

ZORA OBSTOVÁ
(Università Carlo IV, Praga)

Esiti di un processo unicizzante o parole storicamente sprovviste di autonomia collocazionale? Uno sguardo alla diacronia delle cranberry words in italiano¹

The present study deals with extremely restricted collocability in Italian and analyses the phenomenon from a historic perspective. By following the diachronic development of a sample of lexemes and word-forms with strongly limited combinatorial potential and often lacking in both syntactic and semantic autonomy (these forms are usually referred to as cranberry words or monoclocable words), extracted from corpora, and the fixed combinations in which they occur, it tries to shed light on the origin of these peculiar words and on the lexicalisation processes of multiword expressions with a monoclocable component in Italian. One of the principal aims of the paper is to verify the widespread hypothesis that cranberry words are the result of a diachronic process leading to a progressive loss of combinatorial possibilities of originally autonomous lexemes. The diachronic analysis shows that the situation is much more complex: only some of the examined words can be considered archaisms (or historisms), whereas many monoclocable forms are characterised, historically, by combinatorial restrictions, appearing in the language already as components of multiword units. Furthermore, some of the studied examples suggest that the “unicalisation” process is not unidirectional and that there is a constant exchange between free lexemes and distributionally idiosyncratic items contributing to an incessant enrichment of the lexicon.

1. *Introduzione*

Il fenomeno della collocabilità estremamente ristretta di parole e forme flesse in italiano è finora rimasto ai margini dell’attenzione degli studiosi e risulta ancora poco esplorato. Alcune lingue dispongono ormai di liste di *cranberry words* – o *parole monoclocabili* (ai quali mi riferirò d’ora in poi con l’abbreviazione PM), usando il termine coniato da František Čermák nel contesto della linguistica ceca – e di studi approfonditi su questo tema: citiamo qui, fra tanti altri, gli importanti contributi di Dobrovol’skij (1988), Fleischer (1989) e Dobrovol’skij e Piirai-

¹ Il presente studio è stato finanziato dal progetto dell’Università Carlo IV *Progres 4, Lingua nei cambiamenti di tempo, di spazio e di cultura*. Desidero ringraziare i due revisori anonimi di *Linguistica e Filologia* per i loro preziosi commenti e suggerimenti.

nen (1994) per il tedesco², di Trawiński et al. per il tedesco e l'inglese (cfr. CoDII) e di Čermák (2014) per il ceco. Delle PM in italiano si sono finora occupati solo Veland (2004, 2006), Konecny (2014) e recentemente, nell'ambito di una ricerca più ambiziosa effettuata su corpora in quattro lingue (cfr. Čermák et al. 2016), anche Obstová (2016a, 2016b, 2017, in stampa). In tutti gli studi sopracitati (eccetto quello di Konecny) il fenomeno della monoccollocabilità viene analizzato sotto il profilo sincronico; infatti, il raggio collocazionale dei lessemi varia nel tempo e una parola può essere considerata monoccollocabile solo sincronicamente, rispetto ad altre parole che dispongono, in un determinato momento storico, di più ampie possibilità combinatorie.

Scorrendo le liste di queste parole molto particolari, per lo più prive di autonomia semantica e sintattica e limitate nella loro distribuzione a pochissime combinazioni con altre parole o addirittura a una sola locuzione, come *vanvera*, *malapena*, *squarciagola*, *battibaleno*, *perdifiato*, *malincuore*, *auge*, *bilico*, *briga* o *espiatorio* (cfr. Veland 2004, 2006 e Obstová 2016), sorge spontanea la domanda: come sono nate le locuzioni contenenti una PM? Quali sono i presupposti che portano a un restringimento così radicale del raggio combinatorio di una parola?

Qualsiasi riflessione diacronica sulle combinazioni lessicali con elementi non autonomi viene inevitabilmente resa più complicata dalla loro enorme eterogeneità (sotto vari profili, cfr. Obstová 2017). Come osserva Veland (2004: 331), “la categoria dei vocaboli non autonomi ha contorni sfilacciati. Non può essere altrimenti, dato che essi appartengono a due gruppi: parole storicamente sprovviste di autonomia collocazionale, categoria cui sembra appartenere il già citato *battibaleno*, e parole che in altre epoche hanno avuto più ampie possibilità combinatorie ridottesi poi successivamente a una sola nell'italiano corrente di oggi”.

Secondo Fleischer (1989: 118-120) le PM³ attuali sono esiti di un processo unicizzante (*Unikalisierungsprozess*) che è graduale e può avere modalità diverse:

² Un riassunto esauriente della ricerca precedente su *Unikalia* in tedesco è fornito da Stumpf (2015, cap. 4, accessibile on line da <https://www.peterlang.com/view/9783653956368/chapter04.xhtml?print>)

³ Fleischer e molti altri studiosi che si sono finora occupati del fenomeno intendono “l'unicità” come caratteristica propria di alcune locuzioni idiomatiche (infatti parlano di *Phraseologismen mit unikaler Komponente*), mentre la nostra accezione di PM è più vasta (si veda sotto).

1. le PM sono unità lessicali diventate obsolete nell'uso autonomo (l'autore fornisce esempi dal tedesco come *Trutz* nella locuzione *zu Schutz und Trutz*, 'per la protezione e la difesa'). Infatti le PM vengono tradizionalmente considerate "relitti lessicali" (*tradierte Sprachrelikte*, cfr. Dobrovolskij 1988: 87) che originariamente possedevano autonomia semantica e sintattica, appartenevano al linguaggio quotidiano e si comportavano come lessemi liberi; tuttavia, col passare del tempo hanno perso la loro autonomia, conservandosi solo in un determinato ambiente sintagmatico. Si tratterebbe dunque di arcaismi (o "storicismi"), parole che vengono percepite come sempre più desuete fino a sparire completamente dall'uso (cfr. il termine "necrotismi", coniato da Asimova 1963: 207 ss.). Ciò non implica che le locuzioni nelle quali le PM si sono conservate debbano essere considerate arcaiche o periferiche dai parlanti di oggi: anzi, non di rado si tratta di espressioni frequentissime;
2. alcune PM sono parole di origine dialettale entrate nel lessico comune (*Hucke* nella locuzione idiomatica *j-m. die Hucke voll hauen*, 'dare un sacco di botte a q.');
3. le PM possono essere prestite da lingue straniere (*ad absurdum*);
4. una PM può passare nel lessico comune provenendo da un linguaggio settoriale: il significato originario della parola è solitamente sconosciuto ai parlanti non specializzati che pertanto percepiscono spesso la rispettiva locuzione come non trasparente (*den Drehwurm haben*, 'sentirsi girare la testa');
5. una PM può nascere, nonostante questa possibilità sia del tutto marginale, come parte di un fraseologismo d'autore, una costruzione letteraria creata *ad hoc* che entra a far parte del lessico quotidiano (*Blüenträume* in un verso di Goethe *nicht alle Blüenträume reifen*, 'non tutti i sogni fioriti maturano').

In questo contributo, basandomi su un campione delle PM più frequenti nell'italiano attuale estratte dai corpora (si veda sotto), cercherò di verificare se e fino a che punto le ipotesi proposte da Fleischer (e da altri studiosi) per il tedesco valgano anche per l'italiano. Intraprendendo un viaggio a ritroso nel tempo, seguirò lo sviluppo di alcune PM italiane e delle loro possibilità combinatorie, dalla loro prima attestazione fino al giorno d'oggi, per dare una risposta alle seguenti domande:

- I. quando e in che modo queste parole sono entrate nel lessico italiano? Continuano direttamente una parola latina o prelatina oppure sono lessemi creati all'interno del sistema linguistico italiano? Sono state attinte alle lingue classiche o sono prestiti da un'altra lingua straniera o da un dialetto italiano?
- II. come si andava sviluppando il loro raggio combinatorio? Erano, dal momento della loro apparizione, caratterizzate da restrizioni di collocabilità o potevano invece, in epoche più remote, combinarsi liberamente con altre parole, subendo solo successivamente – in diversi momenti storici e per vari motivi – un processo più o meno lungo di monoccollocabilizzazione?
- III. valgono per l'italiano le possibilità ipotizzate da Fleischer e quale è il percorso più tipico che porta alla nascita delle locuzioni monoccollocabili?

Il lavoro è strutturato come segue: nel paragrafo 2 viene presentato in sintesi il fenomeno della monoccollocabilità sotto il profilo sincronico, nel par. 3 viene descritto il campione delle PM analizzate e le fonti sulle quali si è basata la ricerca diacronica, il par. 4 è dedicato al modo in cui le PM studiate sono giunte in italiano, nel par. 5 si cerca di seguire lo sviluppo del raggio collocazionale di alcune di esse nella storia. I risultati vengono infine riassunti nel par. 6.

