

La procedura cloze, la ridondanza e la valutazione della competenza della lingua italiana

1. Introduzione

In questo articolo si intendono affrontare i problemi legati all'uso della procedura *cloze* nella valutazione della competenza di una lingua straniera, e in particolare della lingua italiana. Specifico riferimento si farà alla nozione di *ridondanza*, centrale ogni qual volta entrino in questione fattori di prevedibilità, ricostruzione e compensazione linguistica. Il parlante nativo e il discente di una seconda lingua è incessantemente sottoposto alla presenza di enunciati, ascoltati e prodotti da lui stesso, in cui vi sia una riduzione di qualche tipo della ridondanza che solo una competenza *globale* di tipo, come vedremo, *linguistico e metalinguistico* permettono di valicare e portare la comunicazione al successo che quotidianamente ottiene. Ci si concentrerà inoltre in particolare su alcune differenze necessarie alla considerazione dei risultati della procedura *cloze*, applicata come test di valutazione di una lingua straniera, nella lingua inglese e nella lingua italiana.

2. La procedura cloze: funzioni di verifica, funzioni didattiche e criteri di leggibilità nello studio della L₂

La procedura 'cloze' fu introdotta ufficialmente nella letteratura psicolinguistica con il saggio del giornalista Wilson Taylor (*Cloze* 43–64) come metodo di misurazione della leggibilità dei testi. Una *unità cloze* viene definita come qualunque occorrenza di un tentativo riuscito di riprodurre accuratamente una parte cancellata da un 'messaggio' (qualunque prodotto linguistico) attraverso la valutazione della parte mancante, basata sul contesto rimanente (Taylor, *Cloze* 46). Come osserva Stefania Nuccorini (15–16), in un volume interamente dedicato all'uso di questa procedura nella valutazione della lingua inglese, già la parola *cloze* rimanda alla capacità di 'chiudere' il testo incompleto ricostituendone la sua interezza attraverso una *closure* (*close/cloze*). L'assunto iniziale era che maggiore risultasse la capacità dei soggetti di ricostruire correttamente il testo, maggiore sarebbe stato da considerarsi il tasso di leggibilità del testo sottoposto ad esame, poiché il processo di lettura e comprensione include come sua parte fondamentale la capacità del lettore di anticipare e predire quello che verrà nella catena (e questo, come si vedrà nel prossimo paragrafo è anche alla base delle ricerche della teoria dell'informazione).

La procedura può trattare allo stesso modo porzioni di significato e di forma (*pattern*), si tratta di un metodo che, a differenza del sistema di 'completamento delle frasi' (*sentence-completion test* o *fill-the-gap*), non è frutto di una pre-analisi e selezione degli elementi eliminati, sfrutta invece quasi sempre la cancellazione casuale (o regolata secondo meccanismi determinati). L'interpretazione dei meccanismi che l'uso di questa procedura mette in atto non è tuttavia univoca: da una parte Taylor (*Cloze* 1953) la considera il frutto di un processo di completamento simile a quello che avviene nella percezione all'interno della visione dei gestaltisti, mentre per Weaver (16-33) il processo è più simile a un sistema probabilistico all'interno del quadro della teoria dell'informazione. In quest'ultima prospettiva l'uso della procedura cloze come misura della leggibilità di un testo viene considerata in rapporto diretto alla mancanza di ridondanza dei testi, conducendo a una peggiore prestazione dei soggetti (Weaver 25); anche Bensoussan (18-37) ritiene che siano la ridondanza semantica e sintattica a permettere la ricostituzione dei testi mutilati dalla 'cloze' in una prospettiva che mira alla valutazione della comprensione dei testi da parte di soggetti che apprendono una lingua straniera.

