

Adattamento dei testi indice di leggibilità Gulpease

A fronte di particolari disabilità come, ad esempio, quelle:

- intellettive,
- uditive (in linea generale è da tenere presente che la comprensione della lingua da parte di alunni con disabilità uditiva può essere compromessa e può essere povera dal punto di vista lessicale),

si può avere la necessità di presentare dei contenuti **adattati**.

Adattamento

- In linea generale un percorso di adattamento prevede le seguenti fasi:
 - sostituzione
 - semplificazione

Sostituzione

- Con la **sostituzione**, in linea con l'UDL
 - 1° princ.: Fornire molteplici modalità di rappresentazione e
 - la prima linea guida: Fornire differenti opzioni per la percezione
- non si modifica la difficoltà del testo che deve essere letto, ma si utilizzano canali (modalità) alternativi per la fruizione dell'informazione.

Semplificazione

- Qualora la sostituzione non dia esito positivo si può passare alla **semplificazione**: processo graduale che mira a facilitare la comprensione del testo, modificandolo per adattarlo allo studente; si può procedere:
- «integrando il testo di supporti didattici come la sua divisione in blocchi di testo, la sottolineatura delle parole chiave, l'accompagnamento di immagini [video], la sintesi [visualizzazione] mediante mappe concettuali, ecc; [...]»

Semplificazione

- utilizzando un approccio [multimediale](#); in questo caso la semplificazione deriva dall'uso sinergico di più codici che restituiscono un significato aumentato che facilita la comprensione del testo;
- infine si può anche semplificare il testo rendendolo più **leggibile**.

Leggibilità e comprensibilità dei testi

- La leggibilità riguarda la decifrazione della superficie del testo; si fa riferimento all'impianto linguistico: un testo risulta più o meno chiaro e comprensibile sulla base di caratteristiche linguistiche combinate fra loro, a prescindere dalla complessità degli argomenti contenuti.
- Ovviamente, però, tanto più un testo risulta di facile lettura, tanto maggiore è la probabilità che esso risulti di facile comprensione!

Leggibilità

- Si tiene conto, ad esempio:
 - dell'aspetto grafico (organizzazione del testo in capitoli, paragrafi e sottoparagrafi, la titolazione di queste partizioni, ..);
 - della lunghezza delle frasi (tanto più una frase è lunga e ricca di subordinazioni, tanto meno sarà di facile ed immediata lettura; paratassi e ipotassi);
 - della lunghezza delle parole all'interno di ogni singola frase (quanto più una parola è lunga, tanto maggiore è il carico di informazioni che essa trasmette e anche la difficoltà di “raccolgere” tutto il significato; la presenza di molte parole lunghe può rendere una frase troppo densa di significato e quindi di non facile lettura e comprensione).

Leggibilità del testo

- Alcune applicazioni software permettono di ricavare gli indici di leggibilità di un testo scritto; sono indicatori che individuano la maggiore o minore difficoltà nella comprensione del testo;
- uno di questi è l'indice Gulpease.

