

# Corpora e Lessici di Italiano Parlato e Scritto

## **W1- a1**

(Definizione Caratteristiche del Corpus)

**Title:** **Criteria per la predisposizione delle liste di lettura**

**Document No:** CLIPS/W1a1/DCC/002/v3

**Document Name:** LEDL002.rtf

**Status:** **pubblico**

**Date:** 10/3/2006

**Authors:** *Salvatore De Masi*  
*Università di Lecce*

## *Presentazione*

Il *corpus c)* è costituito da 16 ore di “materiale letto in laboratorio da più parlanti”. I testi sottoposti alla lettura si possono suddividere in due tipologie: *liste di frasi* e *liste di parole*.

Per quanto concerne le *liste di frasi*, esse consistono in testi atti a garantire la copertura lessicale di base, rappresentata dalle frequenze alte e medie dei lessici di frequenza correnti, come il LIF e il LIP, cui sono stati aggiunti, nel corso del lavoro, anche il VELI e il *Lessico elementare* di Marconi et alii (Bologna 1994). In questo modo è stato possibile identificare le parole con la massima probabilità di essere d’uso più frequente nell’italiano contemporaneo, tenendo conto degli strumenti oggi a disposizione e delle rispettive, in parte diverse, metodologie di raccolta e analisi dei dati.

Sulla base dell’elenco dei lemmi più frequenti (v. § 1) è stato costruito un primo corpus costituito da 20 brevissimi testi, nei quali sono distribuite 100 parole riconducibili ai 70 lemmi d’uso più frequente e ad altri 30 di rango compreso nell’intervallo 71-240. Il testo così ottenuto è stato successivamente testato in area salentina e sottoposto ai conseguenti aggiustamenti (senza, ovviamente, cambiarne i requisiti fondamentali).

Per quanto concerne, invece, le *liste di parole*, esse sono state realizzate sulla base del materiale utilizzato per la costituzione del corpus dialogico, che consiste in due esempi (A e B) di *mappe* e *vignette* utilizzate rispettivamente per la registrazione dei dialoghi *map-task* e *test differenze*.

Nel presente documento saranno specificati, per quanto concerne la costituzione delle *liste di frasi*, i criteri utilizzati per la selezione dei lemmi, i 240 lemmi più usati nella lingua italiana contemporanea, con misure di dispersione e indice d’uso<sup>1</sup>, inoltre sarà presentata la lista completa dei testi sottoposti alla lettura. Infine, anche per quanto riguarda la costituzione delle *liste di parole*, saranno specificati i criteri utilizzati e saranno presentate le liste complete di parole sottoposte alla lettura.

---

<sup>1</sup> L’indice d’uso è ottenuto moltiplicando la frequenza di ogni lemma per la misura della sua dispersione.

## 1 Liste di frasi<sup>2</sup>

### 1.1 Criteri per la selezione dei lemmi

L'elenco delle parole ad alta frequenza è stato ottenuto attraverso il confronto tra i seguenti lessici di frequenza:

- *Lessico di frequenza dell'italiano parlato*<sup>3</sup> (LIP);
- *Lessico di frequenza della lingua italiana contemporanea*<sup>4</sup> (LIF);
- *Vocabolario Elettronico della Lingua Italiana*<sup>5</sup> (VELI);
- *Lessico elementare*<sup>6</sup> (LE).

I differenti criteri di classificazione delle parole e di selezione dei testi campione non consentono una utilizzazione diretta dei risultati delle ricerche lessicografiche indicate. Alcuni esempi:

**NIENTE:** per il LIP è un pronome, mentre per il LIF è un sostantivo;

**MOLTO - SOLO - CERTO:** nel LIF sono unificate le occorrenze di tipo avverbiale con quelle aggettivali;

**FINE:** Il LIF non distingue le occorrenze maschili (*il fine*) da quelle femminili (*la fine*);

**ANDARE:** Il LIP registra una accezione come verbo con *uso* pari a 2593 e *rango* 26 ed una come ausiliare, con *rango* 224 ed *uso* 167; diversamente gli altri lessici;

**SIGNORA - SITUAZIONE -PRATICAMENTE - ATTIMO:** questi ed altri lemmi occupano posizioni molto differenti nelle diverse liste ordinate per *uso*.