2. *Le PM dal punto di vista sincronico*⁴

Parole monoccollocabili⁵ (PM) o *cranberry words*, come esse vengono designate in inglese (in analogia con *cranberry morphemes*, cfr. Bloomfield 1933 e Aronoff 1976), sono lessemi o forme flesse caratterizzate da un raggio collocazionale estremamente ristretto e una distribuzione limitata a pochissime combinazioni lessicali (o addirittura a una sola). Queste parole sono spesso prive di autonomia semantica e sintattica: essendo difettive del proprio significato, non vengono definite nei dizionari (semanticamente è definibile solo l'unità polirematica a cui fanno

⁴ In questo capitolo viene offerto solo un riassunto elementare della problematica: per le analisi più dettagliate delle PM nell'italiano attuale si veda Obstová (2016b); in una prospettiva contrastiva con il tedesco, Obstová / Vachkova (2017) e con il cecco, con particolare riguardo alle differenze tipologiche, Obstová (2017 e in stampa).

capo) bensì accompagnate dall'indicazione “usato solo nella loc.” e simili. La collocabilità ristretta può riguardare il lessema (tutte le sue forme) o solo una forma flessa o una variante morfologica, mentre le altre forme non presentano particolari limiti di collocabilità (ad es. *quartier generale x quartiere*). Alcune PM occorrono, per lo meno nella lingua attuale, solo in una forma flessa – in italiano spesso mancano del singolare o del plurale (*squarciagola*) e in lingue con flessione nominale fortemente sviluppata, ad es. in ceco (e nelle lingue slave in generale), possono comparire perfino in un solo caso di declinazione⁶. Queste forme non possiedono realmente il lemma (nonostante la loro forma di citazione venga spesso artificialmente costruita o ricostruita nei dizionari)⁷ e risultano difficilmente classificabili dal punto di vista delle parti del discorso: tali anomalie hanno condotto alcuni studiosi a contestare il loro stesso statuto di parole (cfr. Filipec-Čermák 1985: 174, che le designano come *verboidi*).

Le PM sono state finora studiate prevalentemente nell'ambito della fraseologia come costituenti di locuzioni idiomatiche. Recenti ricerche basate sull'analisi dei corpora (cfr. Čermák et al. 2016) hanno tuttavia permesso di riconsiderare alcuni aspetti del fenomeno, mostrando che un numero non trascurabile di queste parole è riscontrabile anche nei termini tecnici e nelle collocazioni (o in quelle combinazioni che Cosevri [1971] definisce solidarietà semantiche) e confermando il carattere scalare della monoccollocabilità. Infatti vanno considerate monoccollocabili non solo parole – del resto rarissime – occorrenti in una sola locuzione, ma anche lessemi o forme flesse con più collocati che formano un paradigma collocazionale molto ristretto e in genere chiuso, contenente solo pochi elementi⁸ (ad. es. *correre/sparare/battere [cuore] all'impazzata*).

⁵ Termine usato in linguistica ceca, coniato da Čermák (1982, 2014 ecc.). In tedesco si parla di *unikale Komponenten* o *Unikalia*, per l'italiano Veland (2005, 2006) propone l'espressione *componenti a collocazione unica*.

⁶ Cfr. la locuzione idiomatica ceca *jít k DUHU* ‘giovare a q.’, in cui il nome esiste solo nella forma dativa singolare (*duh). Si fa menzione qui della flessione nominale considerando che la maggioranza delle PM nelle lingue che abbiamo finora analizzato è attribuibile, almeno formalmente, alla categoria del nome o dell'aggettivo.

⁷ Questa pratica lessicografica viene esplicitamente criticata da Čermák (2014: 10-12).

⁸ Secondo Čermák et al. (2016: 7, 17) il paradigma collocazionale delle PM è solitamente formato da 1 a ± 7 elementi.

3. *Il campione delle PM analizzate e fonti utilizzate per la ricerca diacronica*

Per la ricerca diacronica mi baserò sulla lista delle PM elencate in Čermák et al. (2016: 89-97), estratte dai corpora (per l'italiano è stato utilizzato il CORIS⁹ con 130 milioni di parole) tramite un'analisi automatica basata sull'indice Herfindahl-Hirschman (HHI)¹⁰. Questo indicatore è in grado di fornirci informazioni su quanto sia ristretto il raggio collocazionale delle singole parole: considerando la natura scalare del fenomeno e il fatto che l'analisi diacronica delle 300 PM italiane estratte dal CORIS sarebbe un'impresa troppo impegnativa, ho deciso di limitarmi in questa sede alle PM con i valori più alti dell'indice (HHI a sinistra /L1/ o a destra /R1/ della $PM \geq 9$)¹¹, caratterizzate dalla monocollocabilità pressoché assoluta. Le 27 PM che soddisfano questa condizione vengono indicate qui in ordine alfabetico e sottolineate nelle rispettive locuzioni. La lista riflette l'enorme eterogeneità, già menzionata nei paragrafi precedenti, della categoria: accanto a locuzioni verbali o avverbiali troviamo anche locuzioni nominali, che hanno spesso carattere di termini tecnici, e locuzioni grammaticali:

per autonomasia, anidride carbonica, a casaccio, (andare, tenersi) a braccetto, datore di lavoro, in extremis, (essere, tenersi, venire, tornare) a galla, all'incirca, (correre, sparare, battere) all'impazzata, all'insù, a malapena, a malincuore, in primis, a priori, quartier generale, mettere a repentaglio qc., decreto rettorale, a ridosso di qc., (andare, viaggio) a ritroso, (cantare, gridare, urlare) a squarciagola, stragrande maggioranza, in subordine, a suon di qc., un tantino (di qc.), all'unisono, (i) non vedenti, (fecondazione) in vitro.

⁹ CORIS = Corpus di Italiano Scritto [in linea], http://corpora.dslo.unibo.it/coris_ita.html, gentilmente messo a disposizione dal prof. F. Tamburini per l'analisi automatica.

¹⁰ HHI è un mezzo statistico, usato in economia principalmente per misurare il grado di concorrenza presente in un determinato mercato. Questo indice, adattato da Cvrček (2013) agli usi della linguistica dei corpora, permette di quantificare la diversità del contesto di una parola. In Čermák et al. (2016) sono state prese in esame solo parole con frequenza assoluta superiore a 200, considerando che per frequenze più basse l'analisi automatica potrebbe non generare risultati statisticamente affidabili.

¹¹ L'indice HHI può assumere valori da 0 a 1: più alto è il suo valore, più omogeneo (formato da meno tipi) dovrebbe essere il contesto immediato della parola esaminata. La maggioranza delle parole nel corpus raggiunge valori molto bassi: il valore medio di tutte le *word-forms* nel CORIS è 0,16; il valore 0,3 viene superato solo da 2650 forme delle 23 395 analizzate.

Cercare di ricostruire, almeno a grandi tratti, la strada che queste parole hanno percorso dalla loro nascita fino ai nostri giorni non è certo un compito da poco: nonostante ci si possa basare su eccezionali opere lessicografiche¹² come il *Dizionario etimologico della lingua italiana* (DELI), *Dizionario etimologico italiano* di C. Battisti e G. Alessio, il *Grande dizionario della lingua italiana* di S. Battaglia e il *GRADIT*, ma anche su altri dizionari come il *Sabatini Colletti* che riportano le date di prima attestazione delle parole, su archivi elettronici come la preziosissima *Letteratura italiana Zanichelli* (LIZ), sostituita recentemente dalla *Biblioteca italiana Zanichelli* (BIZ)¹³, le informazioni che se ne possono ricavare sono spesso insufficienti per tracciare con precisione le tappe di un eventuale processo di monoccollocabilizzazione. Le date e le fonti della prima attestazione finora conosciuta indicate nei dizionari sono solo provvisorie e soggette a ulteriori precisazioni e ancora più problematico è lo studio del contesto delle parole esaminate: le occorrenze trovate nella LIZ risultano spesso troppo basse per essere statisticamente rilevanti e non permettono di stabilire con sicurezza in quale momento una parola perde definitivamente l'autonomia sintattica (o/e semantica) e viene ormai usata esclusivamente come componente di una locuzione. Inoltre la LIZ contiene solo opere letterarie, mentre il CORIS, il corpus sincronico e bilanciato da cui le attuali PM sono state estratte, è composto da vari tipi di testi e tra i sottocorpora domina quello della stampa. Infine, il problema di gran lunga più spinoso resta la scarsità di testimonianze sulla lingua parlata e sui rapporti tra la lingua orale e la lingua scritta nelle epoche più o meno remote, elemento che impedisce di osservare il comportamento delle PM nell'uso quotidiano.