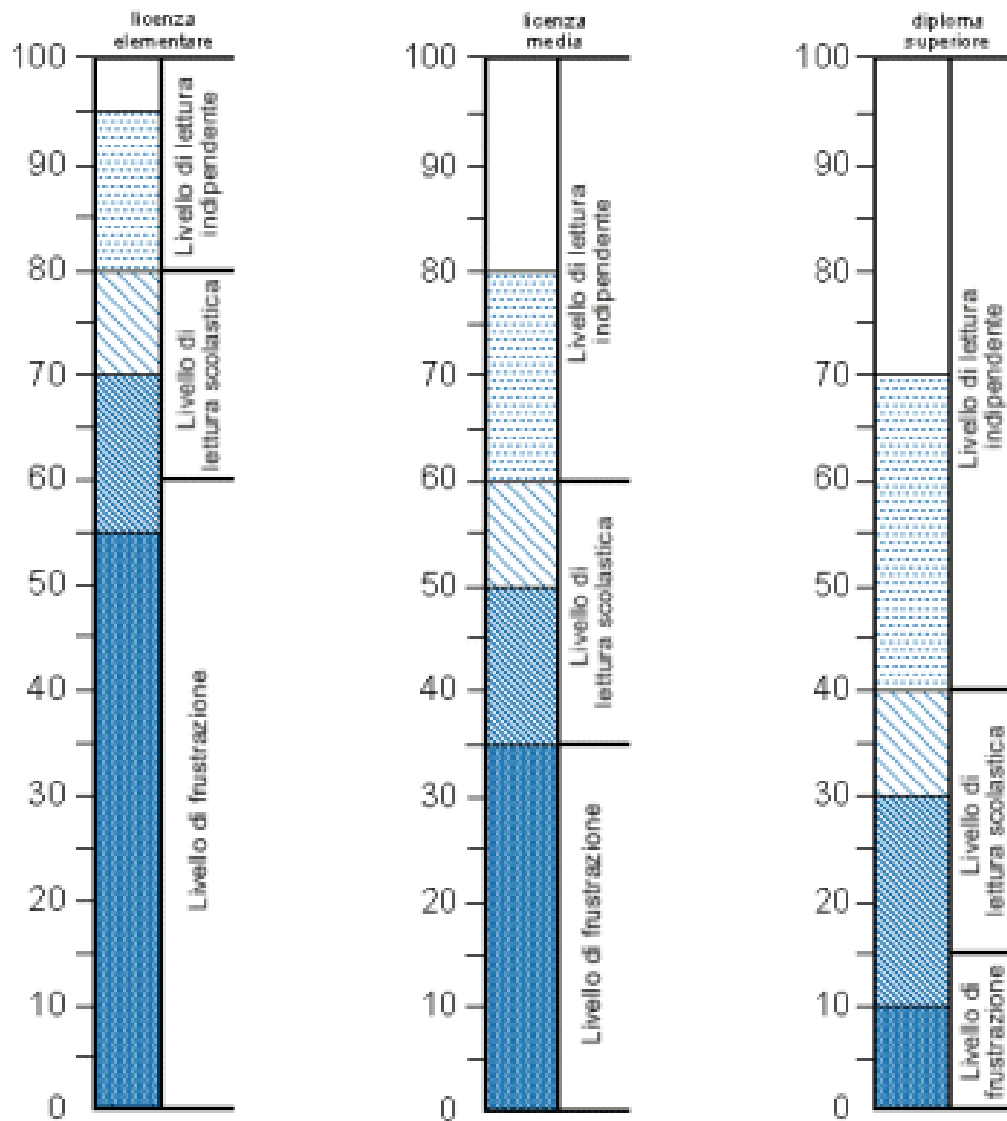
Indice GULPEASE

- Un gruppo di linguisti dell'Università La Sapienza di Roma riunito attorno a Tullio de Mauro ha costituito il GULP (Gruppo Universitario Linguistico Pedagogico); nasce l'indice GULPEASE;
- deriva da uno studio sperimentale su diversi tipi di lettori e di testi;
- tra i fattori considerati nella formula ci sono il numero di lettere per parola e il numero di parole per frase.