---

<sup>2</sup> La sigla identificativa utilizzata all'interno dell'intero corpus CLIP per le liste di frasi è LF.

<sup>3</sup> T. DE MAURO-F. MANCINI-M. VEDOVELLI-M. VOGHERA, *Lessico di frequenza dell'italiano parlato*, ETASLIBRI, Milano 1993.

<sup>4</sup> U. BORTOLINI-C. TAGLIAVINI-A. ZAMPOLLI, *Lessico di frequenza della lingua italiana contemporanea*, Garzanti, Milano 1972.

<sup>5</sup> AA. VV., *Vocabolario Elettronico della Lingua Italiana*, IBM Italia, Milano 1989.

La procedura di selezione si è, perciò, sviluppata in quattro passi:

1. Si sono presi i primi duecentocinquanta lemmi, secondo il loro rango di uso, di ciascun vocabolario, e si sono fuse le quattro liste così ottenute.
2. Dall'elenco risultante sono state eliminate le *parole funzione* (congiunzioni, esclamazioni, preposizioni, articoli, verbi ausiliari); in un momento successivo sono stati eliminati avverbi, possessivi, indefiniti e numerali, sia per le incertezze classificatorie riscontrate, sia perché gli avverbi di modo (gli unici che potevano essere considerati *parole contenuto*) sono abbastanza vicini, per il *rango*, ai corrispondenti aggettivi.
3. Si è calcolato l'indice d'*uso* di ciascun lemma utilizzando i corrispondenti indici riportati nei diversi vocabolari e si sono scelte le settanta parole con indice più alto.
4. Si è verificato che i lemmi così ottenuti fossero tra quelli che il DIB<sup>7</sup> considera ad *alta disponibilità*.

### *1.2 I 240 lemmi più usati nella lingua italiana contemporanea, con misure di dispersione e indice d'uso*

Nella tabella che segue è riportato il risultato delle operazioni ora elencate. La tabella riporta, per ogni lemma:

- il rango
- la categoria grammaticale (CG),
- la misura della DISPERSIONE,
- l'indice d'uso (USO)
- il riferimento all'item del corpus c in cui il lemma è stato successivamente inserito (cfr. § 3).

---

<sup>6</sup> L. MARCONI-M. OTT-E. PESENTI-D. RATTI-M. TAVELLA, *Lessico elementare*, Zanichelli, Bologna 1994.

<sup>7</sup> T. DE MAURO-G.G. MORONI, *DIB. Dizionario di base della lingua italiana*, Paravia, Torino 1996.

<i>N</i>	<i>LEMMA</i>	<i>2</i>	<i>DISPERSIONE</i>	<i>USO</i>	<i>TESTO N.</i>
1	fare	v	0,23958085	25153,3543	3
2	dire	v	0,18962808	15128,1489	17
3	andare	v	0,32805829	14759,3427	4
4	potere	v	0,10150762	9830,09965	16
5	vedere	v	0,31108265	9492,37605	17
6	volere	v	0,29321071	9149,93359	16
7	cosa	s	0,30261032	8269,43244	4
8	dovere	v	0,11917641	7918,79630	2
9	stare	v	0,17151854	7023,51304	2
10	sapere	v	0,29386009	6105,23733	4
11	venire	v	0,11695332	5849,77116	16
12	casa	s	0,28031511	5585,55904	4
13	giorno	s	0,14294749	5098,79414	18
14	dare	v	0,14851354	4891,59061	5
15	volta	s	0,16521106	4433,60423	8
16	anno	s	0,04333141	4411,65794	9
17	mettere	v	0,17809595	4288,37249	15
18	prendere	v	0,21296734	4146,04834	15
19	trovare	v	0,17627250	3870,41533	10
20	bello	ag	0,51567609	3838,17716	15
21	grande	ag	0,09763900	3818,66132	6
22	chiamare	v	0,31054808	3712,91287	14
23	portare	v	0,19915567	3272,32687	12
24	parlare	v	0,15042038	3217,04080	8
25	sentire	v	0,42241229	3171,89393	14
26	arrivare	v	0,19531794	3030,94382	12
27	tempo	s	0,08527014	2832,41854	15
28	pensare	v	0,18410461	2751,44340	17
29	guardare	v	0,41713691	2738,50382	11
30	piccolo	ag	0,15517766	2666,10747	12
31	giocare	v	0,4359112	2618,51858	6
32	uomo	s	0,13498008	2525,74731	13
33	bambino	s	0,44769851	2524,57190	6
34	parte	s	0,05822576	2440,06722	11
35	scuola	s	0,24943620	2407,55826	2
36	certo	ag	0,15121865	2388,04506	19
37	capire	v	0,27344109	2349,40588	16
38	piacere	v	0,44200431	2324,50066	20
39	passare	v	0,13598634	2156,74345	1

