

## Appendice C.

*SAMPA* (con aggiunte e modifiche, per la trascrizione allineata dell'italiano e di alcune sue varietà)<sup>258</sup>

### C.1. Simboli per vocoidi

Descrizione	<i>IPA</i>	<i>SAMPA</i>	<i>SAMPA</i> per it.
vocale anteriore alta non arr.	i	i	i
voc. ant. alta arrotondata	y	y	
voc. ant. quasi alta non arr.	ɪ	I	
voc. ant. quasi alta arrotondata	ʏ	Y	
voc. ant. semi-alta non arr.	e	e	e
voc. anteriore semi-alta arrotondata	ø	2	
voc. ant. semi-bassa non arr.	ɛ	E	E
voc. ant. semi-bassa arrotondata	œ	9	
voc. ant. quasi bassa non arr.	æ	{	
voc. ant. (centrale) bassa non arr.	a	a	a
voc. ant. bassa arrotondata	ɶ	&	
voc. posteriore bassa non arr.	ɑ	A	
voc. post. bassa arrotondata	ɒ	Q	
voc. post. semi-bassa non arr.	ʌ	V	
voc. post. semi-bassa arrotondata	ɔ	O	O
voc. post. semi-alta non arr.	ɤ	7	
voc. post. semi-alta arrotondata	o	o	o
voc. post. quasi alta arrotondata	ʊ	U	
voc. post. alta arrotondata	u	u	u
voc. post. alta non arr.	ʉ	m\	
vocale centrale alta non arr.	ɨ	l	
vocale centrale alta arrotondata	ɥ	}	
vocale centrale semi-alta non arr.	ɘ	@\	
vocale centrale semi-bassa non arr.	ɜ	3	
vocale centrale semi-alta arr.	ø	8	
vocale centrale semi-bassa arr.	ɞ	3\	
vocale centrale quasi bassa non arr.	ɐ	6	

258 Le convenzioni qui presentate si basano su quelle di un documento operativo messo a punto per il "Tirocinio di Fonetica Applicata" della Fac. di Lingue e Lett. Straniere dell'Università di Torino (edizioni 2002-2006). Il documento trae spunto dalle linee-guida definite nell'ambito del progetto *CLIPS* (a cura di R. Savy; v. sito *web*) e, ovviamente, si basa sulle convenzioni originarie *SAMPA* (del progetto omonimo coordinato da J.C. Wells; v. sito *web*). La necessità di ricorrere a queste convenzioni per la segmentazione e l'etichettatura di materiali di parlato (di laboratorio o spontaneo) si giustifica con le specifiche di uniformità e di trattabilità computazionale delle annotazioni eseguite (v. Appendice D).

vocale centrale (*schwa*)                    ə                    @                    @<sup>259</sup>

**C.2. Simboli per contoidi**

Descrizione	IPA	SAMPA	SAMPA per it.
cons. (occlusiva) nasale bilabiale	m	m	m
cons. occlusiva bilabiale sorda	p	p	p
cons. occl. bilabiale sonora	b	b	b
cons. costr. (o appross.) bilabiale sorda	ɸ	p\	
cons. costr. (o appross.) bilabiale sonora	β	B	
cons. (occlusiva) nasale labio-dentale	ɱ	F	M
cons. costrittiva labio-dentale sorda	f	f	f
cons. costr. labio-dentale sonora	v	v	v
cons. semi-occlusiva labio-dentale sorda	ɸ̄	pf	
cons. appr. labio-dentale sonora	ɸ	v\	
cons. (occlusiva) nasale alveolare	n	n	n
cons. occlusiva alveodentale sorda	t	t	t
cons. occl. alveodentale sonora	d	d	d
cons. costr. (o appr.) (inter-)dentale sorda	θ	T	
cons. costr. (o appr.) (inter-)dentale sonora	ð	D	
cons. costr. alveodentale sorda	s	s	s
cons. costr. alveodentale sonora	z	z	z
cons. semi-occlusiva alveodentale sorda	ts̄	ts	ts
cons. semi-occl. alveodentale sonora	d̄z	dz	dz
cons. (appross.) laterale alveolare	l	l	l
cons. (poli)vibrante alveolare	r	r	r
cons. (mono)vibrante alveolare	ɾ	4	4
			[(</r/ in certi contesti)]
cons. approssimante alveolare	ɹ	r\	
cons. costrittiva postalveolare sorda	ʃ	S	S
cons. costr. postalveolare sonora	ʒ	Z	Z
cons. semi-occlusiva postalveolare sorda	t̄ʃ	tS	tS
cons. semi-occl. postalveol. sonora	d̄ʒ	dZ	dZ

