

Fisionomia lessicale del romeno letterario contemporaneo

Autor(en): **Maneca, Constant**

Objekttyp: **Article**

Zeitschrift: **Revue de linguistique romane**

Band (Jahr): **31 (1967)**

Heft 121-122

PDF erstellt am: **27.04.2024**

Persistenter Link: <https://doi.org/10.5169/seals-399406>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

FISIONOMIA LESSICALE DEL ROMENO LETTERARIO CONTEMPORANEO

La struttura etimologica del lessico ha preoccupato già da molto gli studiosi della lingua romena, che hanno affrontato questo problema da diversi punti di vista ¹.

Quello che ha intuito il valore circolatorio delle parole (enunciando un'idea che rappresenta un punto di vista tutto originale) è stato B. P. Hasdeu, il quale ha mostrato l'importanza della frequenza nel definire la fisionomia lessicale di una lingua: *În lingvistică marele principiu al circulațiunii, uitat pînă aci aproape cu desăvîrșire, s-ar putea privi ca piatra cea angulară a edificiului. Ceea ce se cheamă fizionomia unei limbi nu este altceva decît rezultatul circulațiunii*².

I metodi della statistica moderna ci offrono la possibilità a determinare con precisione la struttura etimologica della lingua. Sfortunatamente, fino agli ultimi tempi, sono poche le ricerche di questo tipo, date le difficoltà del metodo (calcoli laboriosi, numerazioni esaurienti, etc.)³.

Per poter determinare più rigorosamente la struttura e le tendenze evolutive del lessico romeno moderno ci è sembrato necessario un nuovo studio sopra l'origine delle parole nella luce della loro frequenza. A tale

1. Il primo fu A. de Cihac col suo *Dictionnaire d'étymologie daco-romane*, Francfort s/M, vol. II, 1879, p. VIII.

2. *Cuvente den Bătrîni*, tom. III, *Istoria limbei române* de B. Petriceicu Hasdeu, Par-tea I; *Principii de lingvistică*, București, 1881, p. 104.

3. Uno studio che riguarda appositamente la composizione etimologica del lessico romeno è quello di D. Macrea, *Fizionomia lexicală a limbii române* in DR, X/II, 1943, p. 362-373, limitato all'opera di M. Eminescu. Altri studi di questo genere nei quali ci sono anche dei dati riguardando l'oggetto della nostra ricerca sono: V. Suteu, *Observații asupra frecvenței cuvintelor în operele unor scriitori români*, in ȘCL, X 1959, p. 194, Gh. Bolocan, *Unele caracteristici ale stilului publicistic al limbii române literare*, in SCL, XII, 1961, p. 35. A questa occasione porto un omaggio al mio padre, Michele Maneca, che mi ha dato un prezioso soccorso allo spoglio del materiale per la mia statistica.

scopo abbiamo utilizzato i dati offertici da una nostra ricerca statistica sopra la frequenza dei vocaboli nella lingua romena contemporanea¹.

Siccome i testi spogliati da noi costituiscono una prova che rappresenta i principali stili della lingua² estenderemmo sulla lingua romena letteraria in generale alcune conclusioni riguardando il loro vocabolario.

Osserviamo — come prima constatazione della nostra ricerca — che in questi testi abbiamo incontrato un numero di 6231 lessemi³ che totalizzano 49.161 occorrenze con una frequenza media di 7,88, sensibilmente differente da quelle stabilite negli studi precedenti⁴. Il numero medio di lessemi per cento occorrenze è di 12,67, il che mostra ancora una differenza

1. La nostra ricerca è stata fatta su testi lunghi di 50.000 occorrenze (a questo numero l'errore tipo è praticamente trascurabile), colte da cinquanta lavori. Le opere studiate sono apparse negli ultimi venti anni o appartengono a degli autori che gli hanno scritte nel periodo tra le due guerre mondiali, così che esse riflettano lo stato del vocabolario della lingua romena attuale. La composizione del materiale spogliato è la seguente: 5.200 occorrenze sono state estratte dalla poesia (versi di Tudor Arghezi, Ion Minulescu, Gh. Topîrceanu, Mihai Beniuc, Camil Petrescu); 10.400 dalla prosa artistica (racconti, teatro e romanzi di Ion Agârbiceanu, George Călinescu, Liviu Rebreanu, Eusebiu Camilar, Cezar Petrescu, Zaharia Stancu, traduzioni dalla letteratura universale di Mihai Codreanu, Francisc Munteanu); 10.400 occorrenze dalla stampa attuale (giornali, riviste senza carattere di specialità); 19.000 occorrenze sono state estratte dai lavori di specialità dei diversi campi della scienza (fisica, chimica, biologia, geografia, linguistica, legge, storia, critica e storia letteraria) e 5.000 dai testi amministrativi. Quando abbiamo incominciato la nostra ricerca non era ancora apparso il *Frequency Dictionary of Rumanian Words* by Alphonse Juilland, P. M. H. Edwards, Ileana Juilland, Mouton & Co., London — The Hague — Paris, 1965, che d'altronde, non ci avrebbe offerto tutti i dati che ci occorrevano (in quanto non presenta che l'uso, la frequenza e la dispersione delle parole del « vocabolario elementare » della lingua romena, stabilito secondo certi principi selettivi dagli autori del *Dizionario*). Un paragone tra i dati della nostra statistica e quelli del menzionato Dizionario di frequenza sarebbe interessante e mostrerebbe numerevoli concordanze.