Oltre alla sopra citata destinazione originaria della procedura come criterio di leggibilità dei testi (che in questa sede tralascieremo), la cloze è stata considerata un principio di misurazione delle capacità di composizione scritta, principio di misurazione delle generali competenze globali superiori di una lingua (Hinofotis 60). Nella didattica delle lingue (in particolare negli Stati Uniti) la procedura svolge numerose funzioni individuate e studiate nella letteratura relativa: favorire il discente nel collegamento di informazioni date e nuove, farlo riflettere in maniera più approfondita sulla struttura e i contenuti dei testi sotto esame, permettergli una migliore memorizzazione del materiale. Tali funzioni sono ampiamente documentate in studi precedenti (soprattutto nel caso in cui i soggetti abbiano competenze sotto o nella media). Interpretata in questo senso la procedura cloze svolgerebbe una rilevante funzione didattica e non solamente un compito di verifica della comprensione e della competenza attiva dell'allievo.

Una variante interessante, ma problematica da numerosi punti di vista è la cosiddetta *aural cloze*, in cui le porzioni sono eliminate da testi di parlato registrati, mediante sovrapposizioni di rumore bianco o di altri disturbi. La proposta appare indubbiamente interessante, sia per la sua maggiore naturalità rispetto alla cloze scritta (nel parlato spontaneo i fenomeni di disturbo che ci costringono alla ricostruzione sono ordinari), sia per il suo possibile utilizzo in relazione a misurazioni delle abilità di ricezione e *listening comprehension*. Tuttavia questa variante della cloze classica non ha avuto pressoché nessun seguito e nessuna fortuna come test di competenza di una lingua, né di una lingua materna né di una lingua seconda, probabilmente anche in relazione alla maggiore difficoltà nella preparazione e nella somministrazione del test rispetto alla

variante scritta. Il tipo di problemi che sorgono con la *aural cloze* sono in generale gli stessi che concernono la cloze classica, a questi si aggiungono anche specifiche questioni legate alle caratteristiche specifiche del parlato parlato (come la presenza di ipoarticolazione, ellissi e ridondanze, convenzioni, ecc.).

Uno dei punti cardine discussi nella letteratura sulla cloze sembra comunque essere il *grado di globalità* della competenza che l'uso di questa procedura permette di valutare: competenze specifiche, relative a determinate porzioni della lingua (quindi di volta in volta fonologiche od ortografiche, morfologiche, sintattiche, semantiche, testuali o conversazionali e pragmatiche come sostiene per esempio Hughes 161–168), oppure competenze complessive, cumulative e generali che in ciascun caso di completamento mettono in gioco tutti i livelli precedentemente citati o comunque una buona parte.

In relazione a tale questione, che è pertanto anche una delle interrogazioni centrali di questo articolo, ci si è domandati se fosse preferibile una strategia di cancellazione nella cloze chiamata *fixed-ratio* (ossia cancellazione di una parola — o una lettera — ogni cinque, sei, sette, ecc. sempre secondo lo stesso intervallo, denominata anche cloze 'classico') o *natural cloze test* (Brown 93–116), in modo da prendere in considerazione un gran numero di diverse tipologie di strutture linguistiche (come sostengono per esempio Chavez-Oller, mettendo alla prova la competenza su collocazioni, tempi verbali, coesione testuale, ecc.); oppure la strategia cosiddetta della *rational-deletion* (in cui le cancellazioni seguono precise ipotesi e presupposti, per cui gli intervalli non sono regolari, ma sono guidati da modelli appositamente disposti, denominata anche *pseudorandom* o *tailored*, o in italiano cloze 'mirato' [Nuccorini 19–20]) la cancellazione avviene secondo determinati principi al fine di misurare specifiche porzioni di competenza, per esempio l'uso di preposizioni, articoli o verbi (e questa è la soluzione preferita da Bachman 535–55). Si sono tenuti in considerazione differenti fattori che incidono sul risultato dell'applicazione del test come la frequenza di cancellazione, il metodo di calcolo dei punteggi, la lunghezza dei bianchi, la leggibilità dei testi che vengono sottoposti, la loro lunghezza e l'essere parlante nativo o di seconda lingua. Tutti questi fattori incidono in modo determinante nelle scelte pratiche e nell'applicazione del test agli scopi più svariati. In questo articolo non si entrerà nel merito di tali questioni portando la discussione su un livello più astratto e teorico per capire in modo più preciso la procedura cloze cosa misura e come, e in cosa si differenzia la valutazione della procedura in inglese e in italiano. In questa direzione ci si domanda dunque se la procedura cloze misuri abilità globali (*general proficiency*) oppure specifiche capacità di natura diversa a seconda del particolare disegno della cloze (Bachman 535–555), se questa misuri competenze linguistiche di livello alto oppure basso come sostiene (Alderson 59–76) e in che modo eventualmente la valutazione vada tarata nell'applicazione alla lingua italiana.