259 Il simbolo @ è usato nella trascrizione del parlato spontaneo italiano per segnalare il grado massimo di riduzione timbrica che può verificarsi per le vocali non accentate in condizioni d'ipoarticolarità: gradi intermedi di riduzione possono produrre modificazioni timbriche tali da far considerare il timbro risultante come non corrispondente a quello atteso. In tali casi, quando questo si presenti non diversamente descrivibile, si premette un punto esclamativo (!) al simbolo della vocale. Ad es. la seconda, l'ultima e – in buona misura – anche la terza /a/ di *abracadabra* (/abrakad"abra/<sub>SAMPA</sub>), in molti casi, non presentano un timbro descrivibile come [a]. La debole energia con cui sono prodotti questi suoni rende il loro timbro poco precipuo e tuttavia ancora dissimile dagli altri per cui si dispone di simboli appositi: la soluzione adottata è quindi quella di trascrivere [!a]<sub>SAMPA</sub>: quando il timbro è invece totalmente centralizzato si può ricorrere a [@]<sub>SAMPA</sub>.

APPENDICI

cons. (occlusiva) nasale palatale	ɲ	J	J
cons. occlusiva palatale sorda	c	c	
cons. occl. palatale sonora	tʃ	j\	
cons. costrittiva palatale sorda	ç	C	
cons. costr. palatale sonora	ʝ	jj	
cons. (costr. o appross.) laterale palatale	ʎ	L	L
cons. approssimante palatale	j	j	j
cons. (occlusiva) nasale velare	ŋ	N	N
cons. occlusiva velare sorda	k	k	k
cons. occl. velare sonora	g	g	g
cons. costrittiva velare sorda	x	x	
cons. costr. velare sonora	ɣ	G	
cons. approssimante velare	ʝ	M\	
cons. occlusiva uvulare sorda	q	q	
cons. costrittiva uvulare sorda	χ	X	
cons. costr. uvulare sonora	ʁ	R	
cons. vibrante uvulare	r	R\	
cons. costrittiva faringale sorda	ħ	H\	
cons. costr. faringale sonora	ʕ	?\	
cons. costr. (o appross.) laringale sorda	h	h	
cons. costr. laringale sonora	ɦ	h\	
cons. occlusiva laringale sorda	ʔ	?	? <sup>260</sup>

Concludiamo questa lista di corrispondenze ricordandone alcune che, seppur insolite nell'italiano parlato comune, possono rivelarsi utili in casi speciali. Ad es. i contoidi retroflessi sono notati di solito con i simboli corrispondenti del luogo alveolare cui si giustappone il diacritico dell'accento grave ( ` ): i contoidi occlusivi sono quindi t ` (sordo) e d ` (sonoro); i costrittivi s ` (sordo) e z ` (sonoro); il nasale n `; il vibratile r `; l'approssimante r `; l'approssimante laterale l ` . Il più raro suono approssimante velare laterale può essere notato L ` , così come il nasale uvulare è N ` , il vibrante bilabiale è B ` e quello uvulare è R ` <sup>261</sup>.

260 Nonostante il colpo di glottide possa comparire in italiano solo per fini espressivi, il simbolo ? è adottato, nella trascrizione del parlato spontaneo, per trascrivere i fenomeni di laringalizzazione che possono presentarsi in vari contesti, ma – soprattutto – assai frequentemente nella realizzazione di casi di dialefe (vocali a contatto in iato fonosintattico, v. cap. III).

