

# Indice

---

<b>Introduzione</b>	<b>7</b>
<hr/>	
<b>I. I fondamenti della linguistica dei corpora</b>	<b>13</b>
1. Premesse teoriche	13
2. Il corpus nella tradizione	22
3. Il corpus elettronico	29
<hr/>	
<b>II. Struttura di un corpus</b>	<b>43</b>
1. Il corpus come campione	43
2. I tratti definitori di un corpus	52
<hr/>	
<b>III. Codifica e annotazione</b>	<b>73</b>
1. Codifica del carattere e markup testuale	73
2. Le unità linguistiche e la loro annotazione	84
<hr/>	
<b>IV. L'analisi dei corpora: strumenti di base</b>	<b>109</b>
1. La distribuzione del lessico	109
2. Analisi delle concordanze	118
3. Associazioni lessicali: le collocazioni	126

---

<b>V. La situazione italiana</b>	<b>141</b>
1. I centri italiani per la linguistica dei corpora e le principali risorse scritte	141
2. I corpora di italiano parlato	155

---

## **Quadri**

2.1. Corpora e questioni legali	69
3.1. La tokenizzazione	87
3.2. Le espressioni polirematiche	91
3.3. Il formato di trascrizione CHAT-LABLITA	100
3.4. La pragmatica e lo studio del parlato	104
4.1. Vocabolario del corpus e <i>type/token ratio</i>	111
4.2. Classi aperte e classi chiuse	115
4.3. Ricerche grammaticali <i>corpus-based</i>	137

---

<b>Appendice</b>	<b>179</b>
------------------	------------

---

<b>Riferimenti bibliografici e sitografia</b>	<b>187</b>
---	------------

---

<b>Indice analitico</b>	<b>207</b>
-------------------------	------------