Nonostante tutte queste difficoltà ho deciso di tentare l'impresa sperando che le fonti di cui sopra mi permettano di tracciare almeno un quadro approssimativo e di individuare le tendenze principali che portano alla monoccollocabilizzazione delle parole in italiano.

¹² Un insostituibile aiuto per le ricerche di questo tipo rappresenterà sicuramente in futuro, quando sarà portato a termine, *Il Lessico etimologico italiano* (LEI), edito dal 1979 dalla *Akademie der Wissenschaften und der Literatur* a Magonza.

¹³ La più recente BIZ non sembra differire molto dalla LIZ per quanto riguarda i testi contenuti: dopo un confronto delle due biblioteche digitali ho deciso di lavorare con la LIZ perché permette di visualizzare meglio alcune informazioni, specialmente quelle sulle occorrenze totali delle PM cercate.

4. L'entrata delle PM analizzate nel lessico italiano

Se consideriamo il modo in cui le PM analizzate sono comparse nel lessico italiano, possiamo dividere le 27 parole esaminate in quattro gruppi principali. Le date di prima attestazione finora conosciuta, tratte da dizionari etimologici (spec. dal DELI) e segnate tra parentesi, sono solo indicative. Viene qui citata la data di prima attestazione della parola (forma) indipendentemente dai suoi possibili significati: cambiamenti di tipo semantico e problemi legati alla polisemia saranno discussi più in dettaglio nel paragrafo seguente. Le date, anche se solo approssimative, danno una prima idea delle notevoli differenze sull'asse temporale presenti tra le PM studiate, anche all'interno dello stesso gruppo.

4.1. Neoformazioni

La classe più numerosa è rappresentata da parole formate in seno al sistema linguistico italiano mediante i processi di derivazione, composizione e conversione. A questa categoria sono attribuibili 14 delle 27 PM esaminate, citate qui in ordine cronologico a partire dalla data di prima attestazione:

vedente (av. 1321), *malincuore* (av. 1342), *insù* (av. 1342), *incirca* (av. 1499), *ridosso* (av. 1540), *casaccio* (1540), *impazzata* (av. 1600), *tantino* (av. 1642), *braccetto* (1666), *stragrande* (av. 1744), *malapena* (1809), *squarciagola* (1873), *rettorale* (1950), *subordine* (1960).

È interessante notare che quattro PM – *braccetto*, *casaccio*, *stragrande*, *tantino* – sono forme alterate¹⁴ mediante i suffissi *-ino*, *-etto* e *-accio* e il prefisso *stra-*. Due PM (*subordine* e *ridosso*) sono derivate con i prefissi *sub-* e *ri-*¹⁵, mentre l'aggettivo *rettorale* è formato con uno dei suffissi relazionali più produttivi (*-ale*).

Due delle PM possono essere percepite come risultato di un processo di conversione piuttosto che di derivazione¹⁶: *impazzata* dal participio passato femminile al nome (*impazzata* nella locuzione avverbiale *all'i*.

¹⁴ L'alterazione sembra svolgere un ruolo importante nel processo unicizzante anche in altre lingue, come dimostra Piirainen (1996: 325) sull'esempio di alcuni diminutivi in tedesco.

¹⁵ Un'altra ipotesi etimologica considera *ridosso* un nome deverbativo dal latino parlato *redossiāre*.

¹⁶ Sulle peculiarità dei nomi formati sulla base del participio passato e presente cfr. Grossmann-Reiner (eds.) (2004: 338, 357).

viene abitualmente considerata un sostantivo, la possibilità che si tratti di un'espressione ellittica con l'aggettivo – di tipo *alla svelta* – pare improbabile); *vedenti* dal participio presente all'aggettivo e poi, attraverso un processo di ellissi che comporta l'assorbimento del significato della testa del sintagma nominale, al nome.

Tra le PM formate per composizione troviamo un composto V+N (*squarciagola*). Nel caso delle rimanenti quattro PM (*malapena*, *malincuore*, *incirca*, *insù*) sarebbe più opportuno fare ricorso al termine di univerbazione¹⁷.

4.2. *Prestiti lessicali*

È abbastanza numerosa la categoria dei prestiti, integrati e non integrati (10 PM), dominata da latinismi, giunti in italiano in epoche diverse. Alcune PM sono componenti di locuzioni latine, assunte in blocco¹⁸: (*in primis* (sec. XVII), (*in extremis* (1612), (*a priori* (1630) e (*in vitro* (1933). Come voci dotte vengono qualificate anche le parole *datore* (av. 1292), *antonomasia* (sec. XIV) e *unisono* (1561). Tre PM sono probabilmente di origine francese: *quartier*¹⁹ (av. 1348), *repentaglio*²⁰ (av. 1388) e *carbonica* (il franco-latinismo *carbonico* è attestato nel 1795, il termine *anidride carbonica* nel 1889).

4.3. *Parole del fondo latino*

Le parole che continuano direttamente un vocabolo latino sono solo 3: *suon* (*suono*: fine sec. XIII), *ritroso* (av. 1292) e *galla* (1320).

Le date di prima attestazione mostrano che le parole caratterizzate, sincronicamente, dalla collocabilità estremamente ristretta sono entrate in italiano in momenti diversi: molte (quasi 40%) sono attestate negli

¹⁷ Per i termini di composto e univerbazione qui adottati si veda Grossmann-Reiner (eds.) (2004: 51).

¹⁸ Nonostante la scelta possa risultare discutibile, ritengo che anche le locuzioni di origine straniera possano essere „scomposte“ e le loro componenti analizzate, dal punto di vista sincronico, sotto il profilo collocazionale. Oltre a molti costituenti di locuzioni latine e francesi risultano monocollocabili numerose parole inglesi (queste però, avendo il valore HHI leggermente più basso, non vengono analizzate nel presente studio) come *mountain bike*, *mailing list*, *fast food* ecc. Perfino le componenti delle locuzioni straniere subiscono, nel corso del tempo, cambiamenti del raggio combinatorio. Si può ad es. osservare una progressiva „demonocollocabilizzazione“ di alcune PM: ad es. accanto a *fast food* troviamo, anche se con frequenza minore, *slow food*.

¹⁹ Cfr. la nota 24.

²⁰ Secondo una delle due ipotesi etimologiche più solidamente accreditate, cfr. sotto § 5.1.

stadi più antichi della lingua, tuttavia un numero non trascurabile (14%) delle PM compare in tempi abbastanza recenti, alla fine dell'Ottocento e nel corso del Novecento.

5. Sviluppo del raggio collocazionale delle PM analizzate

Se vogliamo esaminare l'evoluzione del contesto sintagmatico ed eventuali cambiamenti diacronici del raggio collocazionale delle attuali PM, dobbiamo tener conto delle diverse accezioni delle parole esaminate e dei mutamenti di tipo semantico o formale che si sono prodotti nel corso dei secoli. Alcune PM, specialmente quelle del fondo latino, hanno subito importanti estensioni di significato, sviluppando un alto grado di polisemia. La tendenza inversa può essere invece osservata nelle PM di origine straniera giunte in italiano nell'ambito di una locuzione idiomatica o di un termine tecnico, che nella maggioranza dei casi non sono state soggette a ulteriori variazioni.

Nelle pagine seguenti cercherò di esaminare a grandi tratti l'evolvere del contesto sintagmatico delle singole PM basandomi sui dizionari menzionati sopra (spec. il DELI e il *Grande dizionario della lingua italiana* di S. Battaglia) e innanzitutto sul corpus dei testi raccolti nella *Letteratura italiana Zanichelli* (LIZ). Per ragioni di spazio non descriverò qui tutte le PM, bensì mi limiterò a illustrare le tendenze generali più rilevanti su alcune di esse, soffermandomi solo occasionalmente su qualche dettaglio che mi sembra significativo.

5.1. Prestiti

Le PM prese in prestito da altre lingue vanno distinte dal punto di vista collocazionale almeno in due categorie: i prestiti non adattati che sono giunti in italiano nell'ambito di una locuzione idiomatica (spesso letteraria) o di un termine tecnico e non hanno, nella maggioranza dei casi, subito cambiamenti considerevoli né a livello formale, né a quello semantico (solo in alcuni casi si possono osservare delle estensioni di significato dell'intera locuzione) e i prestiti pienamente integrati, il cui comportamento combinatorio non ha seguito ovviamente uno scenario uniforme, essendo legato alle vicende diacroniche di ogni singola parola.