Indice GULPEASE

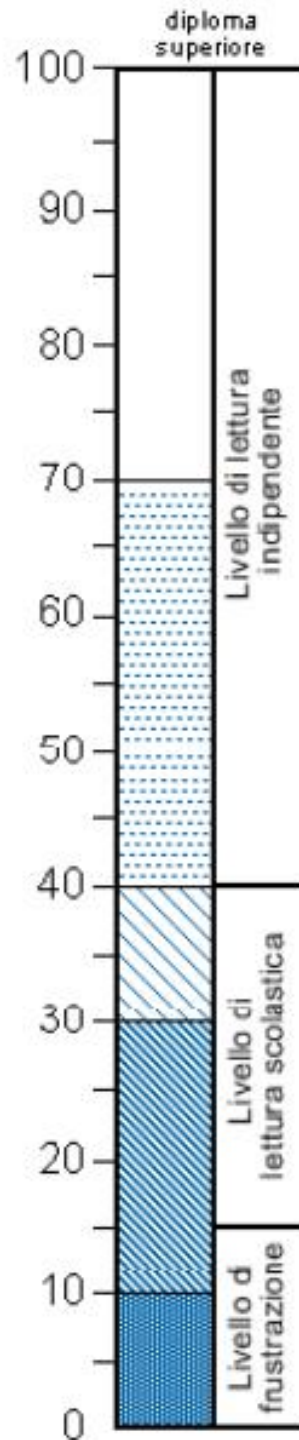
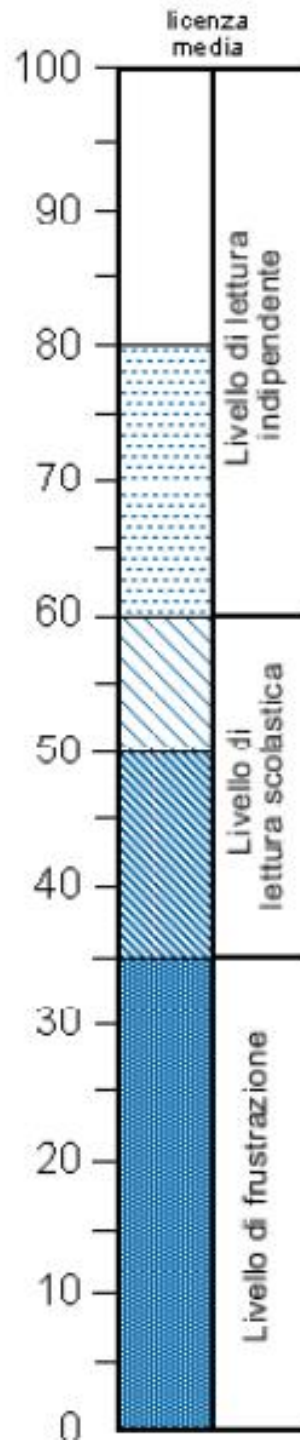
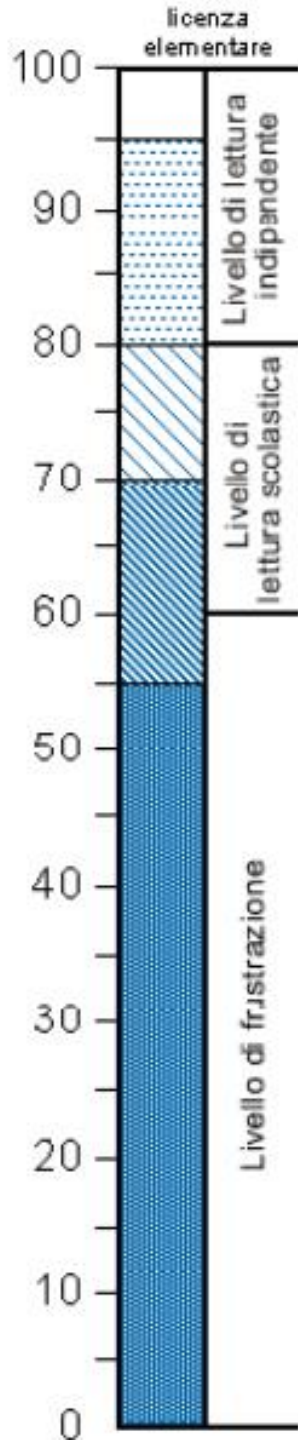
- I valori che si ottengono sono compresi in una scala che va da 0 a 100; più è alto l'indice, più è leggibile il testo;
- i lettori che hanno un'istruzione primaria leggono facilmente i testi che presentano un indice superiore a 80;
- i lettori che hanno un'istruzione secondaria di primo grado leggono facilmente i testi che presentano un indice superiore a 60;
- i lettori che hanno un'istruzione superiore leggono facilmente i testi che presentano un indice superiore a 40.