261 Per i suoni costrittivi laterali (alveolari) i simboli proposti sono K (per il sordo) e K ` (per il sonoro). L'approssimante labiale–velare sordo è W, mentre quello labiale–palatale sonoro è H. Tra i simboli fuori tabella, vi sono inoltre quello per i costrittivi epiglottidali sordo, H ` , e sonoro, < ` . L'occlusivo epiglottidale è invece trascritto come > ` . Per i costrittivi alveolo–palatali si può usare s ` , per il sordo, e z ` , per il sonoro. Il vibratile alveolare laterale è l ` ; la presunta articolazione simultanea di S e x (usata per il fonema svedese /ɣ/, v. Appendice B) è infine indicata da x ` .

### C.3. Simboli diacritici

Altri simboli d'indiscussa utilità sono quelli per la notazione: dell'accento primario, *virgolette banali* ( " ), da premettere alla vocale accentata; della lunghezza (vocalica o consonantica), *due punti* ( : ); della sillabicità, *uguale* ( = ); della nasalità, *tilde* ( ~ ); questi ultimi tre giustapposti al simbolo di base.

Altri diacritici adottati per indicare il risultato di alcuni processi, da giustapporre al simbolo di base, sono i seguenti:

_0	per indicare la desonorizzazione	[◌̥] <sub>IPA</sub>
_v	per indicare la sonorizzazione	[◌̚] <sub>IPA</sub>
_h	per indicare l'aspirazione	[◌̚ <sup>h</sup> ] <sub>IPA</sub>
_j	per indicare la palatalizzazione	[◌ <sup>j</sup> ] <sub>IPA</sub>
_w	per indicare la labializzazione	[◌ <sup>w</sup> ] <sub>IPA</sub>
_G	per indicare la velarizzazione	[◌ <sup>ɣ</sup> ] <sub>IPA</sub>
_e	per indicare velarizzazione/faringalizzazione	[◌̠] <sub>IPA</sub>
_?\	per indicare la faringalizzazione	[◌̠̠] <sub>IPA</sub>
_n	per indicare il rilascio nasale di un'occlusione	[◌ <sup>n</sup> ] <sub>IPA</sub>
_l	per indicare il rilascio laterale di un'occlusione	[◌ <sup>l</sup> ] <sub>IPA</sub>
_}	per indicare un'occlusione senza rilascio udibile	[◌̚̚] <sub>IPA</sub>

Sono stati inoltre introdotti i seguenti diacritici specifici: *\_f* (per indicare la spirantizzazione; ad es. *p\_f* = p con elementi di frizione) e *\_q, \_s, \_x* (per indicare forme di affricazione parziale rispettivamente di p, t e k). Altri simboli adottati per etichettare fenomeni di cancellazione e/o d'inserzione sono invece rispettivamente i simboli – e +, entrambi premessi al simbolo del fono cancellato o inserito<sup>262</sup>.