2. Consideriamo che nella lingua letteraria si distinguono solamente tre stili (che, a loro volta, presentano diverse varietà): lo stile artistico-letterario, lo stile scientifico-amministrativo e quello colloquiale (cfr. il nostro articolo *Considerațiuni asupra stilurilor limbii române literare în lumina frecvenței cuvintelor*, in LR, XV, 1966, p. 427). L'ultimo è rappresentato nella nostra prova di lingua solamente dai testi di letteratura drammatica.

3. Al quale si aggiungono 244 nomi propri e 66 segni diversi. Generalmente, abbiamo incontrato nelle 50.000 occorrenze 6.475 unità distinte (invarianti).

4. Cfr. V. Șuteu, *art. cit.*, p. 421 dove è indicata la frequenza media di 10,99 e Gh. Bolocan, *art. cit.*, p. 42 dove viene indicata la frequenza media di 12,25.

considerevole tra la frequenza delle parole in un certo stile e quella dai testi che appartengono a vari stili¹. Molto diverso, in confronto alle ricerche anteriori si presenta anche l'indice di ricchezza del lessico. Calcolato secondo la formula proposta da Pierre Guiraud² $a = \frac{V}{\sqrt{N}}$, questo indice è 27,86 (superiore agli indici risultati dalle statistiche menzionate, molto meno differenti tra loro), il che si spiega dato che la nostra prova di lingua ha contenuto anche dei testi scientifici.

Vediamo ora i dati riguardando la composizione etimologica del vocabolario dei testi studiati, e poi le conclusioni che potremmo fare sul lessico della lingua romena letteraria contemporanea in generale.

Dalle 6231 parole, un numero di 2860 (cioè il 45,80 per cento) appartengono al fondo più antico della lingua mentre 3371 (il 54,11 per cento) sono neologismi³.

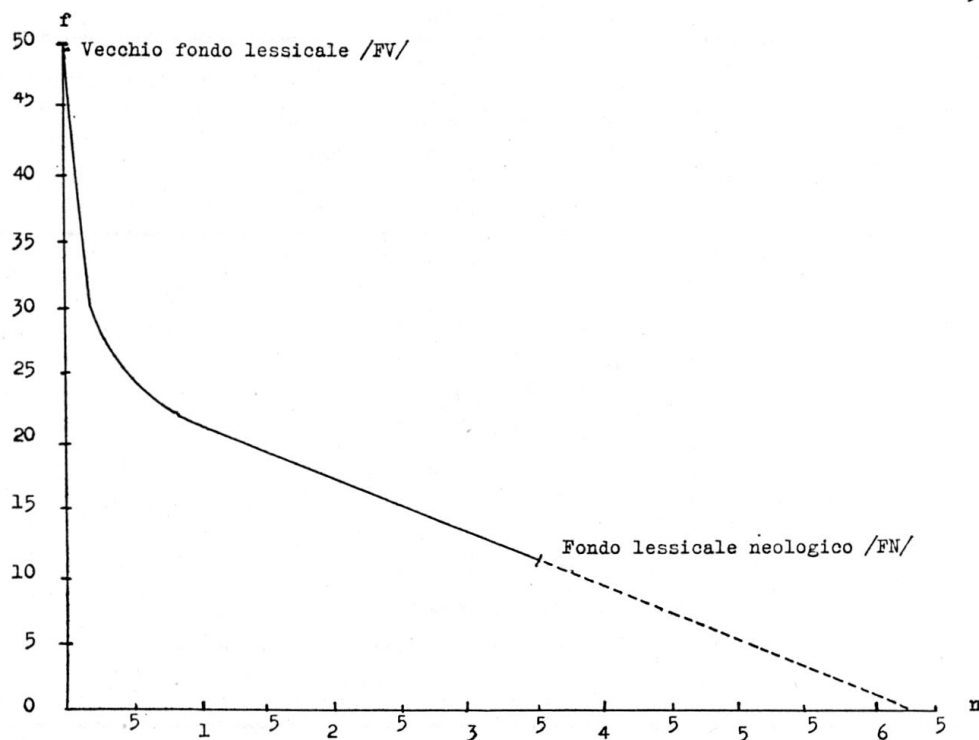
Se sotto l'aspetto numerico il rapporto tra le parole vecchie ed i neologismi $\left(n = \frac{FV}{FN} = \frac{2860}{3371} \text{ oppure, percentualmente, } \frac{45,80}{54,11}\right)$ tende ad assumere un valore vicino all'unità (0,84), dal punto di vista della frequenza la situazione è completamente diversa: in questo caso il rapporto $\left(f = \frac{FV}{FN} = \frac{37198}{11963} \text{ oppure, percentualmente, } \frac{74,39}{25,61}\right)$ presenta il valore 3,10, indice che mostra la grande superiorità del vecchio fondo di parole.

Rapportato alle coordinate frequenza (sull'ordinata, notata in migliaia di parole) e numero (sull'ascissa, notato, ugualmente in migliaia di parole) lo stato delle parole vecchie di fronte ai neologismi si può rappresentare così:

1. Cfr. V. Şuteu, *id.*, p. 420 dove mostra che nei suoi testi il numero medio per cento parole era di 9,09. Cfr. anche Gh. Bolocan, *art. cit.*, p. 42, dove il numero medio per cento parole, calcolato da lui, è stato: 8,15.