Indice Gulpease: scala dei valori

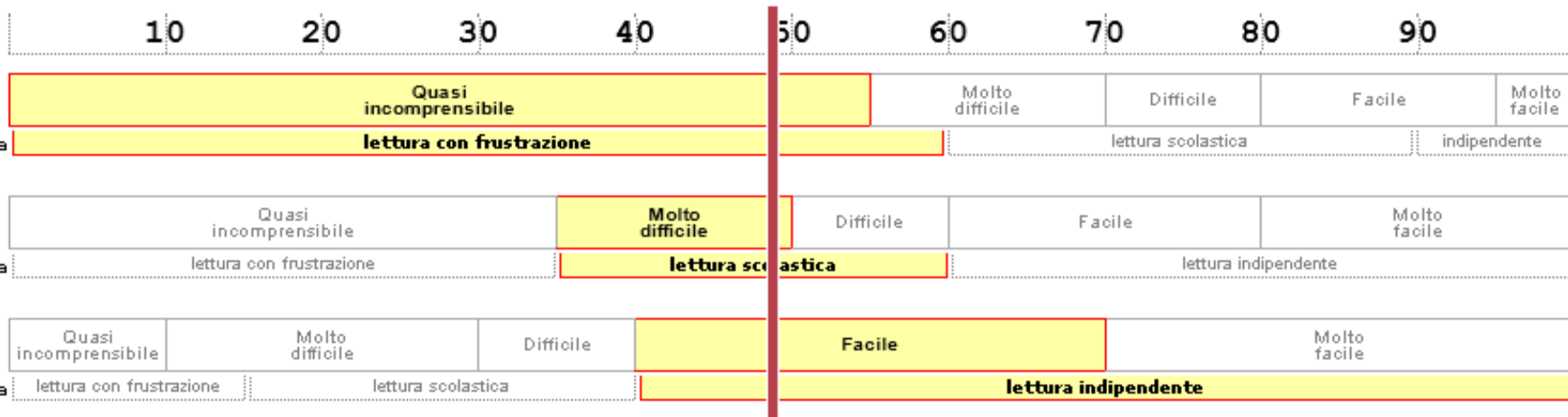


Soglie di Leggibilità





Leggibilità e livello di lettura correlati al grado di istruzione



Indice GULPEASE

- La formula per il suo calcolo è la seguente:

$$89 + \frac{300 * (\textit{numero delle frasi}) - 10 * (\textit{numero delle lettere})}{\textit{numero delle parole}}$$

- I risultati sono compresi tra 0 e 100, dove il valore "100" indica la leggibilità più alta e "0" la leggibilità più bassa. In generale risulta che testi con un indice
 - inferiore a 80 sono difficili da leggere per chi ha la licenza elementare
 - inferiore a 60 sono difficili da leggere per chi ha la licenza media
 - inferiore a 40 sono difficili da leggere per chi ha un diploma superiore

oppure, equivalentemente, la seguente:

Facilità di lettura = $89 - LP/10 + FR * 3$

LP = lettere * 100 / totale parole

FR = frasi * 100 / totale parole

- Si affianca all'indice GULPease, l'analisi del vocabolario di base (VdB); cioè la notorietà dei singoli termini usati.

Vocabolario di base:
FO + AU + AD

FO: fondamentale
AU: altro uso
AD: alta disponibilità

CO: comune

TS: tecnico specialistico

BU: basso uso (rari)