262 L'esempio di aferesi visto al cap. III, *e in terra* → *e 'n terra*, sarebbe trattato etichettando separatamente la realizzazione di *e* e quella dell'*in* seguente che, costituito da un unico fono, riceverebbe una regolare trascrizione mediante un simbolo opportuno (verosimilmente [n]<sub>SAMPA</sub>) preceduto però dall'indicazione [–i]<sub>SAMPA</sub>. Un uso frequente di + è richiesto dalle necessità di notazione dei numerosi fenomeni d'inserzione (epentesi, paragoge). Molto frequente in italiano è ad esempio l'aggiunta, anche da parte di *speaker* professionisti e/o in condizioni di parlato sorvegliato, di vocali paragogiche in parole terminanti per consonante che si trovino in posizione prepausale o che, per ragioni diverse, non si leghino con il suono iniziale dell'eventuale parola seguente. Ad es. *sport* /spɔrt/<sub>il,IPA</sub> (/spɔrt/<sub>il,SAMPA</sub>) è normalmente pronunciato [s'pɔrtə]<sub>il,IPA</sub>; in base a questa convenzione, la sua trascrizione *SAMPA* sarebbe [s'pɔrt+@]<sub>SAMPA</sub>. Per fare un altro esempio: l'espressione *dal Nord e dal Sud* [dal'nɔrded:al'sud]<sub>IPA</sub> / [dal n'ɔrd e d:al s'ud]<sub>SAMPA</sub> è normalmente pronunciata [dal'nɔrded:al'sud:ə]<sub>IPA</sub> / [dal n'ɔrd e d:al s'u\$d:+@#]<sub>SAMPA</sub> (# = pausa potenziale o reale; \$ = geminazione inattesa, v. dopo) con paragoge di @ (pronunce simili sono attestate anche per parlanti settentrionali, in assenza di allungamento delle due /d/). Si noti ancora che, nell'italiano di alcuni *speaker* radiotelevisivi, per perseguire una presunta maggiore chiarezza, si può arrivare ad avere [dal:ə# 'nɔrdə# ed:al:ə# 'sud:ə#]<sub>IPA</sub> / [da\$!:+@# n'ɔrd+@# e da\$!:+@# s'u\$d:+@#]<sub>SAMPA</sub> con pause potenziali anche dopo i due *dal*. Gli esempi più comuni di epentesi in italiano riguardano invece le vocali *svarabhaktiche* che si sviluppano in prossimità di alcune realizzazioni di /t/ (v. esempi al §VI.7).

Nell'applicazione di questi criteri per la trascrizione allineata dell'italiano, è stato inoltre necessario definire un ordine d'uso dei diacritici: salvo eccezioni, i mutamenti che si osservano nella resa di un fono sono etichettati segnando dapprima quelli di modo, poi quelli di luogo e infine quelli che riguardano la sonorità (ad es. alla notazione  $[\text{r}^{\text{h}}]_{\text{IPA}}$  corrisponde  $[\text{r\_f\_0}]_{\text{SAMPA}}$ ).

Risultato di una convenzione adottata diffusamente per l'italiano sono infine: il simbolo della *percentuale* ( % ), usato per segnalare difficoltà di segmentazione (posizionamento dubbio di un demarcatore di confine tra segmenti)<sup>263</sup>; il simbolo del *dollaro* ( \$ ), usato per segnalare casi di geminazione inattesa; il simbolo della *e commerciale* ( & ), usato per segnalare casi di degeminazione (entrambi da premettere al simbolo della consonante soggetta al fenomeno)<sup>264</sup>.

Anche se non sono stati finora adottati nella trascrizione allineata di parlato italiano, ulteriori diacritici sono inoltre disponibili per la notazione di variazioni timbriche nelle vocali (v. sito *web SAMPA*).

263 Si noti che, nella trascrizione allineata, il confine tra i segmenti, dipendente dal *software* usato, è di solito segnalato da un demarcatore di tipo |. Una collocazione incerta del demarcatore ( %|%) è spesso usata nel caso di separazione dei due elementi di un dittongo ascendente (uno 'consonantico' e l'altro propriamente vocalico). Una scelta conservativa (in parte motivata da ragioni fonologiche) è invece quella di non separare i due elementi vocalici di un dittongo discendente anche in quei casi in cui siano evidenti le regioni temporali in cui si estendono. Al contrario si separano di solito i vocoidi che realizzano vocali in iato (col ricorso a %|%) nei casi d'incertezza).