2. Pierre Guiraud, *Problèmes et méthodes de la statistique linguistique*, Dordrecht Holland, 1959, p. 89; in questa formula V = il numero dei lessemi e N = il numero delle occorrenze. Gli stessi indici che riescono dai testi ricercati da Gh. Bolocan e V. Şuteu sono, rispettivamente, 20,85 e 20,33.

3. Consideriamo come neologismi i prestiti (savanti) e le formazioni romene (formate sempre dai prestiti stranieri) apparse nella lingua nelle ultime due decadi dell'Ottocento.



Le parole del vocabolario dei testi studiati si dividono, secondo la loro origine ¹, in ventidue gruppi etimologici : latino, romeno, ungherese, vecchio slavo, slavo ecclesiastico, bulgaro, serbocroato, ucraino, autoctono (albanese), onomatopeico, francese, italiano, latino dotto, greco, medio-greco, neogreco, turco, russo, inglese, tedesco dialettale e tedesco letterario, ai quali si aggiunge il gruppo delle parole con etimologia incerta o sconosciuta.

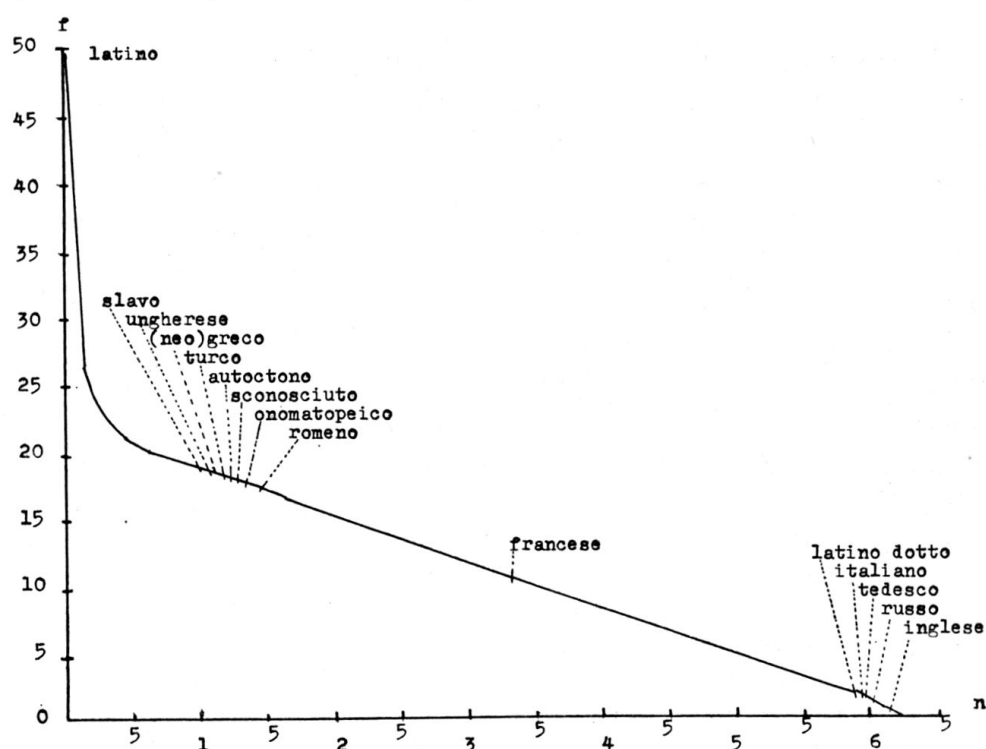
Per poter presentare dati centralizzati più concludenti, abbiamo schierato insieme le parole vecchie d'origine slava (vecchio slavo, slavo ecclesiastico, bulgaro, serbocroato, ucraino), quelle greche (venute in diverse epoche dal antico greco, mediogreco e neogreco) e quelle tedesche (dotte e dialettali), riducendo in questa maniera a quindici il numero dei gruppi.

La composizione etimologica del vocabolario dei testi studiati è la seguente :

1. Abbiamo indicato l'origine dei vocaboli secondo il *Dicționarul limbii române moderne* al Academiei R. S. România (București, 1958) nonostante le nostre obiezioni (che risalgono in parte anche dalla nostra recensione apparsa in RFRG, 1959, p. 264 sq.) in ciò che riguarda la parte etimologica di questo lavoro (specialmente in ciò che riguarda i neologismi per i quali, in molti casi, abbiamo proposto altre etimologie nel *Dicționar de neologisme*, București, 1961.

ORIGINE DELLE PAROLE ¹	NUMERO DEI LESSEMI	PERCEN- TUALE	FRE- QUENZA	PERCEN- TUALE	RAPPORTO TRA LA FREQUENZA ED IL NUMERO	PRODOTTO DELLA FREQUENZA E DEL NUMERO
latina.....	972	15,11	30.455	60,92	31,33	29.602.260
slava.....	436	6,99	1.495	3,04	3,42	651.820
ungherese.....	42	0,69	227	0,45	5,40	9.534
** (neo) greca.....	79	1,26	196	0,42	2,48	15.484
turca.....	56	0,89	112	0,23	2	6.272
autoctona.....	29	0,46	102	0,21	3,51	2.958
sconosciuta.....	67	1,07	191	0,41	2,85	12.797
onomatopeica..	17	0,27	25	0,05	1,47	425
**romena.....	1.613	25,90	5.658	11,52	3,50	9.126.354
* francese.....	2.594	41,63	8.946	18,20	3,44	23.205.924
* latina dotta.....	200	3,20	1.290	2,62	6,45	258.000
**italiana.....	77	1,26	249	0,46	3,23	19.173
* tedesca.....	59	0,94	127	0,25	2,15	7.493
* russa.....	13	0,20	72	0,14	5,53	936
* inglese.....	7	0,11	16	0,03	2,28	112

Somma dei prodotti.....62.919.542



1. I gruppi notati con* sono formati dai neologismi, quelli notati con** sono formati dai neologismi, e dalle parole vecchie.

