Università degli Studi di Trieste

a.a. 2020/21

LINGUISTICA ITALIANA

1 – FORME E STRUTTURE DELL'ITALIANO

Lezione 4 – Lessicologia e lessicografia

Lo studio del lessico è affidato alla LESSICOLOGIA e alla LESSICOGRAFIA

La **lessicologia** è lo studio scientifico del lessico e delle parole; la **lessicografia** ha un intento più pratico, e si occupa di descrivere e di catalogare il lessico di una lingua

Quando entra in un dizionario, un lessema prende il nome di LEMMA (o ENTRATA LESSICALE) e viene descritto secondo le consuetudini lessicografiche

Per prima cosa, occorre distinguere tra il lessico di una lingua, o di una comunità di parlanti, e il lessico di un singolo individuo, sia egli/ella un autore/autrice di lingua letteraria oppure un parlante generico

Il lessico, e la semantica delle parole, costituiscono la struttura più esterna di una lingua: quella più soggetta al contatto con altre lingue, e dunque ad alterazioni (non dimentichiamo però la grande stabilità del nostro lessico di base)

Al contrario, fonetica e morfologia sono strutture profonde e soggette pertanto a cambiamenti assai limitati

Il lessico di ogni lingua si compone di una componente patrimoniale, o ereditaria, di neoformazioni endogene (derivati, composti) e di acquisizioni esogene (prestiti)

Il lessico, tuttavia, oltre a essere “fotografato” sincronicamente nel dizionario, possiede uno spessore storico, una vera e propria stratigrafia, che possiamo analizzare nei suoi sviluppi fondamentali

In particolare, fino al Cinquecento e oltre, il ruolo del latino è stato fondamentale per sostituire la lingua italiana in alcuni settori della cultura, ma anche per rifornirla delle parole di cui via via necessitava e che gradualmente sono entrate nell'uso

Queste parole sono dette CULTISMI, o latinismi: sono parole di origine latina introdotte in italiano in secoli recenti, di solito a partire dall'Umanesimo

I latinismi sono prestiti a tutti gli effetti. Il GRADIT ne censisce circa 30.000: *alleanza, amicizia, difficoltà, arteria, cervello, cemento, intimo, mirabile, essenza, raviolo, salame*. Circa la metà sono ingressi recenti, dei settori scientifici

In alcuni casi si sono creati dei doppioni (ALLÒTROPI) tra parola patrimoniale e latinismo

Dal latino VITIUM, per esempio, si è avuta la parola italiana *vezzo*; ma poi è stato recuperato il latinismo *vizio*, con altra specializzazione semantica

Altro esempio è il latino DISCUM, che ha dato l'it. *desco*; e successivamente è entrato in italiano il latinismo *disco*, che ha assunto molti significati nella storia della lingua

La duplicazione insomma non è semantica, ma etimologica: il significato delle parole è diverso, benché la loro parola di origine sia la medesima

A differenza degli altri prestiti, i latinismi sono stati per lo più facilmente adattati alla struttura fonomorfologica della lingua italiana. Con qualche eccezione: *curriculum, virus, referendum, quorum, omissis, ictus, agenda, grosso modo, auditorium, focus, (mass) media*

Un caso curioso sono i latinismi di origine popolare. Per esempio, la locuz. avv. *in visibilio* deriva dalla recitazione in latino del *Credo* (<<visibilium omnium et invisibilium>>)

Di quante parole è formata la lingua italiana? Non è possibile stabilirlo con precisione, ma possono venirci in aiuto i dizionari

Un comune dizionario contiene 100-150 mila lemmi: per arrivare a stabilire quante parole esistono, dovremo considerare che: ogni nome ha di solito un plurale diverso, ogni aggettivo quattro forme, ogni verbo qualche decina, che vanno aggiunti gli alterati ecc.

Poiché il GRADIT, col suo lemmario allargato e molto completo, assomma 260.000 lemmi (e 130 mila polirematiche), si stima che il lessico italiano ammonti a 2 milioni e mezzo di parole, escluse quelle delle lingue speciali (scientifiche)

Nessun parlante può conoscere tutte le parole; e del resto un dizionario contiene quelle di uso quotidiano, quelle rare, quelle desuete, quelle specialistiche, una parte di quelle regionali

I dizionari si servono di indicatori per segnalare al lettore il gruppo di appartenenza del lessico particolare. Questi indicatori si chiamano MARCHE D'USO (Rar., Dis., Lett., Med., Reg.), a cui si aggiungono le MARCHE DIASISTEMATICHE che invece identificano un ambito o registro: Pop., Bur., Com., Fam., Iron., Parl., Poet., Volg.