40	<b>amico</b>	s	0,25311776	2122,13936	10
41	<b>buono</b>	ag	0,23470870	2100,87759	19
42	<b>tornare</b>	v	0,21127859	2039,68357	18
43	<b>finire</b>	v	0,25421674	1992,04243	18
44	<b>chiedere</b>	v	0,10711400	1899,55968	3
45	<b>lasciare</b>	v	0,16549980	1872,96126	8
46	<b>persona</b>	s	0,14985012	1830,71900	19
47	<b>ora</b>	s	0,14397597	1828,63891	18
48	<b>occhio</b>	s	0,32227327	1770,24712	14
49	<b>sembrare</b>	v	0,11182514	1738,09822	20
50	<b>padre</b>	s	0,25903432	1735,27092	6
51	<b>uscire</b>	v	0,18322981	1721,99383	14
52	<b>cercare</b>	v	0,16285978	1719,96214	11
53	<b>acqua</b>	s	0,23657269	1675,40785	12
54	<b>ragazzo</b>	s	0,30346547	1662,08038	17
55	<b>credere</b>	v	0,18274154	1654,54191	13
56	<b>signore</b>	s	0,35786625	1652,26847	4
57	<b>conoscere</b>	v	0,20254744	1641,64702	7
58	<b>riuscire</b>	v	0,12983361	1595,00600	16
59	<b>entrare</b>	v	0,14255279	1567,93821	7
60	<b>ricordare</b>	v	0,13798522	1563,92453	15
61	<b>diventare</b>	v	0,10543682	1561,09763	19
62	<b>mano</b>	s	0,15409753	1544,51957	20
63	<b>momento</b>	s	0,09012574	1540,51935	10
64	mondo	s	0,11904857	1539,05998	13
65	<b>lavoro</b>	s	0,05350114	1523,87299	9
66	<b>nuovo</b>	ag	0,03056008	1521,21965	9
67	scrivere	v	0,13518212	1518,09520	10
68	<b>alto</b>	ag	0,09972404	1497,25675	20
69	<b>vita</b>	s	0,08443251	1496,81955	8
70	<b>terra</b>	s	0,27830078	1496,14502	18
71	tenere	v	0,10505709	1483,61629	
72	<b>mangiare</b>	v	0,50783347	1452,91157	20
73	vivere	v	0,17130990	1433,52131	
74	sera	s	0,22679060	1410,18400	
75	strada	s	0,14255822	1403,48570	
76	nome	s	0,12229649	1386,35303	
77	punto	s	0,06245416	1361,43841	
78	madre	s	0,34845481	1335,62731	
79	macchina	s	0,20759579	1334,42574	
80	<b>vero</b>	ag	0,07672130	1323,82616	13
81	<b>modo</b>	s	0,06842165	1314,17474	11