264 La degeminazione si segna solo quando non corrisponde alle attese. Notare che il *RF* (radoppiamento fonosintattico, v. cap. III) va considerato una forma di geminazione regolarmente prevedibile (standard). Di conseguenza non è necessario segnare come  $[\text{\$k:}]_{\text{SAMPA}}$  la resa lunga di /k/ in  $a_{\text{RF}}$  *capo* (che si trascrive normalmente  $[\text{a k}^{\text{h}}\text{:a:po}]_{\text{SAMPA}}$ ). Occorre invece trascrivere  $[\text{a \&k}^{\text{h}}\text{:a:po}]_{\text{SAMPA}}$  quei casi in cui l'allungamento non sia stato realizzato dal parlante. Un caso di geminazione inattesa è ad es. una pronuncia non-breve di occlusive sorde postconsonantiche, come può accadere in *tuta*, realizzato come se fosse *tutta*: in tal caso la notazione sarebbe  $[\text{t}^{\text{h}}\text{u}\text{\$t:a}]_{\text{SAMPA}}$ . Un altro esempio di geminazione incontrollata è riferibile anche a casi d'indebita eufonizzazione della congiunzione *e* ( $\rightarrow$  *ed*). Il fenomeno si presenta in alcuni stili di parlato radiotelevisivo (in esempi simili a uno di quelli visti nel caso di paragoge di @). Ad es., invece di formulare la normale espressione *al Nord e al Sud* con la congiunzione *e*, un noto meteorologo televisivo preferisce ricorrere a *ed*, pronunciando però innaturalmente *al Nord ed al Sud* come  $[\text{al}^{\text{h}}\text{no:rd ed:al}^{\text{h}}\text{sud:}\text{\textcircled{a}}]_{\text{IPA}}/[\text{al n}^{\text{h}}\text{ord e}\text{\$d: al s}^{\text{h}}\text{u}\text{\$d:}+\text{\textcircled{a}}]_{\text{SAMPA}}$ ; la geminazione finale riguarda quindi *ed* e *Sud*. Si noti che, in questo caso, l'inattesa presenza del /d/ di *ed* e il suo allungamento creano una potenziale ambiguità, in quanto suggeriscono la ricostruzione della sequenza *e dal Sud* (la quale è resa perfettamente plausibile dato che, in tal caso, si sarebbe in presenza di  $e_{\text{RF}}$ ) di significato diametralmente opposto.



## Appendice D.

### Trascrizione ortografica convenzionale e annotazione di materiali orali

[da Savy *et al.*, 2006]

#### D.1. Trascrizione ortografica convenzionale

Per **trascrizione ortografica convenzionale** s'intende quella che si può adottare nell'eseguire la 'sbobinatura' di una registrazione secondo un sistema di trascrizione convenzionale. Per **annotazione** s'intende invece l'aggiunta di dettagli descrittivi e interpretativi.

Gli obiettivi da perseguire nella trascrizione annotata devono tener conto dei seguenti principi: quello della *leggibilità* e quello dell'*iconicità* cui si può aggiungere un'attenzione particolare nei riguardi di un'opportuna distribuzione spaziale del testo, nel rispetto della prossimità di eventi correlati e della separabilità visiva di eventi diversi. Ovviamente, si accorda una certa priorità logica nella codifica di informazioni utili all'interpretazione di eventi successivi, nell'uso di marche e simboli possibilmente trasparenti e facilmente memorizzabili, nell'ottica di una trattabilità computazionale dei documenti (che richiede sistematicità e uniformità della codifica)<sup>265</sup>.

In tal modo, si può definire **annotazione** un'operazione di arricchimento della trascrizione ortografica, mediante dettagli relativi alla produzione, simboli di categorizzazione basilare degli eventi fonici e commenti extratestuali, mentre l'**etichettatura** può consistere in un insieme di operazioni volte a definire, identificare e classificare le unità linguistiche costitutive di un testo ai vari livelli (fonetico, fonologico, prosodico, morfo-lessicale, morfo-sintattico, testuale, comunicativo etc.). Un'etichettatura multilivello come questa richiede ad esempio le seguenti operazioni:

- a) suddivisione del testo in stringhe, identificabili univocamente, che corrispondano a produzioni unitarie di un singolo parlante;
- b) annotazione d'informazioni sulla sequenza e sovrapposizione di produzioni di parlanti diversi (per esempio turni, nel caso di scambi dialogici o conversazioni);
- c) trascrizione in forma univoca di tutti gli elementi lessicali della produzione, compresi numeri, sequenze di lettere (come acronimi e sigle), interiezioni, abbreviazioni etc.;
- d) trascrizione degli elementi verbali non-lessicali che comprendono disfluenze, errori di pronuncia, frammenti di parole, false partenze, troncamenti, esitazione, nonché pause (piene o vuote) e sequenze inintelligibili;
- e) annotazione dei fenomeni vocali non verbali prodotti dai parlanti (inspirazioni, risate etc.);
- f) annotazione dei rumori presenti nel segnale, prodotti dal parlante o dall'ambiente;

<sup>265</sup> Si vedano a questo riguardo le raccomandazioni espresse da alcune fonti particolarmente autorevoli citate nei documenti operativi diffusi nell'ambito dei progetti nazionali *AVIP*, *API* e *CLIPS* (cfr. Savy, 2002; Savy *et alii*, 2006).

- g) inserimento di brevi commenti del trascrittore relativi alla qualità del segnale, a parti di testo, alle variazioni del volume di voce del parlante (volute o accidentalmente determinate da caratteristiche di registrazione).

❖ *Unità di trascrizione*

Mentre per una lista di parole di parlato di laboratorio l'unità di trascrizione può essere la **parola** o il **sintagma** o l'**enunciato**, per il parlato spontaneo è possibile riferirsi alle nozioni di **unità terminale** e **unità non-terminale** (prevalentemente nel parlato monologico) oppure, nel parlato dialogico, al **turno dialogico**. Fondamentale indicazione di turno dialogico è la coerenza semantico-pragmatica interna alla produzione di uno stesso locutore.

Sebbene in alcuni casi l'individuazione di unità di questo tipo rimanga inevitabilmente condizionata da un certo grado di soggettività, si può comunque definire 'turno' la 'presa di parola' da parte di uno di due interlocutori, sia che essa interrompa effettivamente l'altro locutore, sia che si sovrapponga alle produzioni di quest'ultimo senza dare luogo necessariamente a un'interruzione<sup>266</sup>.

❖ *Aggiunta di commenti extratestuali del trascrittore*

Per commenti extratestuali s'intendono quei commenti espressi dal trascrittore a proposito di parti delimitate di testo, ad esempio cambiamenti localizzati di registro vocale o commutazioni di codice. Questi possono essere annotati tra parentesi quadre [ ] (ad es. [sussurrato] oppure [gridato], nel primo caso, oppure [dialettale] nel secondo).

Per la collocazione nei riguardi del testo, si seguono due principî: si pone l'annotazione dopo l'elemento lessicale, se questa si riferisce a una sola parola (ad es.: *va bbuò' [dialettale]*), oppure prima di una sequenza, inclusa tra {}, se il fenomeno annotato si estende su più di una parola (ad es.: *{{[gridato] no , aspetta !}}*)<sup>267</sup>.

❖ *Trascrizione della sequenza lessicale*

- La sequenza lessicale va trascritta in linea di massima senza utilizzare lettere maiuscole. Queste sono limitate alla trascrizione di nomi propri e sigle (es.: *Maria, AIDS*).
- Le parole in forma ridotta vanno trascritte secondo la pronuncia (es.: *prof* per *professore*); nel caso di forme con aferesi o elisione si utilizza l'apostrofo (es.: *'ste* per *queste*, *m'ha* per *mi ha...*)

266 Non è considerata interruzione di turno la presa di parola con funzione fatica (per esempio, manifestazioni di assenso, espressioni di esitazione, sorpresa, disappunto da parte dell'interlocutore, false partenze e simili).

267 Notare che, contrariamente all'ortografia tradizionale, in questi ambiti la punteggiatura viene sempre separata dal testo precedente (e seguente).