Rapportata alle due coordinate, frequenza, (notata sull'ordinata in migliaia di parole) e numero (notato sull'ascissa, sempre in migliaia di parole) si può rappresentare così lo stato dei gruppi etimologici della tabella in su presentata (p. 194).

La linea che collega i punti che rappresentano la posizione dei diversi gruppi etimologici scende d'un tratto sull'ordinata da 50 a 20 e da qui incomincia una doppia variazione lenta, fino alla zero. Se consideriamo i dati ottenuti, possiamo affermare che la posizione delle parole latine è predominante.

In quanto alle relazioni che si possono stabilire tra l'origine delle parole ed il rango della loro frequenza, la nostra statistica ha dimostrato che tra le prime 102 parole, 84 sono latine ereditate, 8 sono formate su terreno romeno da elementi latini, 3 sono d'origine slava, una è imprestata dal ungherese, 3 provengono dal francese ed una dal latino dotto.

Le parole da noi studiate possono essere divise dal punto di vista della frequenza, in due categorie: quelle di cui la frequenza assoluta è per lo meno uguale al numero intero della frequenza media e quelle che presentano una frequenza assoluta inferiore a questo numero.

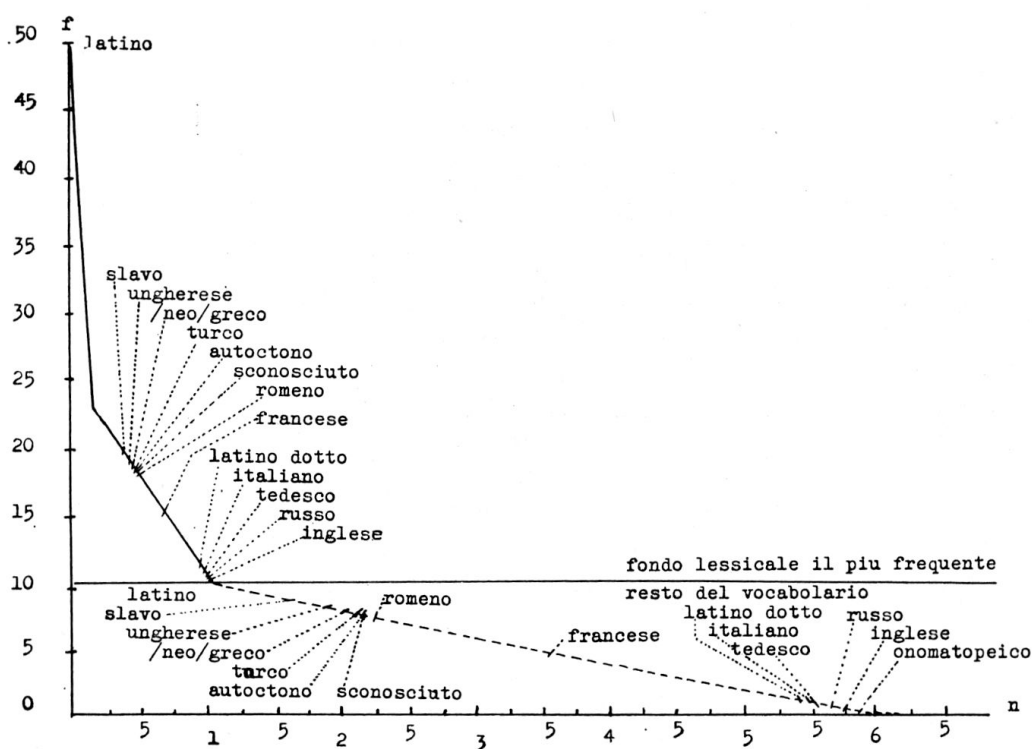
ORIGINE DELLE PAROLE	NUMERO		PERCENTUALE		FREQUENZA		PERCENTUALE	
	Fondo più freq.	Resto del voc.	Fondo più freq.	Resto del voc.	Fondo più freq.	Resto del voc.	Fondo più freq.	Resto del voc.
latina	349	593	34,65	11,35	28.827	1.628	73,75	16,27
slava	54	382	5,35	7,31	819	676	2,09	6,75
ungherese . . .	8	34	0,79	0,66	163	64	0,41	0,64
(neo) greca . .	6	73	0,59	1,43	75	119	0,18	1,18
turca	2	54	0,19	1,03	19	93	0,04	0,91
autoctona . . .	5	24	0,49	0,46	60	42	0,15	0,41
sconosciuta . .	3	64	0,29	1,22	87	104	0,22	1,03
**romena	193	1.420	19,16	27,18	3.093	2.565	7,91	25,63
* francese	324	2.270	32,17	43,44	4.769	4.177	12,20	41,75
* latina dotto . .	48	152	4,76	2,94	975	315	2,49	3,14
* italiana	9	68	0,89	1,3	107	142	0,27	1,41
* tedesca	3	56	0,29	1,02	30	97	0,07	0,96
* russa	2	11	0,19	0,2	50	22	0,12	0,21
* inglese	1	6	0,09	0,11	9	7	0,02	0,69
onomatopeica		17		0,27		25		0,05

Abbiamo chiamato la prima categoria « il lessico più frequente »¹;

1. Ci siamo decisi per questo termine che ci sembra il più adatto in quanto riferisce al criterio di determinazione.

l'analisi dei dati riguardando l'etimologia dei vocaboli appartenendo a questa categoria ¹ è molto interessante in quanto essa riflette l'origine delle parole che hanno la massima circolazione nella lingua. La composizione etimologica del lessico più frequente paragonata a quella del resto del vocabolario dei testi è la seguente (p. 195).