Il nucleo del lessico è costituito dal VOCABOLARIO DI BASE, composto di circa 6700 lessemi

Esso si divide in LESSICO FONDAMENTALE (2000 lemmi ca.), VOCABOLARIO DI ALTO USO (2700), VOCABOLARIO DI ALTA DISPONIBILITÀ (2000)

Queste articolazioni del lessico sono possibili grazie alla compilazione e all'uso di lessici di frequenza. Il più recente e importante è il LIP (Lessico di frequenza dell'Italiano Parlato)

Tra i primi 20 lemmi registrati compaiono molte parole grammaticali: *il, uno, di, a, in, da, e, che*, e poi *essere, avere, questo, lui, dire, fare* ecc. Dopo le parole fondamentali vengono quelle di "alta frequenza", o "alto uso", come *fine, fiocco, fischiare*, ecc.

Il gruppo successivo è invece costituito da parole di "alta disponibilità", cioè da parole che, pur non comparendo entro le prime 5000 posizioni per frequenza, sono di significato noto e utilizzabili secondo necessità: *batuffolo, carrozzeria, forchetta*. Si tratta dunque di un'aggiunta dovuta a considerazioni di tipo lessicografico

La successiva sottodivisione lessicale comprende il LESSICO COMUNE (40-45 mila parole circa), che aggiunto a quello di base forma il vocabolario corrente, accessibile alla maggior parte delle persone che possieda un livello di istruzione medio superiore (ma con molte riserve)

Ben 55000 parole del GRADIT sono etichettate come desuete, letterarie o rare. Sono conservate per consentire la lettura dei testi letterari; ma possono anche rientrare nell'uso

Per quanto riguarda invece le percentuali di presenza delle categorie grammaticali, gli invariabili costituiscono circa il 3%. Sarà importante notare che, a fronte di un 9% di verbi e a un 26% di aggettivi, per il 62% il lessico italiano è formato da nomi

La lessicologia si occupa ad esempio di definire i rapporti di significato tra le parole. Essi possono essere paradigmatici (associativi) o sintagmatici (combinatori)

Se sono davanti a un paesaggio, posso decidere se lo sto *guardando, osservando, contemplando, ammirando*, ecc. Se una cosa è *vicina* o *lontana*, *bella, bellissima, orribile*, sto operando sui rapporti di paradigma lessicale

Se ho acceso l'automobile e devo partire, devo premere il pedale dell'*acceleratore* (e non del *freno*, o della *frizione*). Se dunque il contesto mi suggerisce quale parola riempie un vuoto lessicale, sto operando sui rapporti sintagmatici

Se finora abbiamo conteggiato numericamente le parole diverse di un repertorio lessicale, non si deve dimenticare che abitualmente le parole possiedono più accezioni. Questa particolarità è detta POLISEMIA

Che cosa significa *acqua*? Il liquido incolore, inodore, insapore ecc., ma anche la pioggia, la trasparenza delle pietre preziose (diamante), il liquido della ferita, un gruppo di segni zodiacali; e inoltre ci sono le acque del mare, della placenta, quelle termali, l'acqua benedetta, l'acqua ossigenata, l'acqua di colonia, e la classe non è acqua

Reciprocamente, le parole possono essere in rapporto orizzontale, di <u>sinonimia</u> e <u>antonimia</u> (i "contrari"), oppure di omonimia (parole casualmente uguali: *sale* 'cloruro di sodio' e *sale* (voce del v. *salire*)

Oppure, ci sono rapporti verticali tra iperonimi e iponimi. I primi sono parole di significato più generale, i secondi sono più particolari. Un esempio di sequenza è *animale – mammifero – gatto – soriano*

Esistono anche i meronimi, cioè parole incluse in altre: *volante, cruscotto, sedile*, ecc. sono meronimi di *automobile*

Le parole possono combinarsi occasionalmente, oppure possono formare combinazioni stabili, di forza inferiore alle locuzioni (e infatti possono essere separate da elementi interposti). Queste sono chiamate COLLOCAZIONI. Per esempio: *caffè macchiato* (N+A), *stendere un documento, prendere una decisione* (V+N), *pagare profumatamente* (V+Avv)

Il lessico comune è anche impiegato per tarare gli indici di leggibilità. Essi sono dei valori numerici che stabiliscono la difficoltà presunta di un testo per il suo lettore.