82	lavorare	v	0,13870112	1312,25132	
83	fratello	s	0,31896221	1291,79696	
84	paese	s	0,04489910	1279,44498	
85	cominciare	v	0,09441527	1276,77779	
86	rispondere	v	0,10873682	1264,17435	
87	problema	s	0,05583107	1258,54410	
88	<b>aspettare</b>	v	0,32066503	1254,12094	15
89	posto	s	0,11531573	1218,65663	
90	lungo	ag	0,09519039	1213,39202	
91	morire	v	0,29381934	1209,06658	
92	mamma	s	0,36478721	1192,85420	
93	gioco	s	0,19831208	1188,08772	
94	nero	ag	0,22247658	1182,01807	
95	famiglia	s	0,11074392	1175,98969	
96	colore	s	0,33941823	1172,35059	
97	<b>gente</b>	s	0,20803860	1165,64031	13
98	mare	s	0,25747308	1165,06572	
99	correre	v	0,32950097	1097,56773	
100	storia	s	0,10297412	1073,29934	
101	leggere	v	0,17771326	1071,96639	
102	aiutare	v	0,33128377	1069,71531	
103	parola	s	0,13054209	1067,31217	
104	figlio	s	0,12039429	1067,05466	
105	esempio	s	0,06917516	1060,87039	
106	aprire	v	0,10702937	1048,56678	
107	città	s	0,09500276	1037,05012	
108	bianco	ag	0,18706825	1032,80381	
109	perdere	v	0,13262690	1029,31742	
110	cane	s	0,49184440	1027,95480	
111	ultimo	ag	0,03895111	1010,15810	
112	donna	s	0,11817116	1009,18171	
113	notte	s	0,26565541	1005,24008	
114	fiore	s	0,39289241	998,339635	
115	cadere	v	0,28092101	986,313682	
116	<b>bisognare</b>	v	0,10778252	985,779006	13
117	partire	v	0,08508496	978,391985	
118	<b>fine</b>	sf	0,06073092	977,646344	16
119	diverso	ag	0,05194632	932,332631	
120	<b>mese</b>	s	0,04573694	930,106530	1
121	paura	s	0,31459178	927,416585	
122	vecchio	ag	0,09219653	918,646275	
123	<b>numero</b>	s	0,07577580	913,022704	20

124	animale	s	0,46505761	911,047870	
125	pieno	ag	0,13352093	910,078677	
126	sole	s	0,40691998	901,734679	
127	<b>servire</b>	v	0,13496632	892,937183	20
128	aria	s	0,26665404	882,358219	
129	rosso	ag	0,22954899	880,320378	
130	compagno	s	0,33032329	875,0264	
131	caso	s	0,03830933	870,426468	
132	<b>cambiare</b>	v	0,11365418	870,022795	9
133	testa	s	0,18273163	863,224257	
134	bravo	ag	0,50593655	862,621818	
135	<b>tipo</b>	s	0,08774598	846,573297	20
136	fermare	v	0,34068659	838,429713	
137	via	s	0,06172909	835,935358	
138	secondo	nu	0,05650017	816,427467	
139	letto	s	0,42884665	814,37979	
140	piede	s	0,24448544	809,980290	
141	<b>decidere</b>	v	0,06221865	798,203143	9
142	comprare	v	0,28122479	794,741276	
143	continuare	v	0,06204001	791,382385	
144	<b>importante</b>	ag	0,06761712	791,255631	19
145	tirare	v	0,33480624	790,812338	
146	scendere	v	0,16390415	790,181927	
147	felice	ag	0,31336496	789,993067	
148	preparare	v	0,19842657	788,745615	
149	porta	s	0,17272809	786,258284	
150	capello	s	0,46422291	782,215603	
151	fatto	s	0,04061485	782,120307	
152	classe	s	0,19300441	779,737842	
153	albero	s	0,47139953	772,152438	
154	dormire	v	0,53220041	758,917796	
155	succedere	v	0,27737895	743,098229	
156	sorella	s	0,46987789	733,949276	
157	caro	ag	0,30696708	731,809534	
158	salire	v	0,15021721	719,390231	
159	<b>bastare</b>	v	0,10661436	714,636088	5
160	gatto	s	0,45135334	711,784227	
161	parere	v	0,17647248	709,066435	
162	restare	v	0,06611686	699,185856	
163	nonno	s	0,42051535	696,793945	
164	genitore	s	0,35378117	690,934629	
165	verde	ag	0,16796513	689,328908	