Caulfield e Smith (58–83) in un saggio molto interessante testano la procedura cloze in comparazione ad altri metodi per la verifica della competenza di una lingua seconda. In particolare si mettono a confronto i test classici usati dalla *Modern Language Association* che mirano alla rappresentazione delle quattro abilità discrete di ascolto, parlato, lettura e scrittura e la classica *interview*, che in modo soggettivo valuta le competenze globali del discente. La procedura cloze è risultata generalmente molto ben correlata con le due precedenti metodologie presentando in più anche un'economicità dei mezzi e dei tempi molto più favorevole alla somministrazione (ma risultati contrastanti ha ottenuto successivamente Brown [93–116]). In questo senso, e per noi giustamente, la procedura cloze viene utilizzata come metodo di misurazione e valutazione delle competenze globali sulla lingua specifica, e non come una tipologia di test mirata allo studio e giudizio su determinate competenze linguistiche su particolari ambiti linguistici (per esempio, l'uso delle preposizioni, dei *phrasal verbs*, ecc.).

Molto interessante in particolare per il nostro studio è risultata la comparazione ulteriore proposta da Caulfield e Smith (58–83) tra cloze test e *reduced redundancy test* (ossia test a ridondanza ridotta): come si vedrà in seguito infatti la procedura cloze in tutte le sue forme e molti altri sistemi di verifica delle competenze linguistiche (sulla prima o seconda lingua) chiamano in causa la nozione di ridondanza, che riprenderemo in maniera più dettagliata nel seguito dell'articolo. Ciò che gli autori chiamano appunto *reduced redundancy test* consiste in un test di ascolto in cui sono inseriti dei disturbi del tipo del rumore bianco.¹ Il presupposto evidente è che per un parlante non nativo sia necessaria la presenza nel cotesto o nel contesto di una maggiore ridondanza per portare a termine il compito di comprensione o completamento e ricostruzione del testo. Il parlante nativo è infatti continuamente sottoposto alla presenza di enunciati ascoltati e prodotti da lui medesimo in cui vi sia una riduzione della ridondanza che solo una competenza globale di tipo, come vedremo, linguistico e metalinguistico permettono di superare e portare a successo la comunicazione.

3. La procedura cloze usata nella teoria dell'informazione

Nel quadro della teoria dell'informazione, elaborata da Claude Shannon tra gli anni quaranta e cinquanta, viene elaborata e applicata una metodologia di quantificazione dell'informazione matematica contenuta nei testi (le prime applicazioni avvennero su codici di trasmissione e successivamente si passò all'analisi di testi in lingue naturali). Numerosi sono stati i tentativi di specificare con maggior accuratezza il valore dell'entropia (così è chiamata la quantità di informazione di un messaggio) dell'inglese scritto e di altre lingue secondo le proposte di Shannon (per la lingua italiana si dispone del lavoro di Manfrino [4–29]), uno dei quali — certamente considerato centrale — è il metodo sperimentale che utilizza la procedura cloze. I primi esperimenti di questo tipo utilizzavano la procedura usando