Ecco il grafico di questa situazione notata nello stesso sistema (la linea continua rappresenta il « fondo più frequente » e quella puntata interrotta, il resto del vocabolario).



La struttura etimologica di queste due categorie è molto diversa. Mentre nel « fondo più frequente » l'elemento latino è il più importante, tanto dal punto di vista del numero, quanto — e soprattutto — della frequenza, nel resto del vocabolario le parole d'origine latina sono solamente al terzo posto.

Le parole d'origine non latina si situano, dal punto di vista della frequenza, alla periferia del vocabolario, dove sono i termini speciali, che hanno una circolazione più limitata. Tra questi, un posto predominante occupano i

1. Di questi dati ci siamo occupati in un rendiconto presentato alla « Società romena di linguistica romanza », nel febbraio di 1966.

neologismi dei quali la maggior parte proviene dal francese, fonte principale della terminologia scientifica e tecnica romena moderna ¹.

Nonostante la loro frequenza ridotta, i neologismi hanno una importante parte nel lessico perchè, a differenza delle parole di grande circolazione, essi portano nel processo della comunicazione una grande quantità d'informazione ².

Paragoniamo ora i risultati della nostra ricerca a quelli ottenuti prima di noi dal prof. D. Macrea, nel lavoro menzionato ³.

ORIGINE DELLE PAROLE	STATISTICA DEL PROF. MACREA ⁴				STATISTICA DEL MANECA			
	numero	percentuale	frequenza	percentuale	numero	percentuale	frequenza	percentuale
latina.....	1.756	48,68	28.095	83	972	15,11	30.455	60,92
slava.....	609	16,81	2.318	6,93	436	6,99	1.495	3,04
ungherese.....	59	1,63	285	0,84	42	0,69	227	0,45
(neo) greca.....	78	1,33	196	0,29	79	1,26	196	0,42
turca.....	44	1,20	95	0,28	56	0,89	112	0,23
autoctona.....	32	0,88	188	0,55	29	0,46	102	0,21
sconosciuta.....	125	3,46	615	1,80	67	1,07	191	0,41
*romena.....	—	—	—	—	1.613	25,90	5.658	11,52
*francese.....	432	11,97	855	2,52	2.594	41,63	8.946	18,20
*lat. dotta.....	123	3,41	484	1,13	200	3,20	1.290	2,62
*italiana.....	21	0,52	55	0,16	77	1,26	249	0,46
*tedesca.....	18	0,49	32	0,09	59	0,94	127	0,25
*russa.....	5	—	7	—	13	0,20	72	0,14
*inglese.....	—	—	—	—	7	0,11	16	0,03

1. Cfr. D. Macrea, *Terminologia științifică și tehnică în limba română contemporană*, in CL, XI, nr. 1. 1966, p. 23.

2. Si sa che l'informazione recata da una parola è inversamente proporzionale alla sua frequenza : quanto una parola è più rara, tanto essa porta una quantità maggiore d'informazione (cfr. P. Guiraud, *op. cit.*, p. 70 sq.).

3. Si veda la nota 3. Non paragoniamo i nostri dati a quelli degli studi già menzionati del Suteu e del Bolocan, perchè loro non hanno studiato che il « vocabolario fondamentale » dei testi ricercati, corrispondente con quello che noi chiamiamo « il fondo più frequente ».

4. I dati dello studio del prof. D. Macrea debbono essere considerati con alcune correzioni : agli elementi latini, slavi, ungheresi, etc. l'autore ha aggiunto anche le formazioni romene, che sono formate dagli elementi originari dalle rispettive lingue ; nel suo studio l'etimologia delle parole è stata indicata secondo I. A. Candrea, Gh. Adamescu, *Dicționarul enciclopedic ilustrat « Cartea românească »*, București, 1931 ; la lunghezza totale dei testi studiati (le poesie di M. Eminescu) è stata di 33.846 occorrenze (si veda *art. cit.*, p. 364).

I dati delle due statistiche non concordano tra di loro che per certi gruppi etimologici. Le differenze, talvolta abbastanza grandi, si spiegano tanto per la differenza di tempo tra le due « prove di lingua » studiate ¹, quanto per il loro carattere diverso, dal punto di vista della varietà ².

Si vede dunque, dopo aver fatto le correzioni ³ necessarie, che le similitudini si riferiscono specialmente al vecchio fondo lessicale della lingua e soprattutto all'elemento latino.

Anche le differenze ci sembrano molto interessanti in quanto esse mostrano la linea dello sviluppo del lessico e le differenze tra i vocabolari dei diversi stili della lingua. Osserviamo dunque che il lessico poetico è meno aperto agli influssi stranieri, a differenza di quello scientifico, porta largamente aperta ai neologismi.