In PRIMIS, in EXTREMIS

La locuzione *in primis*, dal latino *in primis rebus*, attestata già in Aretino e nei testi della LIZ talora seguita, per maggior efficacia, da *et ante omnia*, veniva e tuttora viene usata come equivalente di ‘anzitutto’, ‘in primo luogo’.

La locuzione *in extremis* (un’ellissi di *in extrēmīs diēbus/tempōribus*) appare nei testi italiani dal XVII sec. nel senso di ‘in fin di vita’. Dal 1905 è attestata anche nel significato ‘all’ultimo momento’ ed è proprio questa accezione che domina nel corpus attuale, specialmente nei testi giornalistici (“accordo *in extremis* alla Camera”).

In VITRO

Risultano poco soggetti ai cambiamenti anche i tecnicismi come *in VITRO*, documentato per la prima volta nel 1933 nei *Quaderni del carcere* di Gramsci – qui comunque in senso figurato (‘in modo astratto, teorico’):

L’uomo trasforma il reale e non si limita a esaminarlo sperimentalmente in vitro per riconoscerne le leggi di regolarità astratta.

(A. Gramsci, *Quaderni del carcere*, Quaderno 15, §51)²¹

Oggi, almeno stando ai dati del CORIS, questa locuzione viene usata prevalentemente in contesti medico-biologici per indicare vari tipi di sperimentazione in provetta, in particolare nel termine *fecondazione in v.*

Mettere a REPENTAGLIO qc.

Lo sviluppo diacronico di altri prestiti, specialmente quelli molto antichi e integrati, è più difficile da seguire, come nel caso della parola *repentaglio* (attestata nella locuzione *mettere a r.* prima del 1388, come lessema autonomo tuttavia solo nel 1804). L’origine francese di questa PM (dal francese antico *repentaille*) viene – come menzionato sopra nella nota 20 – contrastata da alcuni studiosi che ne propongono la derivazione da un neutro plurale latino ricostruito **repentalia*²². Nei testi della

²¹ Consultabile online su [http://www.nilalienum.com/Gramsci/QC\(GS\)int.html#QUADERNO_15](http://www.nilalienum.com/Gramsci/QC(GS)int.html#QUADERNO_15).

²² Cfr. il DELI, voce *repentaglio*, e http://www.treccani.it/magazine/lingua_italiana/articoli/parole/repentaglio.html.

LIZ questa PM figura solo nella locuzione *mettere* (ev. *porre*) *qc. a repentaglio* (o anche *ripentaglio*), talvolta rafforzata con *grave/gran*. Più spesso vengono messe a r. la vita o l'onore, o uno mette a r. se stesso:

Per una donna non si mette a *repentaglio* l'onore.

(C. Goldoni, *L'amante militare*, At.1, sc.11.7)

Gli usi autonomi della parola (nel senso di 'grave rischio, pericolo'), cfr.

Tale fu l'antica costumanza; nulladimeno i Re Longobardi, tra per l'uso de' sacramentali e per la pratica delle giadette purgazioni volgari, sforzaronsi di restringere al possibile il *ripentaglio* del duellare, riducendolo a pochi casi, siccome si può vedere da molte loro leggi (C. Pecchia, *Storia civile e politica del regno di Napoli da servire di supplimento a quella di Pietro Giannone*),

probabilmente abbastanza rari, sono documentati solo dall'inizio dell'Ottocento e si può ipotizzare che siano il risultato di un processo inverso rispetto a quello che porta alla lessicalizzazione di espressioni monocollocabili; si potrebbe dunque trattare di autonomizzazione (o democollocabilizzazione) di un elemento originariamente non autonomo²³.

QUARTIER generale

Un'altra PM di origine francese è *quartier*²⁴. La parola *quartiere*, caratterizzata da un alto grado di polisemia, presenta nell'italiano attuale restrizioni collocazionali solo nella variante apocopata, la cui esistenza sembra – almeno stando ai dati dei corpora – più o meno limitata alla locuzione *quartier generale*. Nei testi più antichi, come ci mostrano le concordanze della LIZ, la forma apocopata veniva usata autonomamente e in accezioni diverse, specialmente nei componimenti poetici (qui sotto come termine araldico):

²³ Per considerazioni più dettagliate sul processo di "autonomizzazione" delle PM si veda Stumpf (2015, cap 4, accessibile on line da <https://www.peterlang.com/view/9783653956368/chapter04.xhtml?print>).

²⁴ Il DELI e il GRADIT derivano la parola direttamente dal francese *quartier*, mentre Battaglia indica: „der. di quarto, sul modello del fr. quartier“. Il modello francese è ricalcato anche nelle locuzioni *chiedere q.*, *dar q.* e *senza q.* (*demander quartier*, *donner q.*, *sans q.*), cfr. DELI, voce *quartiere*.

Vengo a te per provar, se tu m'attendi, come ben guardi il *quartier* rosso e bianco; che, s'ora contra me non lo difendi, difender contra Orlando il potrai manco.

(L. Ariosto, *Orlando furioso*, Canto 18, 149)

Nei testi settecenteschi comincia a prevalere l'accezione militare del termine (la prima attestazione conosciuta di *quartier generale* è del 1601): *quartier mastro* (2 occ. + 4 occ. di *quartier-Mastro*), *quartier d'inverno* (6 occ.) e *quartier generale*. Dall'Ottocento in poi, la forma *quartier* occorre nella LIZ esclusivamente nella locuzione *q. generale* (24 occ. su 69 occ. totali di *quartier*). Esaminando invece la forma piena *quartiere* (408 occ.), la combinazione *quartiere generale* è presente solo in 8 attestazioni.

DATORE di lavoro

Un caso particolare è rappresentato dalla parola *datore*, considerata una voce dotta e attestata – nel senso di ‘chi dà, chi concede, chi elargisce (doni, ricchezze, elemosine, beni), chi procura (pace, gioie, felicità, grazie)’²⁵ – prima del 1292. Questa PM che sembra essere usata per secoli senza particolari vincoli di collocabilità, compare nell'italiano attuale quasi esclusivamente nella locuzione nominale *datore di lavoro* (assolutamente marginali altri collocati: ad es. *d. di vita* – 8 occ., *d. di ipoteca* – 2 occorrenze su un totale di 3 489 nel CORIS) che è un calco dal composto tedesco *Arbeitgeber*, introdotto in italiano nel 1923 da Mussolini²⁶.

5.2. Neoformazioni

Cantare, gridare/urlare a SQUARCIAGOLA

Abbastanza stabile risulta anche il comportamento collocazionale della maggior parte delle neoformazioni. Nei testi letterari raccolti nella LIZ, molte delle PM citate nel §4.1. sono generalmente limitate nell'uso alle rispettive espressioni monoccollocabili correnti tutt'oggi e non ho trovato indizi né di eventuali cambiamenti formali né di estensioni di significato. Ad es. il composto V+N *squarciagola* compare, sin dalle pri-

²⁵ Cfr. S. Battaglia, *Grande dizionario della lingua italiana*, voce *datore*.

²⁶ B. Mussolini, *Scritti e discorsi*, Milano, III, 1934, p. 76.

me attestazioni, nella locuzione *a squarciagola*, usata prevalentemente con i verbi *cantare*, *gridare/urlare*. Nella LIZ lo troviamo, sia pure marginalmente, anche con altri verbi che implicano l'uso della voce, come *vociare*, *strillare* o *rispondere*. Questo tipo di struttura formale (V + PREP A + N [V+N]) è tra l'altro riscontrabile anche in altre espressioni monocollocabili che tuttavia non erano oggetto della presente ricerca, ad es. (*correre*) *a perdifiato* o (*ridere*) *a crepapelle*²⁷.

Andare, prendere/prendersi, tenere/tenersi a BRACCETTO

Analogamente, il diminutivo *braccetto*, il cui uso sincronico pare più o meno limitato alla combinazione con la preposizione *a* e un piccolo gruppo di verbi (*andare, prendere/prendersi, tenere/tenersi* e molto meno frequentemente *camminare/passeggiare, uscire a b.*), nonostante i dizionari ne diano più accezioni²⁸, dimostra simili restrizioni di collocabilità anche nei testi raccolti nella LIZ. Infatti la PM compare autonomamente solo nel plurale *braccetti*, col significato di 'piccole braccia' (3 occ.), mentre la forma singolare (158 occ.) è attestata esclusivamente come componente della locuzione *a b.*, tuttavia la gamma dei verbi, con cui si combina, è più vasta: accanto ai verbi tipici come *prendere/prendersi* e *tenere/tenersi* troviamo anche *essere, entrare, ripassare, precedere, dondolarsi, tornare, condurre* e in alcuni casi il verbo è omesso:

I ben pensanti sul tardi cominciarono a farsi vedere di nuovo per le strade; l'arciprete dinanzi al caffè; Peperito su e giù pel Rosario; Canali *a braccetto* con don Filippo verso la casa della ceraiuola [...].