- Le forme dialettali si trascrivono adottando, quando possibile, un criterio di normalizzazione per la scrittura, mantenuto invariato per ogni occorrenza del termine (es.: *guaglione*)
  - Le sequenze di numeri non vanno trascritte in forma di cifre, ma secondo il modo in cui sono state pronunciate (es.: *ventinovemila*, *centouno* oppure *cento e uno...*).
  - Le parole o sequenze inintelligibili vanno indicate come <inintelligibile>.
- ❖ *Inserimento di elementi semi-lessicali, non lessicali, non verbali, non vocali e commenti*

È opportuno distinguere, su questo piano, vari elementi che possono essere classificati come segue:

#### **Fenomeni semilessicali**

- Frammenti di parole non finite (disfluenze) che vanno marcati con ‘+’ alla fine del frammento (ad es.: *non lo ve+ non lo vedo*).
- Eventuali interruzioni interne all’elemento lessicale vanno marcate con ‘\_’ (es.: *mon\_tato*).
- Errori di pronuncia e lapsus che danno luogo a non-parole vanno marcati con ‘\*’ all’inizio della parola (es.: *\*altanelante* per *altalenante*) in modo da distinguerli da errori di battitura.
- Le false partenze senza pausa di interruzione sono marcate con ‘/’ (es.: *ma tu / dov’è la chiave?*).

#### **Fenomeni verbali non lessicali** (indicati tra ganci <>)

- PAUSE
  1. Le pause vuote sono quelle pause di silenzio più o meno brevi che non interrompono il flusso logico del discorso (l’enunciato continua dopo la pausa): ‘<pb>’ per una pausa breve, ‘<pl>’ per una pausa lunga (*la vedi ? <pb> sulla sinistra <pl> c’è scritto fiume*). Un’interruzione (più o meno duratura) può invece essere segnalata da <P>.
  2. Le pause piene possono essere di due tipi:
    - se riempite da fenomeni interiettivi, sono trascritte con <eeh> per la semplice vocalizzazione, <ehm> per quella con nasalizzazione (ad es.: <ehm> allora <eeh> fai mezzo cerchio);
    - se riempite da allungamenti dell’ultima vocale o consonante di parola, sono marcate con la duplice ripetizione della vocale <vv> o consonante <cc> alla fine dell’elemento lessicale interessato (es.: allora<aa> ...; con<nn> ...; indipendentemente dal timbro e durata effettivi).
  3. Le articolazioni di esitazione vanno trattate come le pause piene di primo e secondo tipo; va segnalato inoltre l’eventuale allungamento all’inizio di un elemento lessicale (es.: <ss>si).

- INTERIEZIONI E SIMILI

1. Le segnalazioni di assenso possono essere annotate simbolicamente: <eh>, <ah>, <mh>, <ahah>. Espressioni di dubbio possono essere invece annotate: <mhmmh>.
2. Le esclamazioni (sorpresa, soddisfazione etc.) sono seguite da '!' (<ah!>, <oh!>, <eh!>).

**Fenomeni vocali non verbali** (indicati tra ganci <>)

Sono previsti i seguenti tipi di annotazione: <risata>, <tosse>, <respiro>, <inspirazione>, <click>, <schiarimento-voce>; altri fenomeni non classificabili sotto queste etichette sono annotati genericamente come <fenomeno-vocalico>.

**Fenomeni non vocali** (indicati tra ganci <>)

Tutti gli eventi acustici non prodotti dal parlante (come rumori di strada, di fondo, di carta, etc.) sono stati annotati indistintamente come <RUMORE> (es.: *ma che stai <RUMORE> dicendo?*, oppure {<RUMORE> *che mi stai dicendo?*})

## D.2. Codici per l'annotazione di testi in trascrizione ortografica

### Elementi linguistici lessicali e semi-lessicali

Simbolo (e descrizione)	Applicazione	Esempio d'uso
A, B, C... (maiuscole)	Nomi propri, sigle e acronimi	Maria, AIDS
- (trattino)	Sequenze di lettere pronunciate singolarmente	A-I-Di-Esse
+ (simbolo di addizione)	Frammenti di parole troncate (apposto a fine parola)	da+ ( <i>dato</i> )
' (apostrofo)	Forme con aferesi ed elisione	'ste ( <i>queste</i> ), m'ha ( <i>mi ha</i> )
* (asterisco)	Non-parole da lapsus ed errori (apposto a inizio parola)	*cimena (< cinema)
? (punto interrogativo)	Frase interpretata come interrog. (separato dal testo con uno spazio)	vieni ?
! (punto esclamativo)	Frase interpretata come esclamativa (separato dal testo con uno spazio)	vieni !