Consideriamo di poter estendere alla lingua romena attuale le conclusioni riguardando la struttura etimologica del vocabolario dei testi da noi studiati. Le modificazioni che potrebbero sorgere, dato che lo stile colloquiale e la « lingua parlata » ⁴ vengono rappresentate scarsamente nella nostra prova di lingua, non cambierebbero che i dati dei gruppi etimologici secondari che non sono determinanti per la fisionomia della lingua.

Eccettando gli elementi con una situazione speciale nel lessico ⁵ (elementi che, pur non essendo privi di senso, non presentano un contenuto semantico ben definito) lo studio della frequenza delle parole ci può svelare anche altri aspetti della struttura etimologica della lingua. I valori della frequenza ed il numero dei lessemi cambiano, e la fisionomia della lingua apparisce sotto un'altra luce.

Le parole ausiliari totalizzano 17.097 occorrenze (il che rappresenta

1. Le poesie di M. Eminescu, studiate dal prof. D. Macrea, appartengono al Novecento, mentre i nostri testi sono di quasi mezzo secolo più tardi, fatto che conta, specialmente per alcuni settori periferici del lessico.

2. I testi del prof. D. Macrea rappresentano solamente una variante dello stile artistico (l'opera di un unico poeta) mentre i nostri testi rappresentano tutti gli stili della lingua.

3. Si veda la nota 4 p. 197.

4. Lo stile colloquiale e « la lingua parlata », per il loro scarso vocabolario, composto dalle parole usuali, s'avvicinano di più al « lessico più frequente ».

5. I così detti « strumenti grammaticali » o « vocaboli ausiliari » tra quali noi includiamo solamente l'articolo, le preposizioni, le congiunzioni ed i verbi ausiliari. I numerali, le interiezioni ed il pronome, che vengono schierati da altri autori nello stesso gruppo, presentano un'altra situazione semantica, per lo meno nella lingua romena.

36,65 per cento dalla frequenza totale), benchè sono solamente 49 (cioè 0,72 per cento dal vocabolario dei testi). Se omettiano le parole ausiliari, nei nostri testi rimangono 6182 lessemi che appaiono in 32.064 occorrenze.

Ecco come si presenta, da questo nuovo punto di vista, la composizione etimologica dei testi da noi studiati:

ORIGINE DELLE PAROLE	NUMERO DEI LESSEMI	PERCEN- TUALE	FRE- QUENZA	PERCEN- TUALE	RAPPORTO TRA LA FREQUEN- ZA ED IL NUMERO	PRODOTTO DELLA FREQUENZA E DEL NUMERO
latina.....	935	15,10	13.895	43,33	14,86	12.991.825
slava.....	436	7,05	1.495	4,66	3,42	651.820
ungherese.....	42	0,70	227	0,7	5,40	9.534
** (neo) greca.....	79	1,31	196	0,6	2,48	15.484
turca.....	56	0,9	112	0,34	2	6.272
autoctona.....	29	0,48	102	0,31	3,51	2.958
sconosciuta.....	66	1,06	119	0,32	1,80	7.854
onomatopeica.....	17	0,29	25	0,07	1,47	425
** romena.....	1.603	25,93	5.193	16,19	3,23	8.324.379
* francese.....	2.593	41,96	8.944	27,89	3,44	23.191.792
* latina dotta.....	200	3,21	1.290	4,02	6,45	258.000
* italiana.....	77	1,27	249	0,77	3,23	19.173
** tedesca.....	59	0,96	127	0,33	2,15	7.493
* russa.....	13	0,21	72	0,22	5,53	936
* inglese.....	7	0,11	16	0,04	2,28	112
Somma dei prodotti.....45.488.057						

Rapportato alle due coordinate — frequenza e numero ¹ — lo stato dei gruppi etimologici della tabella precedente può essere rappresentato in questo modo (p. 200).

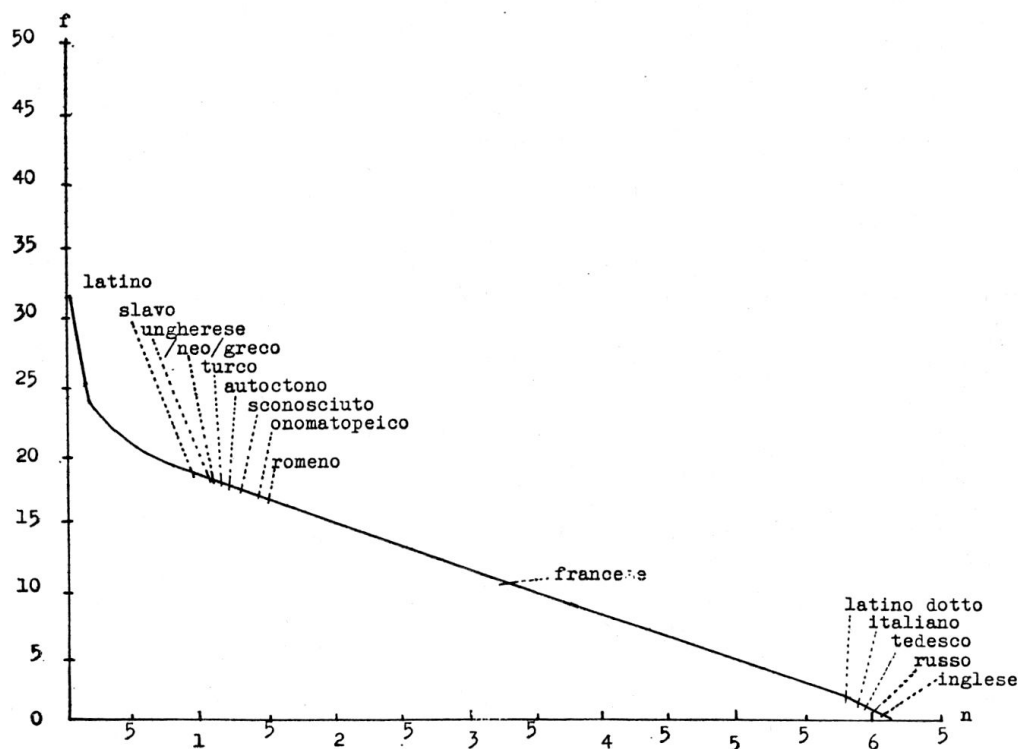
Come si vede, la fisionomia lessicale della lingua romena è ora tutt'altra. Benchè i dati assoluti sono cambiati solamente per alcuni gruppi ², la situazione generale è completamente diversa.