(G. Verga, *Mastro-don Gesualdo*, Parte 2, cap. 3.47)

Mentre nella LIZ prevale di gran lunga l'accezione letterale della locuzione ('con il braccio stretto a quello di un'altra persona'), nell'italiano attuale è molto frequente anche l'uso figurato (nel senso di 'essere molto d'accordo con q.'), specialmente nei testi giornalistici:

²⁷ Queste PM con alto grado di monocollocabilità non raggiungono nel CORIS la frequenza minima di 200 occorrenze indispensabili per l'analisi automatica (cfr. nota 10).

²⁸ Cfr. il GRADIT: "braccetto (s.m.), 1. dim. di braccio, 2. BU tipo di nuotata in cui la testa rimane fuori dall'acqua e le gambe eseguono una battuta ad ogni bracciata, 3. TS mar. cavo fissato ai velacci e ai controvelacci per orientarli. Polirematiche: a braccetto, loc. avv. [...]". Nel CORIS ho trovato solo quattro occorrenze dell'uso autonomo di *braccetto*, utilizzato come diminutivo di *braccio* nell'accezione di 'parte allungata e sporgente, mobile o fissa (di mezzi, oggetti, costruzioni)'.

Ma attenzione, l'entusiasmo deve *andare a braccetto* con la cautela.
(CORIS, sottocorpus Stampa)

È interessante notare che *braccetto* come lessema autonomo è attestato, secondo il DELI, dal 1666 (al f. pl. dal 1550), mentre la locuzione *a b.* viene ricondotta al precedente, non toscano, *a braze* (1496, Forlì; cfr. il DELI, voce *braccio*).

A CASACCIO

Anche la parola *casaccio*, derivata con un suffisso peggiorativo, sembra essersi lessicalizzata in breve tempo nella locuzione *a casaccio*. Nonostante la parola venga documentata dal Battaglia nel significato di 'accidente increscioso, malaugurato':

Maestro Mignamau da Norcia è d'opinione, che ella si finisse di far nera per l'esequie di Mecenate, che le voleva grandissimo bene: ed a quest'esequie intervenne quel *casaccio* strano, per conto del procedere.

(L. Panciatichi, *Cicalata in lode della padella e della frittura*)

non troviamo nessuna traccia del suo uso autonomo nella LIZ, dove questa PM è riscontrabile solo come componente della locuzione *a c.* (più una occorrenza di *per c.*), in funzione avverbiale, nel senso di 'senza ordine, alla rinfusa', o in quella attributiva, nel senso di 'sconsiderato, stravagante' (corrispondendo generalmente, in entrambi i casi, alla loc. *a caso*):

M'aggiravo *a casaccio* fra le maschere, ora spingendo, ora spinto, allorché sentii tirarmi per le falde dell'abito. (G. Verga, *Eva*, [3],1)

Anzi mi ricorda infra gli altri d'un certo Messer Michele Marullo, il quale fu un di quei Greci che si fuggiron di qua per la perdita di Costantinopoli, che era dottissimo, secondo che si diceva, e niente di manco era un certo uomo *a casaccio* e fantastico [...].

(G. B. Gelli, *I capricci del bottaio*, Ragionamento 4.61)

A MAPALENA, a MALINCUORE

Quasi scontate sono invece le restrizioni di collocabilità nelle univerbazioni²⁹. La parola *malapena*, datata dai dizionari 1809, compare nella LIZ anche nei testi più antichi, sempre in combinazione con la preposizione *a*, nel significato di 'a fatica', 'appena':

²⁹ Per la nozione di univerbazione qui adottata si veda la nota 17.

E ho fatto prova di venderla a credenza e promettovi non ne troverei a malapena sedici.

(*Motti e facezie del piovano Arlotto*, Facezia 133.3)

Questa locuzione avverbiale che nella LIZ troviamo in frequenza nettamente superiore in grafia separata, *a mala pena* (70 occorrenze contro 16 di grafia unita), nasce come “quasi superlativo di appena”³⁰; le due varianti sono spesso presenti anche all’interno dello stesso testo.

Situazione analoga anche per la PM *malincuore*, attestata nella LIZ in 36 occorrenze su 40 totali nella locuzione *a malincuore*³¹.

STRAGRANDE maggioranza

Molte delle neoformazioni studiate sembrano dunque appartenere alla categoria delle PM storicamente sprovviste di autonomia collocazionale. Bisogna tuttavia prendere in considerazione che tutte le parole finora menzionate in questo paragrafo sono componenti di locuzioni avverbiali o verbali, mentre altre PM, specialmente quelle che fanno capo a combinazioni di tipo nominale, possono presentare uno scenario diverso. L’aggettivo *stragrande*, attestato prima del 1744, si combina nell’italiano attuale quasi esclusivamente con il nome *maggioranza* (635 su 648 occorrenze totali nel CORIS³²). Pertanto, nel Battaglia troviamo un’ampia gamma di accezioni (l’uso odierno corrisponde all’accezione 2b, ‘molto consistente, quasi totale’, eccezionalmente a l’accezione 2a, ‘straordinariamente elevato, molto rilevante [una quantità, un numero]’) e anche i testi della LIZ ci danno, nonostante il basso numero di attestazioni, un quadro complessivo divergente da quello attuale: *stragrande* viene usato qui per lo più in funzione enfaticizzante, e abbastanza liberamente, con sostantivi diversi (come *valore*, *ingegno*, *spettacolo*, *ricchezza*, *cuore*), spesso in postposizione e anche in funzione predicativa:

³⁰ Come, tra l’altro, anche la locuzione *a gran pena*, cfr. *Vocabolario degli Accademici della Crusca*, voce *a mala pena* (<http://www.lessicografia.it/index.jsp>).

³¹ 36 occ. *a malincuore*, 3 occ. *di malincuore* (tutte e tre nei testi di Dossi) e 1 occ. *con tanto malincuore*.

³² Analoga la situazione nel corpus ItWac (corpus di testi tratti dal web contenente 1,5 miliardi di parole) molto più consistente: 11 700 occorrenze su 12 158 totali. Tra gli altri collocati hanno una certa rilevanza solo *parte* con 98 occ. e *quantità* con 24 occ., marginalmente l’aggettivo si combina anche con altri nomi, ad es. *s. numero*, *abbondanza*, *superiorità*, e può comparire anche in testi informali: “Un bacio stragrande a tutti!”

Del fazzoletto tutto ciò che mi ricordo si è che costò 14 paoli: ma bisogna avvertire che era *stragrande*, e che lo speditore fu Paccapelo.

(G. Leopardi, *Lettere*, II - 813 A Paolina L. (1832)

La combinazione *stragrande maggioranza* è nei testi della LIZ del tutto assente³³. Sembra dunque che questo aggettivo abbia subito un radicale restringimento del raggio collocazionale nel corso del XX secolo; sarebbe sicuramente interessante chiarire più in dettaglio le modalità e la dinamica di questo processo.

(I) *non VEDENTI*

Una monocollocabilizzazione recente caratterizza anche il participio *vedente* (e più marcatamente ancora la sua forma plurale *vedenti*) che viene usato nei testi della LIZ, in funzione di aggettivo o di nome, nel significato di ‘che, chi è in possesso della facoltà visiva’ ma compare, fin dai testi più antichi, altrettanto nella forma negativa come eufemismo ‘chi è affetto da cecità’ (qui sotto nel senso figurato):

Quale è colui, ch’adocchia, e s’argomenta di veder eclissarlo sole un poco, che per veder *non vedente* diventa.

(Dante, *Par.* 25, 120)

È proprio in questo significato che la combinazione si è lessicalizzata nell’odierno linguaggio burocratico: il termine *non vedente* (agg. o s.) viene registrato solo nel 1970 da Zingarelli e secondo il DELI è derivato dal francese *non voyant*, attestato in Francia fin dal XII secolo e presente anche nei dialetti³⁴.

5.3. *Parole del fondo latino*

L’ultima categoria raggruppa parole e forme che continuano direttamente un lessema latino o prelatino. Lo sviluppo del raggio collocazionale di queste PM, nel nostro campione meno numerose rispetto a quelle delle altre due categorie, risulta più complesso, considerando la lunga storia che queste parole hanno alle spalle e l’alto grado di polisemia di alcune di esse.

³³ Non ho trovato nemmeno la variante con l’aggettivo in postposizione che viene invece citata dal Battaglia: “Chi ancora (*la maggioranza stragrande*) poltre indifferente e atomica del mestiere, sotto la pressione economica e morale del tempo attuale” (Rebora).