come unità di misura non la parola, bensì il carattere alfabetico.² Questo metodo di misurazione dell'informazione aveva, rispetto ad altri, lo scopo di ovviare al problema delle dipendenze grammaticali a grande distanza. Si tratta di una serie di esperimenti sulla *predicibilità* delle lingue, basati sull'idea che un parlante possieda implicitamente una buona conoscenza intuitiva della struttura statistica della propria lingua materna.³ L'ipotesi di partenza è che un parlante, di fronte a un qualsiasi testo, di cui siano state eliminate delle porzioni, sia in grado di ricostruire gli elementi perduti con un successo migliore rispetto a una strategia casuale, basandosi sulle proprie conoscenze linguistiche (Bateson 413). Queste conoscenze sarebbero le stesse che permettono, per esempio, al parlante di terminare frasi lasciate a metà o di correggere le bozze di un articolo. Il *guessing game* (come i teorici dell'informazione chiamano questo tipo di variazione della procedura cloze) di Shannon consiste appunto nella richiesta a una serie di soggetti di indovinare la continuazione di una sequenza linguistica data. Uno degli esperimenti di predizione proposti consisteva nel chiedere a un soggetto di indovinare la prima lettera di un testo: se la predizione fosse stata corretta, gli sarebbe stato comunicato il successo e gli sarebbe stato chiesto di continuare a indovinare la seconda lettera. Se la predizione fosse stata scorretta, gli sarebbe stata comunicata la lettera giusta e si sarebbe proceduto alla seconda, e così via. Un esempio di questo procedimento viene proposto da Shannon nel 1951 (*Prediction 547*):

1 THE ROOM WAS NOT VERY LIGHT AS ALL OBLONG
2 ---ROO-----NOT-V-----I-----SM-----OB-----

1 READING LAMP ON THE DESK SHED GLOW ON
2 REA-----O-----D-----SHED-GLO--O--

1 POLISHED WOOD BUT LESS ON THE SHABBY RED CARPET
2 P-L-S-----O---BU--L-S--O-----SH-----RE--C-----

La prima riga riporta il testo originale, la seconda indica con un trattino i caratteri indovinati, con la lettera quelli che non sono stati indovinati. Su un totale di 129 lettere, il soggetto analizzato da Shannon ne ha indovinate 89, cioè il 69%. Secondo la teoria entrambe le righe contengono la stessa quantità di informazione, e in linea di principio dovrebbe essere possibile, dalla seconda riga ridotta, ricostruire la prima.

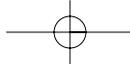
Ciò che è rilevante notare è che il processo di ricostruzione del seguito del messaggio coinvolge *tutte* le proprietà del linguaggio (statistiche e non), per cui il soggetto informante deve mettere in atto strategie euristiche (non molto diverse da quelle che usa quotidianamente) che vanno dalle ipotesi sui nessi consonantici e sillabici ammessi dalla lingua inglese, a

considerazioni sui possibili confini di parola, sulla coerenza semantica, sulla coesione sintattica, ecc. Si mettono dunque in gioco proprietà globali della lingua sottoposta ad analisi e soprattutto meccanismi di riflessione metalinguistica, che nella quotidiana interazione comunicativa sono molto più frequenti di quanto potrebbe apparire.

Altri colleghi di Shannon hanno proseguito su questa strada per la misurazione dell'informazione, utilizzando sistemi ancora più conformi alla classica procedura cloze. Chapanis (496-510) prova a ripetere gli esperimenti di Shannon, per verificare la loro attendibilità e usa un metodo di misurazione della prevedibilità basato sempre sullo stesso sistema. Scelti una serie di passi in lingua inglese, si sono cancellati il 10%, 20%, 25%, 33,3%, 50% e 66,7% dei caratteri attraverso *pattern* regolari e casuali. Ai soggetti veniva chiesto di ricostruire i passi, senza limiti di tempo, e senza la possibilità di consultare dizionari (a differenza dei soggetti di Shannon che facevano uso di tabelle statistiche di frequenza). Del materiale inserito non molto è risultato corretto. Con cancellazioni del 10%, ben il 90% del materiale inserito veniva ricostruito correttamente. All'aumentare della percentuale di materiale cancellato, la percentuale di risposte corrette diminuiva progressivamente. Con il 30% di cancellazioni, circa la metà del materiale inserito era corretto. Con il 50%, soltanto il 10% del testo veniva ricostruito. Quando il *pattern* di cancellazione era casuale e le cancellazioni coprivano una percentuale del testo bassa, allora la ricostruzione appariva più facile che con il *pattern* casuale. Invece per percentuali maggiori di cancellazione con il *pattern* casuale, il testo veniva stato ricostruito con migliori risultati. Gli esiti generali degli esperimenti contrastano con le stime ottimistiche di Shannon. Soltanto se si cancella il 50% del materiale, scegliendo le lettere alternativamente e comunicando lo schema di cancellazione al soggetto, si sono ottenuti risultati concordanti con quelli di Shannon.