166	bisogno	s	0,11551947	684,452859	
167	mandare	v	0,26777327	675,056429	
168	mezzo	s	0,09648228	674,025243	
169	divertire	v	0,48256736	671,251201	
170	cielo	s	0,44576191	668,642876	
171	seguire	v	0,04339797	647,194040	
172	guerra	s	0,07915095	646,346727	
173	ritornare	v	0,36359876	643,206206	
174	zio	s	0,41842177	638,093202	
175	povero	ag	0,26271119	636,549235	
176	spiegare	v	0,05351844	635,370968	
177	forza	s	0,05855264	629,206766	
178	rendere	v	0,05350346	626,365075	
179	chiudere	v	0,10811468	613,118394	
180	accordo	s	0,05773261	593,029393	
181	<b>pagare</b>	v	0,08174313	583,972967	20
182	<b>conto</b>	s	0,04893473	583,840311	20
183	veramente	av	0,24077164	576,648090	
184	rimanere	v	0,04945755	573,163596	
185	papà	s	0,36944275	567,833519	
186	stato	s	0,03571058	553,299748	
187	<b>trattare</b>	v	0,03451249	546,505382	20
188	presentare	v	0,04428961	537,321565	
189	giovane	ag	0,04846494	521,482789	
190	lira	s	0,01904569	518,480813	
191	amore	s	0,16646710	517,879164	
192	senso	s	0,05876617	517,377366	
193	giusto	ag	0,09368229	511,692669	
194	<b>possibile</b>	ag	0,03925456	510,662683	10
195	scusare	v	0,64157018	509,406723	
196	santo	ag	0,15570405	502,612688	
197	amare	v	0,17024515	490,476277	
198	ragione	s	0,07090396	479,948905	
199	gruppo	s	0,02125871	459,783384	
200	<b>discorso</b>	s	0,10876932	441,277158	19
201	contento	ag	0,5045617	438,968678	
202	domanda	s	0,05393668	431,817095	
203	<b>situazione</b>	s	0,04210026	431,317256	19
204	cugino	s	0,42618695	430,875013	
205	unico	ag	0,03744243	429,31493	
206	centro	s	0,03153423	424,608454	
207	moglie	s	0,12465055	419,075151	

208	<b>italiano</b>	ag	0,00938593	399,981428	19
209	dio	s	0,28289056	397,461249	
210	maggiore	ag	0,03217146	382,357825	
211	<b>legge</b>	s	0,02376305	367,875869	19
212	attimo	s	0,47026805	353,171309	
213	particolare	ag	0,02996372	331,938085	
214	riguardare	v	0,03686578	328,437256	
215	rapporto	s	0,02194442	322,539195	
216	scorso	ag	0,02527869	303,243267	
217	servizio	s	0,01965928	295,990243	
218	grazia	s	0,04213728	287,081341	
219	signora	s	0,52414304	272,554383	
220	rispetto	s	0,02437504	267,808565	
221	questione	s	0,04314211	264,202278	
222	<b>società</b>	s	0,01274681	241,297209	19
223	sistema	s	0,01463514	235,977089	
224	mercato	s	0,01590611	222,606126	
225	generale	ag	0,01725173	194,496100	
226	mattina	s	0,26963907	193,870492	
227	ricerca	s	0,01673748	191,878519	
228	pubblico	ag	0,00930180	189,329013	
229	<b>prezzo</b>	s	0,01724828	187,868336	20
230	<b>attività</b>	s	0,01591553	173,288369	9
231	governo	s	0,00893240	149,305135	
232	<b>nazionale</b>	ag	0,01138676	145,306453	20
233	<b>economico</b>	ag	0,01147269	135,917001	20
234	americano	ag	0,00731027	126,650437	
235	presidente	s	0,00503219	121,305964	
236	politica	s	0,01014249	115,908416	
237	<b>ministro</b>	s	0,00861073	103,914298	19
238	settore	s	0,00850338	102,584810	
239	risultato	s	0,00948159	102,552873	
240	<b>politico</b>	ag	0,00390543	93,3243520	19

## 2.1 Testi da sottoporre alla lettura

Con i lemmi evidenziati in grassetto nella colonna 2 della tabella precedente sono stati costruiti i 20 micro-testi che costituiscono le liste di frasi del corpus c. Il tempo di lettura varia, a seconda dello stile di lettura e dell'accuratezza, fra 2 e 3 minuti (tempo medio, durante i test, 2'35").

Si presenta di seguito l'elenco completo delle liste di frasi, con alcune avvertenze:

a. L'ordine dei micro-testi è: dai più brevi, monoproposizionali, ai più complessi: negli ultimi items si arriva a una decina di proposizioni per periodo.

b. Anche i periodi più lunghi sono costruiti in modo semplice (prevalde di gran lunga la coordinazione), sia per esigenze di 'naturalzza' sia per non introdurre una variabile spuria: la variazione del profilo intonativo – con conseguenze anche sulla realizzazione fonetica – dovuta alla microprogettazione dell'attività di lettura ad alta voce di periodi molto articolati e complessi. Si è constatato infatti che la resa fonica di proposizioni subordinate oltre il primo grado è legata al processo di lettura-interpretazione del testo, che è estraneo all'area di indagine della presente ricerca.