³⁴ La combinazione *non vedente* potrebbe dunque essere in un certo senso considerata un calco, al pari di *datore di lavoro* (si veda sopra).

A SUON di qc.

Una delle quattro PM analizzate in questa sede è la forma apocopata *suon*, che, usata autonomamente, può essere considerata, nell'italiano attuale, variante letteraria o poetica, richiesta da esigenze eufoniche, ma risulta quasi assente nell'uso quotidiano (13 occ. nel CORIS, per lo più citazioni di testi poetici). Questa forma è invece lessicalizzata nella locuzione preposizionale *a suon di* (364 occ.), utilizzata oggi molto più frequentemente nel senso figurato ('seguito, accompagnato da qc. '), spesso come sinonimo della locuzione 'a forza di' ("quando ci buttava a terra *a suon di* schiaffi"); "pronte a contenderselo *a suon di* miliardi la prossima primavera")³⁵ che nel significato letterale di 'con l'accompagnamento musicale' ("*a suon di* musica si camminava con passo più marziale"), nel quale viene spesso soppiantata dalla forma piena, preceduta dall'articolo determinativo (307 occ., ad es. "gli uccelli sul ramo non ballavano più *al suono del tamburo*"). È ovvio che per seguire le modalità del processo di lessicalizzazione di una locuzione preposizionale contenente una forma letteraria e poetica, la raccolta dei testi letterari della LIZ non potrà fornirci molte informazioni in merito. L'alto numero di forme apocopate nei testi – soprattutto in quelli poetici – della LIZ (1677 occ. in totale) rende difficile un'analisi più dettagliata. Per questo motivo ho limitato la ricerca alla prosa: nella metà delle 208 concordanze trovate la forma viene usata come lessema libero, in 106 occ. compare come componente della locuzione preposizionale (*a suon di*, 68 occ., *al suon di*, 38 occ.). Prevale nettamente (più del 90% delle concordanze) il significato 'con l'accompagnamento musicale', infatti tra i collocati più frequenti troviamo *tromba/e*, *tamburo* (come nell'esempio proposto sopra), *campana/e*. Molto meno rappresentata è l'accezione figurata della locuzione (combinazioni come *a (al) s. di bastone/bastona-te/busse*), sviluppatesi più tardi: il DELI la data 1543 (G. B. Gelli) ma nella LIZ la troviamo già nella *Mandragola* di Machiavelli:

Dipoi ci travestiremo voi, Liguria, Siro et io, e andrencene cercando in
Mercato nuovo, in Mercato vecchio, per questi canti, et il primo

³⁵ Interessante da questo punto di vista è l'ampia gamma di collocati appartenenti a diversi campi semantici: *a s. di botte, pugni, schiaffi* e sim. (più di 50 occ.), *a s. di miliardi, milioni, dollari* ecc. (più di 40 occ.), *a s. di musica* (26 occ.) ma anche *a s. di gol (goal), debiti, emendamenti*.

garzonaccio che noi troviamo scioperato, lo imbavaglieremo, et a suon di mazzate lo condurremo in casa, et in camera vostra al buio.

(N. Machiavelli, *Mandragola*, At. 2, sc. 6.30)

Nei testi prosaici la forma apocopata nella locuzione si alterna abbastanza liberamente, in tutti i significati descritti sopra, con quella piena – *a suono di* (circa 40 occorrenze) e *al suono di* (più di 300 occ.), che viene usata con una più vasta gamma di collocati (“uomo fiero e selvaggio che non si sia intenerito *al suono* delle lusinghe”).

Mentre nei testi della LIZ risulta frequente, nella locuzione preposizionale, la variazione tra la forma piena e apocopata (e la variabilità riguarda anche la presenza del determinante), nell’italiano attuale la forma troncata si è lessicalizzata soprattutto nel significato figurato della locuzione (nel senso di ‘a forza di’) e la perdita della vocale finale potrebbe essere interpretata, insieme all’assenza di articolo, come sintomo di una fase avanzata di lessicalizzazione della locuzione³⁶.

(Essere, tenersi, venire, tornare) a GALLA

Nonostante i dizionari ci offrano, per la voce *galla*, significati diversi (piuttosto specialistici)³⁷, la maggioranza di queste accezioni non sembra più essere presente nel lessico mentale dei parlanti comuni. Nell’italiano attuale la parola è invece molto frequente nella locuzione *a galla*, che si combina a sua volta con un gruppo ristretto di verbi come *stare, tornare, tenersi, venire*, e all’interno di questa locuzione viene percepita come opaca e immotivata. La PM *galla*, che nella sua accezione botanica (‘cecidio’: si veda la nota 37) continua la parola latina *galla(m)*, è attestata nei testi più antichi e la sua fortuna nel lessico potrebbe essere legata al fatto che l’infuso di galle veniva usato, fin dall’età romana, per la produzione dell’inchiostro. Prima del 1292 è documentata anche la locuzione *a galla*, ‘sul pelo dell’acqua’. Infatti la galla, essendo molto leggera, se immersa in un liquido non va a fondo ma rimane in superficie:

³⁶ Cfr. Situazione analoga, forse, anche per l’altra forma apocopata, *quartier* (si veda sopra al §5.1.).

³⁷ Cfr. il GRADIT, voce *galla*: “1. TS bot. rigonfiamento, deformazione delle foglie, dei rami e delle radici delle piante provocato da parassiti animali o vegetali, 2. BU fig., persona o cosa leggera, inconsistente, 3. TS med. piccola vescica che si forma sulla pelle, spec. a causa di una scottatura, 4. TS vet. => 2molletta, 5. OB LE estens., ghianda; pillola a forma di ghianda” (+ polirematiche a g.)

Io confesso d'esser pesato e molte volte de' miei di essere stato; e per ciò, parlando a quelle che pesato non m'hanno, affermo che io non son grave, anzi son io sì lieve, che io sto *a galla* nell'acqua.

(G. Boccaccio, *Decameron*, Conclusione dell'autore, 10)

Col tempo la locuzione, usata con verbi diversi, assume anche significati figurati (*stare a galla*: 'galleggiare' ma anche 'astenersi', 'tenersi lontano' (G. Boccaccio 1348-53); *venire a galla*: 'scoprirsi', 'manifestarsi' (av. 1735); *tenersi a galla*: 'fare quanto basta per tener fronte agli impegni' (av. 1850); *rimanere, stare a galla*: 'salvarsi da situazioni critiche' (1887)³⁸.

Nella LIZ la parola *galla* viene documentata nella locuzione *a g.* in 127 attestazioni su 179 occorrenze totali. L'uso autonomo (nel significato di 'cecidio' ma anche di 'ghianda' o 'una cosa leggera')³⁹ è relativamente frequente fino ai testi del Cinquecento ma diventa assolutamente marginale nei secoli successivi, lasciando spazio alla monocollocabilizzazione della parola nella locuzione *a g.*

A RITROSO

La parola *ritroso*, così come la PM *galla*, presenta accezioni diverse, di cui alcune tecniche o obsolete⁴⁰. Sincronicamente può essere considerata monocollocabile – sia come aggettivo, sia come nome – solo nella forma maschile singolare⁴¹, dato che 405 sulle 422 occorrenze totali di *ritroso* nel CORIS fanno capo alla locuzione avverbiale *a r.*

Questa locuzione, attestata da Dante in poi, si basa sull'accezione 3 della voce *r.* nel GRADIT (si veda la nota 40) che viene qualificata obsoleta e continua direttamente l'aggettivo latino *retrosus*, 'volto all'in-

³⁸ Si veda il DELI, voce *galla*.

³⁹ La forma è riscontrabile anche come agg. nei significati diversi di 'dei Galla' o di 'gala'.

⁴⁰ Cfr. GRADIT, voce *ritroso*: "CO 1. agg., s.m., che, chi non è socievole per modestia o timidezza [...] 2. agg., restio, riluttante: è ritroso a parlare in pubblico [...] 3. agg. OB che ha movimento inverso 4. s.m. TS pesc., venat. in alcuni tipi di rete, imboccatura a forma di imbuto volta all'interno, attraverso cui un animale entra facilmente, ma non può più uscire 5. s.m. OB vortice d'aria o d'acqua" (+ polirematica *a ritroso*).

⁴¹ Le altre forme non presentano particolari limitazioni di collocabilità. Va comunque osservato che la frequenza degli usi autonomi di questa parola è relativamente bassa (*ritroso* 17 occ.; *ritrosa* 27 occ. – nel sg. spesso nella loc. */non/ fare lo ritroso/la ritrosa; ritrosi* 7 occ.; *ritrose* 7 occ. nel CORIS).

dietro'. Gli altri significati della parola nascono probabilmente più tardi, per estensione:

[...] e dissersi altramente A ritroso, dal latino togliendosi, dalla quale s'è formato il nome et essi detto Ritroso calle e Ritrosa via, come sarebbe quella de' fiumi, se essi secondo la favola ritornassero alle lor fonti; da cui si tolse a dire Ritrosa donna, e Ritrosía il vizio.