Gli esperimenti di Shannon vengono inoltre ripetuti e applicati ad altre lingue da Jamison e Jamison (164-167), su soggetti non madrelingua, mentre altri dimostrano che gli effetti del contesto si fanno sentire in modo significativo sui tentativi di ricostruzione di un soggetto fino a 32 caratteri (circa 6 parole per l'inglese).

Il tasso di ricostruzione aumenta ulteriormente quando la scelta delle porzioni di lingua da sottoporre a test appartengono a lessici specialistici, come in uno studio sui messaggi usati nelle torri di controllo che dimostra che si può giungere fino all'80% nella possibilità di ricostruzione dei testi, se viene fornito il contesto verbale, e al 96% se viene fornito anche quello situazionale (cfr. Frick and Sumby 595-97; Miller 401-20; Fritz and Grier 232-43). Il risultato generale sottolinea come la ridondanza di una porzione di messaggio sia correlata con il suo grado di probabilità, come l'incertezza nella predizione sia correlata con l'esitazione (Taylor, *Cloze* 43-64; Kokkota 115-19).



472 ISABELLA CHIARI

4. *L'anticipazione e la ridondanza*

Come si collega la procedura cloze alla nozione di ridondanza? Per arrivarci passeremo attraverso la nozione di anticipazione:

The cloze procedure is justified on the assumption that a person who is either a native speaker of the language tested or a reasonably proficient non-native speaker should be able to anticipate what words belong in the blanks given the contextual clues of the passage. (Hinofotis 413)

In altri termini si chiama in causa la capacità di un parlante competente in una lingua (materna o seconda) di prevedere, anticipare e quindi ricostruire i messaggi ottenuti in modo incompleto (e questo vale maggiormente per il parlato, anche se la procedura cloze per quest'ambito risulta molto più problematica da applicare per numerose ragioni che qui non affronteremo). Brown (238) esprime in modo chiaro questo punto citando i lavori di Chavez-Oller sulla grammatica dell'aspettativa (*grammar of expectancy*) in cui si sostiene appunto che la procedura cloze misura gli aspetti della competenza linguistica legati all'anticipazione e all'anticipabilità di porzioni di un testo.

In cosa consiste l'anticipazione di una porzione di lingua? Su cosa poggia questa prevedibilità, su specifici fattori ed abilità linguistiche o su una globale competenza di livello alto della lingua? L'anticipazione funziona nelle diverse lingue allo stesso modo, oppure le diverse categorie di porzioni eliminabili da un testo sortiscono effetti diversi a seconda delle caratteristiche strutturali di ciascuna lingua?

Per rispondere a queste domande si userà come tramite o termine medio la nozione di *ridondanza*. Il concetto di ridondanza, già citato di passaggio, viene spesso usato maldestramente e in accezioni negative in relazione a questioni linguistiche. In realtà se colto per alcuni suoi aspetti che lo riconducono alla sua definizione tecnica originaria esso può costituire un cardine della riflessione linguistica, sia a livello teorico, sia a livello applicativo.

La nozione è stata introdotta nella sua accezione scientifica dai teorici dell'informazione (ed è già comparsa nelle pagine precedenti in modo non ancora specificato), e risultava centrale nella teoria di Shannon e nei suoi sviluppi successivi. La ridondanza nasce dunque e si sviluppa appunto all'interno di questo quadro concettuale, tuttavia viene introdotta in linguistica in modi e cornici teoriche diverse. Il contributo essenziale ma controverso fornito dalla teoria dell'informazione è andato quasi perduto per motivazioni diverse: la debolezza dell'estensione della teoria dei codici alla comunicazione linguistica, i maldestri tentativi di estensione della nozione di ridondanza all'ambito linguistico da parte di ingegneri non linguisti, la stretta matematizzazione della nozione e l'insufficiente discussione teorica dei suoi fondamenti. Queste ragioni

hanno fatto sì che alcune intuizioni molto interessanti elaborate in quel quadro teorico andassero disperse o avessero poco seguito.