La lista seguente presenta i 20 brevi testi che costituiscono le liste di frasi del corpus c), con l'evidenziazione degli items lessicali di alta frequenza (sottolineati)

1. Un mese di vacanza passa in fretta
2. Maria dovrebbe stare più attenta, a scuola
3. Luigi fa tutto quello che gli si chiede
4. Quei signori non sanno mai cosa fare né dove andare
5. Ora basta! Non ti darò più un soldo
6. Nel grande parco un bambino giocava con suo padre
7. La ragazza che è appena entrata, non la conosco
8. Se mi lasci un'altra volta solo, non ti parlerò per tutta la vita
9. Quest'anno ho fatto un lavoro nuovo, però ho deciso di non cambiare più attività
10. E' possibile che non trovi mai un momento per scrivere a un amico lontano
11. Luisa fingeva di guardare da un'altra parte, ma cercava il modo per farsi notare
12. Ho chiesto a Gianna di portare subito dell'acqua in un recipiente piccolo, ma non è ancora arrivata
13. Non bisogna credere che sia vero tutto quello che dice la gente: tu

- non conosci ancora gli uomini, non conosci il mondo
14. Chiamai il medico perché avevo male agli occhi, ma quando uscì mi accorsi di non sentire neanche i rumori
15. Dopo tanto tempo non ricordo più dove ho messo quella bella foto, ma se aspetti un po' la cerco e te la prendo
16. Alla fine, non sono riuscito a capire se Marco non voleva o non poteva venire a casa di Sandra
17. Quel ragazzo non dice mai la verità, ma tu non fargli vedere che pensi che sia un bugiardo
18. Questo tormento durerà ancora qualche ora, forse un giorno; poi tutto finirà e tu potrai tornare a casa, nella tua terra
19. Lucio era certo che sarebbe diventato una persona importante, un uomo politico o magari un ministro: aveva a cuore il bene della società, rispettava la legge, se teneva un discorso, trovava le parole adatte ad ogni situazione. Si sentiva proprio un buon italiano
20. Stefano era un bel tipo: gli piaceva farsi servire per primo e mangiava per due, ma quando si trattava di pagare, sembrava non vedere la mano che gli porgeva il conto: il prezzo gli pareva sempre troppo alto. Eppure lavorava in una azienda che, dal punto di vista economico, era la numero uno su scala nazionale.

La lista seguente, invece presenta i 20 testi che costituiscono il *corpus di frasi*, così come sono state presentate al lettore (senza sottolineature).

1. Un mese di vacanza passa in fretta
2. Maria dovrebbe stare più attenta, a scuola
3. Luigi fa tutto quello che gli si chiede
4. Quei signori non sanno mai cosa fare né dove andare
5. Ora basta! Non ti darò più un soldo
6. Nel grande parco un bambino giocava con suo padre
7. La ragazza che è appena entrata, non la conosco
8. Se mi lasci un'altra volta solo, non ti parlerò per tutta la vita

9. Quest'anno ho fatto un lavoro nuovo, però ho deciso di non cambiare più attività
10. E' possibile che non trovi mai un momento per scrivere a un amico lontano?
11. Luisa fingeva di guardare da un'altra parte, ma cercava il modo per farsi notare
12. Ho chiesto a Gianna di portare subito dell'acqua in un recipiente piccolo, ma non è ancora arrivata
13. Non bisogna credere che sia vero tutto quello che dice la gente: tu non conosci ancora gli uomini, non conosci il mondo
14. Chiamai il medico perché avevo male agli occhi, ma quando uscì mi accorsi di non sentire neanche i rumori
15. Dopo tanto tempo non ricordo più dove ho messo quella bella foto, ma se aspetti un po' la cerco e te la prendo
16. Alla fine, non sono riuscito a capire se Marco non voleva o non poteva venire a casa di Sandra
17. Quel ragazzo non dice mai la verità, ma tu non fargli vedere che pensi che sia un bugiardo
18. Questo tormento durerà ancora qualche ora, forse un giorno; poi tutto finirà e tu potrai tornare a casa, nella tua terra
19. Lucio era certo che sarebbe diventato una persona importante, un uomo politico o magari un ministro: aveva a cuore il bene della società, rispettava la legge, se teneva un discorso trovava le parole adatte ad ogni situazione. Si sentiva proprio un buon italiano
20. Stefano era un bel tipo: gli piaceva farsi servire per primo e mangiava per due, ma quando si trattava di pagare, sembrava non vedere la mano che gli porgeva il conto: il prezzo gli pareva sempre troppo alto. Eppure lavorava in una azienda che, dal punto di vista economico, era la numero uno su scala nazionale.