(P. Bembo, *Prose della volgar lingua*, 3, 66)

Anche se nell'insieme dei testi della LIZ prevale l'uso autonomo della parola (118 occ. contro 60 occ. nella locuzione *a r.*), relativamente frequente nei testi più antichi, approssimativamente dalla seconda metà dell'Ottocento si può osservare la tendenza inversa: *ritroso* come lessema libero diventa sempre più raro e marginale, fino ad apparire quasi esclusivamente nella locuzione *a r.*, usata spesso con verbi di movimento come *andare, correre, allontanarsi a r.*

6. Conclusioni

Riassumendo i punti essenziali della breve e sommaria analisi diacronica effettuata su alcune PM italiane nei paragrafi precedenti, cercherò adesso di dare risposta alle domande poste nella sezione introduttiva.

6.1. *Quando e come le PM sono comparse nella lingua italiana*

Il gruppo delle PM studiate si presenta molto eterogeneo non soltanto sotto il profilo sincronico ma anche sotto quello diacronico. Tra le PM analizzate prevalgono le neoformazioni, nate all'interno del sistema linguistico italiano, abbastanza rilevante è anche il numero dei prestiti, integrati e non integrati. Sorprende invece la bassa percentuale di parole di fondo latino.

Una parte consistente delle odierne PM è attestata fin dagli stadi più antichi della lingua: non si tratta solo di parole che continuano direttamente lessemi latini ma anche di alcune neoformazioni e alcuni prestiti, specialmente quelli integrati. Ciò nonostante è evidente che le parole caratterizzate oggi da un raggio collocazionale estremamente ristretto compaiono nel corso di tutta la storia linguistica e alcune sono di data relativamente recente.

6.2. Lo sviluppo del raggio collocazionale delle PM nella storia

Le PM vengono abitualmente considerate “relitti linguistici”, lessemi originariamente liberi che hanno subito nel corso del tempo un progressivo restringimento del paradigma collocazionale. Tuttavia, il breve sguardo al contesto delle PM studiate ha rivelato una situazione molto più complessa: un processo diacronico unicizzante sembra riguardare solo una parte (ca. il 52%) delle parole esaminate e le modalità di questo processo divergono da parola a parola, essendo talvolta legate anche ai fattori extralinguistici⁴². Alcune forme vengono usate, anche per secoli, autonomamente e nello stesso tempo anche come componenti di una locuzione (*galla – a galla*), altre si comportano come lessemi liberi, cristallizzandosi solo in un secondo tempo in una combinazione lessicale (*stragrande maggioranza*, [*i*] *non vedenti*). L’alta frequenza d’uso della locuzione e la sua coesione interna contribuiscono probabilmente, insieme a fattori extralinguistici (come potrebbe essere, nel caso della PM *galla*, la sostituzione dell’inchiostro ferrogallico da inchiostri di china?) a marginalizzare l’uso autonomo della parola, la quale – se il referente che denotava tuttora esiste – viene non di rado sostituita, almeno in alcune accezioni, da un altro significante (*galla* → *ghianda*, *impazzare* [in *correre all’impazzata*] → *impazzire*).

Tuttavia in una parte considerevole delle PM analizzate questo processo diacronico monoccollocabilizzante risulta del tutto o quasi assente: storicamente privi di autonomia collocazionale sono, naturalmente, molti prestiti, specialmente quelli non integrati che sono entrati in italiano nell’ambito di una locuzione straniera (*in vitro*), ma anche alcuni prestiti integrati (*mettere a repentaglio*) il cui uso autonomo, assolutamente marginale anche dal punto di vista diacronico, con rare attestazioni in letteratura che sono solitamente di data più recente rispetto alle locuzioni in cui compaiono, potrebbe invece essere interpretato come esito di un processo inverso che porta alla demonoccollocabilizzazione di elementi originariamente fissi. Colpisce invece la limitata capacità combinatoria di molte neoformazioni, attestate anche dal punto di vista diacronico quasi esclusivamente nelle rispettive locuzioni (*a squarciagola*, *all’impazzata*, *in subordine*). Tra le neoformazioni troviamo alcuni no-

⁴² Come ad es. nel caso di *galla* (si veda qui sotto).

mi alterati, innanzitutto i diminutivi: si potrebbe ipotizzare che si tratti, perlomeno in alcuni casi, di processi derivativi che modificano l'intero sintagma, testimoniando un'alta coesione interna della locuzione che poteva esserne alla base (*a caso* → *a casaccio*; cfr. Voghera 1994: 205-206).

6.3. *Valgono per l'italiano le ipotesi proposte da Fleischer (menzionate nel cap. 1)?*

Rispondere in modo univoco alla terza domanda diventa più complesso almeno per due motivi:

1. Fleischer sottopone all'analisi soltanto le locuzioni idiomatiche che sono quasi tutte di natura verbale (anche se nella sua lista di *Phraseologismen mit unikalen Komponenten* troviamo anche *unbeschränkter Bahnübergang*, 'passaggio a livello senza barriere' (cfr. Fleischer 1997: 39), mentre la nostra concezione di monoccollocabilità è più ampia e include anche locuzioni nominali che hanno spesso carattere di termini tecnici (*anidride carbonica*) o locuzioni di tipo grammaticale (loc. preposizionali come *a suon di*);

2. sotto il profilo diacronico, alcune delle PM analizzate sono difficilmente classificabili in una sola categoria. *Galla* nella locuzione *a g.* può essere considerata un "relitto lessicale" ma anche un termine botanico, *in vitro* è un termine tecnico usato in medicina e biologia ma nello stesso tempo un prestito, *l'unisono* è una voce dotta (dunque un tipo particolare di prestito) ma anche un termine della teoria musicale. La fissità distribuzionale del *suon* è di natura piuttosto grammaticale che lessicale ma date le connotazioni letterarie e arcaizzanti di questa forma, la PM potrebbe essere percepita anche come obsoleta, e quindi come un "relitto" del passato. Particolarmente problematica risulta la categoria dei tecnicismi: dal punto di vista diacronico è molto difficile stabilire se una PM era passata nel lessico comune da un linguaggio settoriale (la *galla* come termine botanico?, il *ridosso* come termine marinaresco?) e neanche sotto il profilo sincronico è sempre chiaro se un termine vada considerato estraneo al lessico quotidiano e sconosciuto ai parlanti comuni, data la massiccia presenza di tecnicismi nell'italiano odierno.

Anche se la complessità dello sviluppo diacronico non permette dunque di far rientrare tutte le PM studiate in modo univoco nelle cate-

gorie proposte da Fleischer, si può ciò nonostante concludere che una parte importante di queste parole in italiano è rappresentata da prestiti (integrati e non), molto più numerosi rispetto agli arcaismi o forme considerate obsolete nell'uso attuale. Non ho trovato nessuna parola di origine dialettale (anche se al dialetto può essere probabilmente ricondotto l'uso dell'intera locuzione *a braccetto*) e nessun fraseologismo d'autore, considerato, tra l'altro, una possibilità estremamente rara dallo stesso Fleischer.

Tuttavia, un terzo delle PM analizzate non risulta attribuibile a nessuna delle cinque categorie. Ciò è dovuto principalmente al fatto che tutte le ipotesi etimologiche proposte, forse con l'eccezione di alcuni prestiti, prevedono un processo unicizzante e dunque un progressivo restringimento del raggio collocazionale della forma, mentre, come abbiamo visto, una percentuale importante delle parole esaminate sembra essere storicamente sprovvista dell'autonomia combinatoria. Per questo motivo ho preferito dividere le PM nel campione (in accordo con quello che osserva Veland, cfr. cap. 1) in due categorie principali: forme che hanno subito un processo diacronico monoccollocabilizzante e altre che hanno fatto la loro comparsa in italiano già come componenti di una combinazione lessicale:

	Processo unicizzante	PM storicamente non autonome
Parole del fondo latino	3	0
Voci dotte	3	0
Prestiti integrati	1	2
Prestiti non integrati	0	4
Neoformazioni	7	7
Totali	14	13

Tabella 1. Sviluppo del raggio collocazionale delle PM analizzate.