OB: obsoleti

DI: dialettali

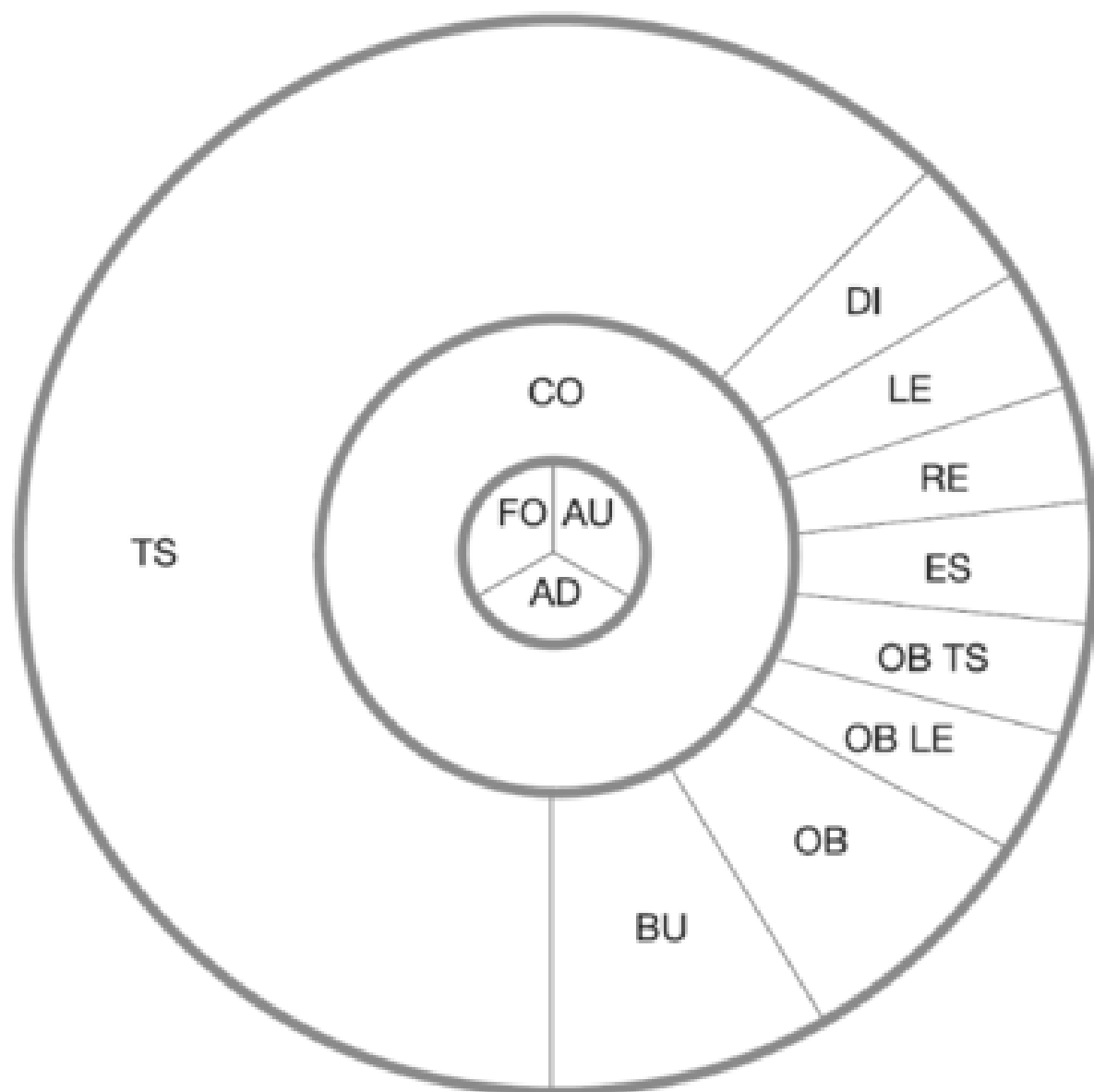
LE: letterari

RE: regionali

ES: stranieri

OB TS: obsoleti di uso
tecnico-specialistico

OB LE: obsoleti letterari

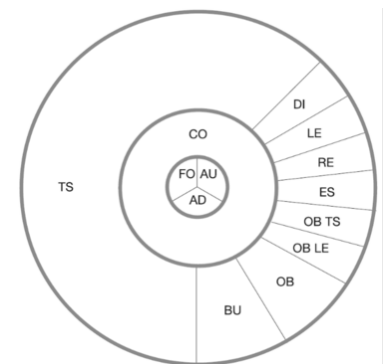


VdB

- 1. **Vocabolario fondamentale**, composto da 1.991 parole. Sono le più usate in assoluto nella nostra lingua (esempi: amore, lavoro, pane).
- 2. **Vocabolario di alto uso**, composto da 2.750 parole. Sono molto usate, ma meno di quelle del Vocabolario fondamentale (esempi: palo, seta, toro).
- 3. **Vocabolario di alta disponibilità**, composto da 2.337 parole. Sono poco usate nella lingua scritta, ma molto in quella parlata (esempi: mensa, lacca, tuta).