In teoria dell'informazione la ridondanza veniva definita come è l'opposto speculare della quantità di informazione: se quest'ultima misura il grado di novità trasmesso dal messaggio, la ridondanza ne trasmette le porzioni prevedibili. È stato detto che la ridondanza misura in qualche modo l'inutilità di un simbolo, di un segno o di un messaggio, poiché la sua eliminazione garantisce la presenza della quantità minima sufficiente di informazione. Interpretata in questo modo la ridondanza è strettamente connessa con le applicazioni della procedura cloze poiché è proprio questa a garantire la possibile ricostruzione di testi mutilati, mediante la sua capacità principale di fornire, attraverso diversi strumenti,⁴ prevedibilità ai messaggi linguistici.

I risultati degli esperimenti e dei modelli matematici proposti da Shannon e dai suoi colleghi sulla ridondanza forniscono infatti un'idea solamente complessiva del fenomeno, non spiegando a quali fattori linguistici (o semio-statistici) sia dovuta una caratteristica così macroscopica delle lingue. I dati rilevati, dai calcoli matematico-statistici, dalle approssimazioni probabilistiche e dagli esperimenti, non rendono conto precisamente, per esempio, dello *scarto* tra il tasso di *ridondanza effettivamente calcolabile* (sulle probabilità poligrammatiche della successione dei simboli della lingua, come si può misurare per esempio sui linguaggi artificiali degli di Shannon), e il tasso di *ridondanza globale* ottenuto attraverso procedure sperimentali come la *cloze*. Questo scarto così ampio mette in rilievo molti aspetti della complessità di sistema (non solo dal punto di vista statistico) delle lingue storico-naturali, e sottolinea proprio quanto la prevedibilità dei messaggi linguistici poggi su tutti i livelli di strutturazione della lingua.

All'interno del quadro elaborato dalla teoria dell'informazione sono stati proposti numerosi tentativi di misurazione di tale ridondanza, tuttavia i risultati sono spesso stati viziati da una certa parzialità di veduta del fenomeno, che nelle lingue naturali è molto più esteso e complesso che nei codici di trasmissione. In particolare, mentre i valori di ridondanza individuati per le lingue naturali dai teorici dell'informazione si basano esclusivamente sulle sequenze di caratteri grafici,⁵ le lingue naturali comprendono forme di ridondanza anche a livello morfologico, sintattico, lessicale e in modo meno definito anche a livello situazionale. In parte queste forme di ridondanza più complesse sono già assorbite nel valore grafemico, dato che le frequenze dei grafemi sono in qualche modo un prodotto diretto di generali condizioni morfologiche e sintattiche, in parte però questo dato risulta fuorviante.

Questo fatto è sottolineato per esempio da Hurford nella voce dell'*Encyclopedia of Language and Linguistics* (3481) a cura di R. E. Asher, in cui si sostiene che il metodo della predicibilità utilizzato nella teoria

dell'informazione è "linguistically crude" in quanto non permetterebbe di distinguere fattori sintattici da fattori semantici. Come esempio di questa insufficienza si sostiene che l'elemento mancante della frase *The . . . is the world's largest land-dwelling mammal* è predicibile sulla base del suo contenuto semantico ed extralinguistico, mentre nella frase *The elephant is . . . world's largest land-dwelling mammal* la predicibilità sarebbe basata su fatti puramente grammaticali. In realtà non ci pare affatto agevole separare in maniera così netta le componenti in gioco quando si usano procedure come la *cloze*. Nel primo caso, per esempio, non figurano affatto fattori solamente semantici: anche la categoria grammaticale è perfettamente prevedibile dalla presenza dell'articolo determinativo prima del vuoto, e dalla presenza del verbo dopo, suggerendo la mancanza di un sostantivo singolare, che copra la funzione sintattica di soggetto. È vero sì che una volta determinati questi elementi grammaticali il corretto *item* è selezionato sulla base di criteri semantici, ma non sono *solo* i criteri semantici a guidare la predizione.

