra i cinque tipi di enunciazione, le enunciazioni interrotte e quelle non modali sono le più difficili da classificare. Particolarmente importante è la distinzione delle enunciazioni interrotte (come in (a)) dalle enunciazioni che contengono una sequenza di riparazione (come in (b)).

(a) Non voglio... Non puoi farmelo fare.

(b) Non vo/ non voglio farlo.

Il criterio decisivo è stabilire se

- abbiamo a che fare con un'enunciazione completa in cui alcuni difetti vengono riparati in superficie (sequenza di riparazione, esempio (b)) o
- la verbalizzazione *in fieri* di un'enunciazione viene interrotta, e al suo posto viene verbalizzata una nuova enunciazione (enunciazione interrotta, esempio (a))

Tipiche enunciazioni non modali sono le interiezioni di un partecipante che al momento funge da ascoltatore e *non* pretende di prendere la parola. Vale a dire, le enunciazioni non modali sono dei segnali tipici degli ascoltatori.

7. Riparazioni

Le riparazioni sono auto-correzioni spontanee. Una riparazione si compone tipicamente di tre parti, un reparandum (la parte da riparare), un intervento (spesso un segnale di esitazione come "ehm") e un reparans (la parte che appunto ripara), anche se non tutte queste componenti sono necessariamente presenti in ogni riparazione. Nella trascrizione, la sequenza di riparazione è rappresentata da una barra subito dopo il reparandum, ad esempio:

Beh, il suo/ eh lui preferirebbe iniziare più tardi.

8. Commenti del trascrittore

Una traccia supplementare per ogni parlante (con la categoria 'k' e di tipo 'a') può contenere gli eventuali commenti del trascrittore. Tali osservazioni dovrebbero essere esattamente allineate con la parte nella trascrizione verbale a cui si riferiscono, cioè il commento e quanto commentato dovrebbero trovarsi in eventi che abbiano i medesimi punti di inizio e di conclusione.

	0	1	2	3
X [v]	Eh hai rotto il	(cavolo)	acciderba!	
X [k]		tavolo?		

I commenti del trascrittore di solito offrono informazioni in merito a:

- un comportamento non verbale parallelo a quello verbale (ad es. "ridendo")

X [v]	Non credo proprio che l'abbia detto.
X [k]	ridendo

- la forma standard di una forma che, nella traccia verbale, se ne discosta fortemente

X [v]	Mi scusassi, li bagni unni sunnu?
X [k]	per: dove sono

- una trascrizione in base all'alfabeto IPA delle peculiarità di pronuncia che non possono essere adeguatamente rappresentate con una trascrizione letterale

- le alternative possibili nel caso di incertezza

X [v]	Eh hai rotto il (cavolo) acciderba!
X [k]	tavolo?

9. Altri fenomeni

Per motivi di semplicità, i seguenti fenomeni non vengono menzionati qui, anche se HIAT prevede delle regole per la loro trascrizione:

- Transizioni lisce da un'enunciazione all'altra ("blocco")
- Aumento dell'atto linguistico
- Discorso diretto, citazione, ecc.
- Sillabe accentuate
- Altri fenomeni non-segmentali (modulazione, velocità, ecc.)
- Comunicazioni
- Toni

Ai principianti HIAT, si suggerisce di ignorare tali fenomeni al meno all'inizio, e iniziare con le otto classi di fenomeni sopra descritte.