Il lessico di una lingua può essere paragonato a una cipolla (Figura 3.2). Il suo nucleo è costituito da un gruppo molto ristretto di parole che tutti i parlanti conoscono e che da sole servono nella stragrande maggioranza dei testi verbali, scritti e orali. Seguendo la terminologia di Tullio De Mauro e del suo gruppo di ricerca (a cui mi riferirò costantemente da qui in avanti, con la precisazione che esistono anche altri sistemi classificatori che non spostano però in nessun modo la sostanza di quanto espongo), questo nucleo prende il nome di *Vocabolario di base*, per l'italiano costituito da 6698 parole (secondo quanto emerge dal *Grande Dizionario Italiano dell'uso*, da qui in avanti *GRADIT*, su cui torneremo più avanti e che sarà il riferimento costante per i dati relativi a questa parte e per le definizioni citate). Il nucleo di base è a sua volta suddiviso in tre sottocategorie: *Lessico fondamentale (FO)*, *Lessico di alto uso (AU)* e *Lessico di alta disponibilità (AD)*. Il *Lessico fondamentale* comprende “circa 2000 vocaboli molto frequenti, che da soli costituiscono circa il 90% delle parole che ricorrono nell'insieme di tutti i testi scritti o discorsi parlati”; con *Lessico di alto uso* si indicano i circa 2500 vocaboli di alta frequenza nei *corpora* rappresentativi dell'italiano che “da soli costituiscono circa il 6% delle parole che ricorrono nell'insieme di tutti i testi scritti o discorsi parlati”; infine il *Lessico di alta disponibilità* riunisce “circa 1900 vocaboli, di uso relativamente raro nel parlare o nello scrivere, a tutti ben noti perché di grande rilevanza nella vita quotidiana (alluce, batuffolo, carrozzeria, dogana)”.

Lo strato della cipolla immediatamente sopra al nucleo è spesso ed è costituito dal *Lessico comune*, cioè i 68.782 “vocaboli che possiamo capire indipendentemente dalla professione o mestiere che esercitiamo e che sono generalmente noti a chiunque abbia un livello medio-superiore di istruzione”. Infine la cipolla lessicale è completata da un ultimo strato ancora più spesso e denso, in larga parte occupato dal *Lessico Tecnico Specialistico (TS)*, vale a dire quello dei 178.234 vocaboli “di uso solo tecnico-specialistico, noti soprattutto in rapporto a particolari attività”; alcune aree di questo strato sono invece occupate da categorie più o meno influenti dal punto di vista quantitativo: vocaboli di uso letterario (8963), vocaboli di uso regionale (7399), vocaboli dialettali (609), vocaboli stranieri (9389), vocaboli rari (33.970), vocaboli obsoleti (21.027), vocaboli obsoleti di uso tecnico-specialistico (3097), vocaboli obsoleti di uso letterario (2470).



- Il sito di Gulpease: www.corrige.it
- Gulpease online free: [sito](#)

Corpora

- I cosiddetti *corpora* (sing. *corpus*) linguistici sono collezioni, per lo più di grandi dimensioni, di testi orali o scritti prodotti in contesti comunicativi reali (per es., registrazioni di discorsi o articoli di giornale), conservati in formato elettronico e spesso corredati di strumenti di consultazione informatici.
- I *corpora* permettono di osservare l'uso effettivo di una lingua e di verificarne tendenze generali su base statistica (Lüdeling & Kytö 2008 e 2009). Rivestono un'importanza fondamentale nella lessicografia contemporanea, all'interno della quale sono utilizzati tra l'altro per selezionare lemmi in base alla loro frequenza d'uso, per identificare le costruzioni tipiche in cui una parola occorre e per coglierne le sfumature di senso in base ai contesti.

Lessico

- **1.** Dizionario, vocabolario, come opera che registra alfabeticamente le parole di una lingua dando di ciascuna la spiegazione;
- 2.** Il complesso dei vocaboli e delle locuzioni che costituiscono una lingua, o una parte di essa, o la lingua di uno scrittore, di una scuola, o di un qualsiasi parlante (in questo senso, è meno com. di *vocabolario*) [...]; *l. familiare*, insieme di vocaboli e modi espressivi, spesso molto personali anche nella formazione, che sono consueti nella comunicazione tra i membri di una famiglia o di una cerchia molto ristretta [...].