Nella tabella 1 sorprende l'alto numero di parole storicamente non autonome che rappresentano quasi la metà di tutte le PM analizzate. Considerando il modo in cui le PM sono entrate nel lessico italiano, nella prima categoria troviamo soprattutto le parole del fondo latino (*galla*) e le voci dotte (*datore*), più marginali sono i prestiti integrati (*quartier*) e alcune neoformazioni (*stragrande*). Nel secondo gruppo dominano i prestiti non integrati (*la/ priori*) e in particolare alcune neoformazioni: univerbazioni (*malapena*) ma anche derivati come *braccetto* (un particolare ruolo sembrano avere i nomi alterati, cfr. la nota 14) o composti (*squarciagola*).

Va sottolineato che la presente ricerca non è che un primo sommario tentativo, basato su un numero limitato di PM, di seguire lo sviluppo diacronico del potenziale combinatorio delle parole a distribuzione fissa e i risultati andrebbero verificati su un campione molto più vasto. L'analisi conferma comunque la natura estremamente eterogenea della categoria anche dal punto di vista diacronico e richiama l'attenzione sulla complessità dei processi che portano alla lessicalizzazione delle locuzioni con elementi non autonomi, che non si limitano al progressivo restringimento del raggio collocazionale di un lessema originariamente libero. Inoltre suggerisce che non si tratti sempre di processi unidirezionali, accennando alla possibilità di autonomizzazione di alcuni elementi originariamente fissi, a favore dell'ipotesi che tra lessemi liberi e parole con distribuzione limitata ci sia uno scambio continuo che contribuisce all'arricchimento costante del lessico.

Zora Obstová
Istituto di Studi Romanzi
Facoltà di Lettere e Filosofia dell'Università Carlo IV di Praga
zora.obstova@ff.cuni.cz

Bibliografia

- Amosova, Natalia Nikolaevna, 1963, *Osnovy anglijskoj frazeologii*, Leningrad, Izd. Leningradskogo universiteta.
- Aronoff, Mark, 1976, *Word Formation in Generative Grammar*, Massachusetts, The MIT Press.

- Battaglia, Salvatore, 1961-2002, *Grande dizionario della lingua italiana*, 21 vol., Torino, UTET.
- BIZ, *Biblioteca italiana Zanichelli*, DVD-ROM, 2010, Bologna, Zanichelli.
- Bloomfield, Leonard, 1933, *Language*, London, Allen Unwin.
- CoDII, *Collection of Distributionally Idiosyncratic Items*, accessibile da: <https://www.english-linguistics.de/codii/>.
- CORIS, *Corpus di Italiano Scritto*, elaborato e coordinato da R. Rossini Favretti, accessibile da <http://corpora.dslo.unibo.it/coris_ita.html>.
- Cortelazzo, Manlio / Zolli, Paolo, 1999, *Dizionario etimologico della lingua italiana (DELI)*, Bologna, Zanichelli.
- Coseriu, Eugenio, 1971, "Solidarietà lessicali" (1967). In: *Teoria del linguaggio e linguistica generale*. Sette studi, Bari, Laterza: 303-316.
- Cvrček, Václav, 2013, *Kvantitativní analýza kontextu*, Praha, Nakladatelství Lidové noviny.
- Čermák, František / Čermák, Jan / Obstová, Zora / Vachková, Marie, 2016, *Language Periphery. Monocollocable Words in English, German, Italian and Czech*, Amsterdam, John Benjamins.
- Čermák, František, 2014, *Periferia jazyka. Slovník monokolokabilních slov*, Praha, Nakladatelství Lidové noviny.
- De Mauro, Tullio, 2000, *GRADIT, Grande dizionario italiano dell'uso*, Torino, UTET, 2000.
- Dobrovolskij, Dimitrij, 1988, *Phraseologie als Objekt der Universalienlinguistik*, Leipzig, VEB [Linguistische Studien].
- Dobrovolskij, Dimitrij / Piirainen, Elisabeth, 1994, "Sprachliche Unikalia im Deutschen. Zum Phänomen phraseologisch gebundener Formative". *Folia Linguistica* 28/3-4: 449-473.
- Filipec Josef / Čermák, František, 1985, *Česká lexikologie*, Praha, Academia.
- Fleischer, Wolfgang, 1989, "Deutsche Phraseologismen mit unikaler Komponente. Struktur und Funktion". In: Gréciano, Gertrud (ed.), *Europhras 88, Phraséologie Contrastive. Actes du Colloque International Kligenenthal – Strasbourg, 12–16 mai 1988*. Strasbourg, Université des Sciences Humaines, Département d'Etudes Allemandes, Collection Recherches Germaniques 1: 117-126.
- Fleischer, Wolfgang, 1997, *Phraseologie der deutschen Gegenwart*, Tübingen, Niemeyer.
- Grossmann, Maria / Rainer, Franz (eds.), 2004, *La formazione delle parole in italiano*, Tübingen, Niemeyer.

- Konecny, Christine, 2014, "Unikale Lexeme – 'Spuren' der Diachronie in der Synchronie? Eine Analyse anhand ausgewählter italienischer Phraseologismen". In: Melchior, Luca / Göschl, Albert / Rieger, Rita / Fischer, Michaela / Voit, Andrea (eds.), *Spuren.Suche (in) der Romania. Beiträge zum XXVIII. Forum Junge Romanistik in Graz (18.–21. April 2012)*, Frankfurt a. M., Lang [Forum Junge Romanistik, 19]: 283-298.
- LIZ 4.0. *Letteratura italiana Zanichelli*, CD-ROM dei testi della letteratura italiana, 2001, a cura di Pasquale Stoppelli / Eugenio Picchi, Bologna, Zanichelli.
- Obstová, Zora 2016a, "Parole monocollocabili alla luce dei corpora", "A frequency dictionary of Italian monocollocable words", "An alphabetical dictionary of Italian monocollocable words". In: Čermák, František et al., *Language Periphery. Monocollocable Words in English, German, Italian and Czech*, Amsterdam, John Benjamins: 17-22, 55-68, 89-97.
- Obstová, Zora 2016b, "Fenomeno della collocabilità ristretta nell'italiano di oggi". *Linguistica Pragensia* 26/2: 33-46.
- Obstová, Zora / Vachková, Marie (in stampa), "Aktuelle Korpusforschung zu monokollokablen Wörtern im Italienischen und Deutschen: Bericht und Ausblick". In: Konecny, Christine / Autelli, Erica / Abel, Andrea / Zanasi, Lorenzo (eds.), *Lexemkombinationen und typisierte Rede im mehrsprachigen Kontext*, Tübingen, Stauffenburg [Stauffenburg Linguistik].
- Obstová, Zora, 2017, "Monokolokabilita ve dvou typologicky odlišných jazycích: srovnání češtiny a italštiny". *Časopis pro moderní filologii* 99/2: 225-244.
- Obstová, Zora, (in stampa), "Cranberry words tra tipologia e diacronia: l'italiano e il ceco a confronto". In: Voicu, Roxana / Balaş, Oana-Dana / Gebăilă, Anamaria (eds), *Atti del IV Congresso di fraseologia e paremiologia PHRASIS*, Saarbrücken, Éditions Universitaires Européennes.
- Piirainen, Elisabeth, "Phraseologismen mit gebundenen Formativen: Deutsch-Niederländisch kontrastiv". In: Weigand, Edda / Hundsnurscher, Franz (eds.), *Lexical Structures and Language Use. Proceedings of the International Conference on Lexicology and Lexical Semantics Münster, September 13-15, 1994*, vol. 2, Tübingen, Niemeyer: 319-328.
- Il Sabatini-Coletti, *Dizionario della lingua italiana*, versione CD-ROM, 2005, Milano, Rizzoli Larousse.
- Stumpf, Sören, 2015, *Formelhafte (Ir-)Regularitäten. Korpuslinguistische Befunde und sprachtheoretische Überlegungen*, Frankfurt a. M., Peter Lang.
- Trawinski, Beata / Sailer, Manfred / Jan-Philipp, Soehn / Lemnitzer, Lothar / Richter, Frank, 2008, "Cranberry Expressions in English and in German". In: *Proceedings of the LREC 2008 Workshop: Towards a Shared task for Multilingual Expressions (MWE 2008)*, Marrakech, Morocco: 35-38.

- Veland, Reidar, 2005, “I vocaboli a collocazione unica nell’italiano di oggi”. *Mémoires de la société néophilologique de Helsinki* 68: 331–339.
- Veland, Reidar, 2006, “Il concetto di collocazione unica e il valore di predizione della dicitura ‘solo nella loc.’ in uso nella pratica lessicografica”. *Zeitschrift für Romanische Philologie* 122: 260–280.
- Voghera, Miriam, 1994, “Lessemi complessi: percorsi di lessicalizzazione a confronto”. *Lingua e stile*, XXXIX/2: 185-214.