APPENDICI

, (virgola)	Confine sintattico-semantic percepito (separato dal testo con uno spazio)	no , mi sembra
-------------	---	----------------

*Fenomeni verbali non lessicali; fenomeni vocali non verbali; fenomeni non vocali*

Simbolo (e descrizione)	Applicazione	Esempio d'uso
<pb> <pl>	Pause vuote <pb>: breve, <pl>: lunga	la chiave <pb> l'hai vista ?
<eeh> <ehm>	Pause piene con vocalizzazione o nasalizzazione	allora<aa> <eeh> ce l'hai ? allora<aa> <ehm>
<vv> (v= vocale) <cc> (c= consonante)	Pause piene con allungamento di vocali o consonanti iniziali o finali	allora<aa>, non<nn>, <ss>senti
/ (barra obliqua - slash)	Falsa partenza senza pausa di interruzione (separata dal testo con uno spazio)	ho tro+ / ce l'hai la chiave ?
_ (trattino basso - underscore)	Interruzioni interne all'elemento lessicale	mon_tato
<risata>, <tosse>, <respiro>, <inspirazione>, <click>, <schiarimento-voce>	Fenomeni vocali non verbali prodotti dal parlante	ma dove ce l'hai <risata> ?
<fenomeno-vocale>	Altri non rientranti nelle etichette precedenti	dove <fenomeno-vocale> ?
<eh>, <ah>, <mh>, <ahah>	Segnalazioni di assenso da parte del locutore	- poi devi andare a destra - <mh>
<mhmf>	Segnalazioni di dubbio da parte del locutore	sulla scrivania <mhmf>
<oh!>, <ah!>, <eh!> <ahah!>	Esclamazioni (sorpresa, stupore, soddisfazione...)	<ah!> è così che rispondi ?
<RUMORE>	Evento non vocale, non comunicativo generico	dov'è <RUMORE> ?
<inintelligibile>	Parole o sequenze inintelligibili	ma <inintelligibile> dov'è?

APPENDICI

{ } (parentesi graffe)	Prima e dopo il testo (senza spazi bianchi) cui si sovrappone un evento di qualsiasi natura. Il testo è preceduto dall'annotazione dell'evento	{<risata> la banana ?} {<risata> non ce l'ho la chiave}
------------------------	--	--

*Commenti del trascrittore*

Simbolo (e descrizione)	Applicazione	Esempio d'uso
[gridato]	Commenti generici del trascrittore, sovrapposti o no al testo - se la parola interessata è una sola basta l'indicazione subito dopo; - per una sequenza, va invece segnalata l'esatta estensione	come hai detto [gridato] ? {[gridato] come hai detto ?}
[sussurrato]	“	gli ho preparato una sorpresa [sussurrato]
[cricchiato]	“	modestamente<ee> [cricchiato] l'ho salvato da un grave crisi
[mormorato]	“	aspetta<aa> [mormorato] ha detto {[mormorato] che me lo lasciava<aa>}
[dialettale] o altro	Parola o sequenza in un codice diverso (in questo caso dialettale)	{[dialettale] nun me ne pò ffregà' dde meno} <sup>268</sup> ho detto guaglió' [dialettale]

268 Si noti la necessità di dettagliare tutti i casi in cui si manifesta un *RF* in varietà non-standard, sub-standard o con uno standard distinto da quello dell'italiano (cfr. cap. III). Si noti ancora il *pò* del romanesco, senza dittongazione, corrispondente al *può* dell'italiano standard (cfr. §VI.8).