Le parole d'origine latina, nonostante il loro scarso numero, continuano ad avere una posizione predominante dal punto di vista della frequenza. Sull'ordinata, la differenza tra il gruppo delle parole ereditate dal latino e quello dei neologismi d'origine francese (che sull'ascissa sorpassano di più il gruppo latino) è molto più piccola che nel caso in cui abbiamo considerate tutte le parole incontrate nei testi. Lo stato dei gruppi

1. Nella stessa notazione.

2. Cambiamento essenziale solamente per le parole d'origine latina e, molto meno importante, per le parole d'origine romena.

etimologici che non oltrepassano il dieci per cento tanto come frequenza, quanto come numero, resta quasi lo stesso.



Il rapporto tra il numero delle parole vecchie e quello dei neologismi ($n = \frac{FV}{FN} = \frac{2812}{3370}$ o, percentualmente, $\frac{45,48}{54,51}$) — secondo i nuovi dati — resta quasi invariabile ($n = 0,83$ in confronto a $0,84^1$), ma il valore del rapporto tra la frequenza totale del vecchio fondo lessicale e quella dei neologismi ($f = \frac{FV}{FN} = \frac{20.101}{11.963}$ oppure, percentualmente, $\frac{62,69}{37,30}$) scade quasi alla metà ($1,68$ in confronto a $3,10$)².

Le nostre osservazioni fanno risalire l'importante contributo dei prestiti recenti (quasi esclusivamente d'origine latino-romanza) allo sviluppo dei mezzi espressivi della lingua romena, alla formazione di un ricco vocabolario e di una vasta terminologia scientifica e tecnica, capaci a denominare l'innunerevoli nozioni della cultura moderna.

Presenta molto interesse anche i valori dei rapporti tra la frequenza

1. Cfr. p. 192.

2. Questo fatto dimostra l'influsso della struttura grammaticale della lingua (tramite i suoi elementi di relazione) sulla fisionomia lessicale.

totale ed il numero delle unità lessicali sui gruppi etimologici ¹. Gli elementi latini ereditati presentano un'indice molto superiore ² a quello più alto degli altri gruppi, ciò che dimostra che queste parole hanno la massima circolazione nella lingua e determinano la sua fisionomia lessicale.

Fermiamoci un po' sulla distribuzione etimologica sui stili. La variazione della frequenza delle parole ci ha mostrato quanto interessante sarebbe una ricerca sulla concorrenza tra gli elementi latini originari e quelli penetrati più tardi nella lingua (slavi, ungheresi ecc.) oppure sulla concorrenza tra i neologismi ed i loro sinonimi più vecchi. Se osserviamo i dati di *vreme* ³ (< sl.) e *timp* ⁴ (< lat.), oppure *veac* ⁵ (< sl.) e *secol* ⁶ (< lat. dott.), vediamo che, mentre *vreme* ha la frequenza massima nello stile artistico, *timp* apparisce il più spesso nei testi scientifici, dove il neologismo *secol* tende a prendere il posto del più vecchio *veac*, che trova rifugio nei testi più conservatori della letteratura artistica.

Un'altra osservazione, collegata alla dispersione delle parole sui stili è quella che i testi scientifici costituiscono l'ingresso dei neologismi nella lingua ⁷. Il vocabolario della stampa rappresenta — dal punto di vista lessicale, l'aspetto intermedio tra la lingua della prosa artistica e lo stile scientifico.

Abbiamo pensato che, per poter stabilire il valore circolatorio delle parole che determinano la fisionomia lessicale di una lingua, un elemento importante è il prodotto ⁸ tra il numero e la frequenza dei lessemi di una certa origine.

In base di questo prodotto, secondo la formula ⁹: $C = \frac{f \cdot n}{P/100}$, ab-

1. Cfr. le rispettive rubriche dai tabelli delle p. 195 e p. 199. Questi valori rappresentano la frequenza media delle parole sui gruppi etimologici.

2. Cinque volte più grande nella prima tabella e due volte e mezzo nella seconda.

3. Rango 100; la dispersione: +ox (i segni della dispersione sono: o = testi scientifici; x = testi della stampa; + = testi dalla poesia o dalla prosa artistica); frequenza assoluta: 23.