### 3 *Liste di parole*

#### 3.1 *Criteri di costituzione delle liste*

Il corpus di parlato letto costituito da *liste di parole* è stato realizzato sulla base del materiale utilizzato per la raccolta del parlato dialogico di CLIP, che consiste in *mappe*, con l'uso delle quali sono stati registrati i dialoghi *map-task*, e *vignette*, che sono state invece utilizzate per la registrazione dei dialoghi del *test differenze*.

Sia le mappe, sia le vignette sono di due tipi, A e B.

A partire dall'osservazione delle mappe e delle vignette, è stato chiesto ad alcuni operatori di stilare una lista di nomi possibili da assegnare a tutti gli oggetti presenti nelle mappe utilizzate per i *map-task*, e a tutti gli oggetti che si potevano individuare nelle vignette e che potevano essere di qualche interesse per lo svolgimento del test delle differenze. Dal confronto tra gli elenchi proposti, sono state create liste di parole, da sottoporre alla lettura, che potessero essere presenti nei dialoghi *map-task* e test differenze, al fine di operare un confronto tra parlato letto e parlato spontaneo.

#### 3.2 *Liste di parole (map-task)*<sup>8</sup>

A partire dalle mappe A e B sono state realizzate 2 liste di parole che fanno riferimento alle possibili e più frequenti designazioni degli oggetti raffigurati nelle mappe.

La lista costruita sulla base della mappa A è la seguente:

1. *televisione*
2. *televisore*
3. *sci*
4. *torta*
5. *dolce*
6. *macchina*
7. *automobile*
8. *macchina rossa*
9. *rossa*
10. *blu*
11. *macchina blu*

---

<sup>8</sup> La sigla identificativa delle liste di parole relative agli oggetti presenti nelle mappe del *map-task* è LM.

12. *automobile rossa*
13. *automobile blu*
14. *maggiolino*
15. *pettine*
16. *farfalla*
17. *camion*
18. *camion rosso*
19. *bicicletta*
20. *occhiali*
21. *dado*
22. *lupo*
23. *stelle*

La lista costituita sulla base della mappa B è la seguente:

1. *limoni*
2. *frigorifero*
3. *gatto*
4. *freccia*
5. *treno*
6. *sedia*
7. *pipa*
8. *barca*
9. *forchetta*
10. *bar*
11. *mano*
12. *cuore*
13. *cuore rosso*
14. *rosso*
15. *giallo*
16. *cuore giallo*
17. *gelato*
18. *cono*

A questo punto è opportuna una precisazione sulla modalità di presentazione delle liste: al fine di evitare, nella parte finale della lista, l'effetto prosodico caratteristico degli elenchi, e al fine di ottenere, durante tutta la lettura della lista, un andamento prosodico analogo, sono state ripetute le prime 5 parole di ciascuna lista.

La lista che si riferisce alla mappa A è costituita da 28 parole (23 + 5 ripetizioni), la lista che si riferisce alla mappa B è invece costituita da 23 parole (18 + 5 ripetizioni).

Qui di seguito si riportano le liste complete così come sono state presentate ai lettori.

Lista di parole (Mappa A)

*televisione*

*televisore*

*sci*

*torta*

*dolce*

*macchina*

*automobile*

*macchina rossa*

*rossa*

*blu*

*macchina blu*

*automobile rossa*

*automobile blu*

*maggiolino*

*pettine*

*farfalla*

*camion*

*camion rosso*

*bicicletta*

*occhiali*

*dado*

*lupo*



*stelle*  
*televisione*  
*televisore*  
*sci*  
*torta*  
*dolce*

Lista di parole (Mappa B)

*limoni*  
*frigorifero*  
*gatto*  
*freccia*  
*treno*  
*sedia*  
*pipa*  
*barca*  
*forchetta*  
*bar*  
*mano*  
*cuore*  
*cuore rosso*  
*rosso*  
*giallo*  
*cuore giallo*  
*gelato*  
*cono*  
*limoni*  
*frigorifero*  
*gatto*  
*freccia*  
*treno*

### 3.3 Liste di parole (test-differenze)<sup>9</sup>

Sulla base delle due vignette A e B utilizzate per i *test delle differenze*, sono state realizzate, analogamente, 2 liste di parole che fanno riferimento alle possibili e più frequenti designazioni di tutti gli oggetti presenti nelle vignette che potessero essere di interesse rilevante ai fini dello svolgimento del compito previsto dal confronto tra due vignette parzialmente differenti.