Lemma

In linguistica si dice **lemma** la citazione di una parola, ossia quella parola che per convenzione è scelta per rappresentare tutte le forme di una flessione.

Per esempio:

- in italiano il *lemma di ogni verbo* è in genere la forma al modo infinito e tempo presente
- in altre lingue (come nel latino e nel greco) il *lemma del verbo* è la prima persona dell'indicativo presente
- In italiano il *lemma del nome* è il maschile o il femminile singolare (per esempio: viso, favola)
- il lemma degli aggettivi è
 - il maschile singolare per gli aggettivi a quattro uscite (per esempio: bello, primo)
 - e il maschile o femminile singolare per gli aggettivi a due uscite (per esempio: utile, presente).
 - Sostanzialmente *il lemma è la forma rappresentativa di tutte le altre forme flesse che una classe di parole può avere*
- L'operazione che porta dalla forma flessa al corrispettivo lemma si dice *lemmatizzazione*.

- Al lessico *fondamentale*, composto di poco più di 2000 unità, appartengono quei lessemi che costituiscono circa il 90% di qualunque testo italiano. Si tratta per lo più di lessemi grammaticali, che servono cioè non a significare concetti ma piuttosto a costruire la struttura delle frasi: preposizioni, articoli, congiunzioni, avverbi, verbi ausiliari, modali e verbi supporto (*essere, avere, fare, venire, potere, dovere*). Rientrano inoltre in questo nucleo verbi, nomi e aggettivi di alta frequenza d'uso, da *abitare, anno* e *azzurro* a *volere, zia* e *zitto*.
- Il lessico *di alto uso* è costituito da quei 2500-3000 lessemi di frequenza immediatamente inferiore: voci come *abbassare, alimento* o *africano*, o come *veneziano, zampa* e *zappare*.
- I 2300 lessemi *di alta disponibilità* (da *abbraccio, accavallare, accogliente* a *zampogna, zoppicare* e *zoppo*) corrispondono a «parole che può accaderci di non dire né tanto meno di scrivere mai o quasi mai, ma legate a oggetti, fatti, esperienze ben noti a tutte le persone adulte nella vita quotidiana» (De Mauro 1980: 148). *Forchetta*, ad es., è un lessema di alta disponibilità: pensiamo alla forchetta quando l'abbiamo davanti, a tavola o in cucina, e cioè proprio in quelle occasioni nelle quali capita più spesso di indicarla a gesti che di chiamarla col suo nome. In un certo senso, i lessemi di alta disponibilità sono pensati molto più spesso di quanto non siano detti o scritti, e ciò rende problematico riscontrarli analizzando un corpus di testi in maniera puramente automatica.

- Tra le diverse componenti del vocabolario di base c'è una differenza significativa. I lessemi fondamentali e di alto uso sono voci di lunga durata, per lo più latinismi patrimoniali (vedi sopra) o neoformazioni antiche, e dunque presenti in italiano da secoli o, più spesso, fin dalle origini. Il lessico di alta disponibilità è invece più esposto alle evoluzioni che avvengono nella società e più in particolare nella cultura materiale. *Atomica, asciugacapelli, abbagliante, aspirapolvere, astronave, autoscuola, cachet, citofono, computer, democristiano, elicottero, frigorifero, nylon, parabrezza, parcheggiare, prenotare, propagandare, registratore, semaforo, shampoo, telecronaca, tergicristallo, tifo, tostapane* nominano oggetti o concetti nati in epoca moderna e sono essi stessi lessemi di formazione recente (sebbene tra età dei nomi ed età degli oggetti nominati non vi sia alcun legame necessario, come mostrano le innumerevoli realtà nate di recente che hanno ricevuto nomi di formazione antica, dalla *spina* elettrica alla *chiavetta* USB al *foglio* elettronico del computer).