4. Rango 48; dispersione: ox +; frequenza assoluta: 87.

5. Rango 108; dispersione: +o; frequenza assoluta: 15.

6. Rango 106; dispersione: o; frequenza assoluta: 17.

7. Cfr. D. Macrea, *Terminologia...*, p. 22.

8. Ci potremmo immaginare questo prodotto come « l'area lessicale » occupata di un certo gruppo etimologico.

9. Nella quale f = frequenza totale di un certo gruppo etimologico; n = numero delle parole del gruppo; P = somma dei tutti i prodotti risultati per i gruppi etimologici.

Gruppo etimolo- gico	Indici		0	10	20	30	40	50	60
	I	II							
latino	I	47,0478							
	II	28,5609							
slavo	I	1,0359							
	II	1,4329							
ungherese	I	0,0151							
	II	0,0209							
/neo/greco	I	0,0246							
	II	0,0340							
turco	I	0,0099							
	II	0,0138							
autoctono	I	0,0047							
	II	0,0065							
sconosciuto	I	0,0203							
	II	0,0173							
onomatopeico	I	0,0007							
	II	0,0009							
romeno	I	14,5048							
	II	18,3001							
francese	I	36,8819							
	II	50,9843							
latino dotto	I	0,4100							
	II	0,5672							
italiano	I	0,0305							
	II	0,0421							
tedesco	I	0,0119							
	II	0,0165							
russo	I	0,0015							
	II	0,0021							
inglese	I	0,00017							
	II	0,0002							

biamo calcolato un « indice di circolazione etimologica » (C.) per ciascuno dei gruppi etimologici dei nostri testi.

Ecco questi indici calcolati secondo i dati delle due tabelle precedenti ¹ :

Come si vede indici più importanti presentano solamente gli elementi d'origine latina, francese e le formazioni romene. Dagli altri gruppi etimologici, solamente il gruppo slavo presenta un indice che oltrepassa un poco l'unità, il resto dei gruppi situandosi sotto questo indice. Dunque, anche secondo questi dati, l'elemento latino è predominante nel lessico di circolazione della lingua romena letteraria. Considerando solamente le parole semantiche l'aspetto cambia a favore dell'elemento neologico, dove predominano i vocaboli d'origine francese.

Per poter farci un'idea esatta sopra la struttura etimologica del lessico romeno in circolazione, dobbiamo completare i dati della nostra analisi sul piano sincronico, con altri parametri, tra i quali, in primo luogo, sono la stabilità e la dispersione.

La stabilità nel tempo di certe parole (che si mantengono in un costante rango di frequenza), e la loro dispersione quasi uniforme nei diversi aspetti della lingua (stilistici o territoriali) le assegnano uno stato tutto speciale.

Nella lingua romena solamente gli elementi d'origine latina mantengono generalmente il rango della frequenza in tutte le epoche ed in tutti gli aspetti della lingua.

Uno studio simile al nostro, fatto su testi provenienti dall'inizio del Seicento oppure dalla metà del Settecento ci fornirebbe certamente i stessi dati in ciò che riguarda gli elementi latini ereditati, ma ci offrirebbe risultati completamente diversi per gli altri gruppi etimologici (mancherebbero le parole d'origine francese, mentre quelle d'origine neogreca, slava ecc. presenterebbero dati completamente diversi). Questo fatto dimostra la posizione speciale dei vocaboli latini originari che costituiscono l'elemento invariabile del lessico romeno.

I prestiti stranieri di tutti i tempi sono gli elementi secondari, variabili che si situano alla periferia del lessico. Il paragone della frequenza e del numero dei neologismi e dei prestiti più vecchi conferma la nostra asserzione: il rapporto tra il numero delle vecchie parole non-latine e

1. Indice I, tracciato sul grafico con una linea continua, rappresenta la situazione dei gruppi etimologici rapportata all'intero vocabolario. Indice II, tracciato sul grafico con una linea puntata, rappresenta la situazione delle parole semantiche.

quello dei neologismi $\left(\frac{613}{2950} = 0,20\right)$ è quasi uguale a quello tra la frequenza dei vecchi prestiti e la frequenza totale dei neologismi $\left(\frac{2034}{10700} = 0,19\right)$. Tutto ciò dimostra l'uguale importanza che hanno questi elementi nel lessico di circolazione della lingua ².

Concludiamo dunque — dai dati ottenuti — che le parole eterogene « si mutano » ³, le più recenti prendendo il posto di quelle più vecchie. Questi cambiamenti si situano però nella stessa « banda di frequenza ».

Sempre molto interessanti sono le costatazioni riguardando le tendenze del lessico romeno attuale.

La lingua letteraria arricchisce il suo vocabolario con nuove parole che man mano divengono più importanti nei diversi stili dell'idioma. I neologismi sono penetrati proprio nel fondo più frequente del lessico dove essi rappresentano numericamente quasi la metà delle parole ed un terzo dalla frequenza delle parole semantiche.

Il fondo neologico è costituito quasi completamente dalle parole latino-romanze, tra le quali la maggior parte sono di origine francese, oppure sono state imprestate tramite la lingua francese. Questo fatto ha determinato un rinforzamento della struttura neolatina della nostra lingua che — grazie agli influssi rinnovatori « consona » ancor meglio con le altre lingue romanze.

Constant MANECA.

1. Non abbiamo calcolato le formazioni romene e quelle onomatopeiche.
2. Osserviamo lo stesso fenomeno anche se paragoniamo le frequenze medie delle due categorie di parole: 3,31 per i prestiti vecchi e 3,66 per quelli recenti.
3. Seguendo, generalmente, il destino delle nozioni che denominano.