La lista costruita sulla base della vignetta A è la seguente:

1. *panchina*
2. *albero*
3. *palazzo*
4. *casa*
5. *cavallo*
6. *base*
7. *statua*
8. *cavaliere*
9. *spada*
10. *chioma*
11. *finestre*
12. *macchina*
13. *automobile*
14. *specchietto*
15. *prato*
16. *cane*
17. *guinzaglio*
18. *corda*
19. *televisore*
20. *televisione*
21. *schermo*
22. *antenna*
23. *camino*

---

<sup>9</sup> La sigla identificativa delle liste di parole relative agli oggetti presenti nelle vignette del test differenze è LT.

24. *tronco*
25. *fari*
26. *luci*
27. *bottoni*
28. *pulsante*
29. *ruote*
30. *uomo*
31. *bocca*
32. *naso*
33. *occhi*
34. *maglia*
35. *pantaloni*
36. *comignolo*
37. *foglie*
38. *cappello*
39. *tetto*
40. *coda*
41. *zampe*
42. *zampa*
43. *scarpa*
44. *piede*

La lista costruita sulla base della vignetta B è la seguente:

1. *salvagente*
2. *pallini*
3. *bambino*
4. *occhi*
5. *naso*
6. *bocca*
7. *mano*
8. *dita*
9. *mare*

10. *uovo*
11. *tondi*
12. *ragazzo*
13. *onda*
14. *nuvola*
15. *barca*
16. *vela*
17. *spiaggia*
18. *tappo*
19. *costume*
20. *becco*
21. *bandiera*
22. *sasso*
23. *pietra*
24. *masso*
25. *ciambella*

Anche nel caso delle liste di parole relative alle vignette del test differenze, come nel già citato caso delle liste relative alle mappe del map-task, per evitare nella parte finale della lista, l'effetto prosodico caratteristico degli elenchi, e al fine di ottenere, durante tutta la lettura della lista, un andamento prosodico analogo, sono state ripetute le prime 4 parole nel caso della lista relativa alla vignetta A e le prime 5 parole nel caso della lista relativa alla vignetta B.

La lista che si riferisce alla vignetta A risulta dunque costituita da 48 parole (44 + 4 ripetizioni), la lista che si riferisce alla vignetta B è invece costituita da 30 parole (25 + 5 ripetizioni).

Qui di seguito si riportano le liste complete così come sono state presentate ai lettori.

Lista di parole (Vignetta A)

*panchina*  
*albero*

*palazzo*  
*casa*  
*cavallo*  
*base*  
*statua*  
*cavaliere*  
*spada*  
*chioma*  
*finestre*  
*macchina*  
*automobile*  
*specchietto*  
*prato*  
*cane*  
*guinzaglio*  
*corda*  
*televisore*  
*televisione*  
*schermo*  
*antenna*  
*camino*  
*tronco*  
*fari*  
*luci*  
*bottoni*  
*pulsante*  
*ruote*  
*uomo*  
*bocca*  
*naso*  
*occhi*  
*maglia*

*pantaloni*  
*comignolo*  
*foglie*  
*cappello*  
*tetto*  
*coda*  
*zampe*  
*zampa*  
*scarpa*  
*piede*  
*panchina*  
*albero*  
*palazzo*  
*casa*

*Lista di parole (Vignetta B)*

*salvagente*  
*pallini*  
*bambino*  
*occhi*  
*naso*  
*bocca*  
*mano*  
*dita*  
*mare*  
*uovo*  
*tondi*  
*ragazzo*  
*onda*  
*nuvola*  
*barca*  
*vela*

*spiaggia*

*tappo*

*costume*

*becco*

*bandiera*

*sasso*

*pietra*

*masso*

*ciambella*

*salvagente*

*pallini*

*bambino*

*occhi*

*naso*