Il più famoso e usato è l'indice Gulpease (GULP sta per Gruppo Universitario Linguistico-Pedagogico: un gruppo di ricercatori dell'Università "La Sapienza" di Roma che elaborò questa ricerca negli anni '80).

Attraverso un software, disponibile anche on line, misura il rapporto tra numero di frasi, numero di lettere e numero di parole, considerando che frasi e parole lunghe siano di più difficile comprensione. Il software calcola questi valori in un testo e dà un risultato numerico che stabilisce se un testo è comprensibile da chi ha almeno la licenza media, il diploma, la laurea ecc.

Il servizio è disponibile (a pagamento) sul sito www.corrige.it

LESSICOGRAFIA

Differenza tra VOCABOLARIO e DIZIONARIO

VOCABOLARIO è un insieme lessicale finito, riferibile anche a singoli parlanti, a opere, ad autori letterari ecc.

DIZIONARIO è invece il termine tecnico dello strumento lessicografico di consultazione, o di lettura, che ci consente di conoscere meglio il significato del lessico

Esistono dizionari monolingui, ma anche bilingui, oppure plurilingui. Ci sono dizionari tecnici e specialistici, dizionari enciclopedici, dizionari etimologici, dizionari dei sinonimi e dei contrari, dizionari inversi, dizionari delle collocazioni, dizionari storici, ecc.

Esistono anche dizionari cartacei e dizionari on line

Storicamente, uno dei primi dizionari è il *Glossario di Monza*, un repertorio del X secolo che conta 60 lemmi tradotti dall'italiano (settentrionale) al greco-bizantino. Per es.: *de capo=cefali; costa=pleura*, ecc.

Luigi Pulci compilò un *Vocabulista*, 700 parole tratte dal suo *Morgante* (1478); Niccolò Liburnio invece pubblicò *Le tre fontane* (1526), repertorio lessicale delle Tre Corone toscane. Il *Dittionario* di Francesco Sansovino (1568) è un primo esempio di lessicografia attenta agli usi locali

Ma esistono anche repertori plurilingui di enorme importanza storica: la Stele di Rosetta (scritta in geroglifico, demotico e greco); il Calepino (antenato dei dizionari pluringui tascabili), il *Vocabolario degli Accademici della Crusca* (I ed. 1612)

Nel dizionario di uso corrente troviamo lemma, categoria grammaticale, eventualmente il genere, la scansione sillabica o la pronuncia, le accezioni, le marche d'uso

Le marche del GRADIT ci consentono di verificare immediatamente l'appartenenza di un termine a uno dei settori lessicali: FO – AU – AD – CO – BU (con possibile specificazione RE – DI – LE – ES – TS)

Nei dizionari dell'uso le parole fondamentali, di alto uso e alta disponibilità sono evidenziate con un espediente: Sabatini-Coletti usa un fondino rosso, Zingarelli un rombo pieno, Devoto-Oli l'inchiostro azzurro

Esistono dizionari storici, che raccolgono l'evoluzione semantica delle parole attraverso i secoli, spesso con esempi d'autore: il *Vocabolario degli Accademici della Crusca*, dal 1612, ha questa caratteristica; così il Dizionario di Niccolò Tommaseo, ottocentesco, fino al *Grande Dizionario della Lingua Italiana* (1961-2002) – GDLI

Esistono dizionari etimologici, nei quali si ritrova l'origine remota della parola, solitamente posta al vertice di tutta la famiglia che si è creata attorno a essa

Esistono dizionari speciali, come i dizionari analogici, i dizionari ragionati, ecc., nei quali il lessico viene illustrato nelle sue ramificazioni: non in modo orizzontale come spesso siamo abituati a pensarlo, ma ragionando su basi e loro derivazioni e composizioni