Un prima questione centrale per la valutazione della ridondanza ci sembra dunque la distinzione tra una ridondanza *globale*, valutabile attraverso diversi metodi (di cui la procedura *cloze* è solo il più noto), o *macroscopica*, il cui tasso risulta da una somma di diversi fattori; e vari tipi di *ridondanze specifiche*, che si possono osservare (e misurare in alcuni casi) a diversi livelli di analisi di una lingua. È indiscutibile che il primo valore scaturirà direttamente dai secondi essendone la conclusione macroscopica; ed è altrettanto evidente quanto sia compito complesso e (probabilmente) irraggiungibile quello di individuare *tutti* i tipi di ridondanze specifiche che entrano in gioco in una lingua naturale. Un po' meno complesso — ma non così scontato — è il discorso applicato ad alcuni codici più semplici delle lingue. Già nei codici meno complessi infatti sono individuabili diverse forme di ridondanza, a differenza di quello che spesso viene asserito. La ridondanza (nei codici più semplici come anche nelle lingue naturali) risponde in qualche modo alla fallibilità del processo di trasmissione ed è a questo fine che entra in gioco per garantire l'individuazione delle unità del messaggio, la loro non-sostituibilità, la correttezza della successione delle unità, e in generale ogni momento dello svolgimento del processo comunicativo. La ridondanza ovviamente svolge numerose altre funzioni, ma ciò che interessa qui mettere in evidenza è la sua stretta connessione con la prevedibilità e l'anticipazione.

5. La ridondanza, la procedura *cloze* e la lingua italiana

Abbiamo visto fin qui come la ridondanza sia legata a tutti i fattori di prevedibilità che riguardano la ricezione di un enunciato linguistico e quindi come sia precisamente questa a permettere a diversi livelli e con diversi meccanismi la ricostruzione dei '*blanks*' creati mediante la *cloze*

procedure. Si è anche sottolineato come la ridondanza sia 'plurale,' nel senso che si manifesta sotto diverse forme nella struttura delle lingue (per esempio a livello fonologico nelle restrizioni alle sequenze di fonemi — per cui in italiano non esistono nessi triconsonantici in posizione iniziale di parola il cui primo elemento non sia una sibilante; a livello morfo-sintattico, per esempio nel fenomeno dell'accordo che ripete alcune informazioni grammaticali quali genere e numero permettendo dunque la prevedibilità di parti di sequenze prodotte; a livello sintattico, per esempio nelle restrizioni nell'ordine di parole).

La ridondanza è dunque legata all'anticipabilità e assume un valore globale o macroscopico e numerosi diversi modi di manifestarsi concretamente nelle lingue naturali. Questo ci spinge a riflettere sul problema di 'cosa misura la cloze' in termini di ridondanza. Nella letteratura vastissima sulla procedura cloze, come si è detto in precedenza, la questione è stata più volte affrontata: a favore della globalità sono Brown (93–116); Chavez-Oller et alii (181–206) a sostegno di questa ipotesi osservano che gli items della cloze sono 'sensibili' a restrizioni che vanno ben oltre le 5–10 parole e possono raggiungere restrizioni fino alla distanza di 50 parole (oltrepassando ovviamente anche i confini di enunciato o di frase [cfr. anche Jonz 61–83]); a favore della specificità delle abilità messe in causa e misurare dalla cloze è, per esempio, Bachman (535–55).

Se, come abbiamo cercato di mettere in luce, a ogni 'blank' di un passaggio testuale sottoposto alla cloze corrisponde la possibilità di ricostruirlo mediante il ricorso alla ridondanza (o meglio alle ridondanze della lingua), allora rispondere alla questione della globalità o specificità delle abilità misurate dalla cloze corrisponde a una riflessione sulle tipologie di ridondanza che entrano in gioco durante ogni processo di comprensione (e lettura).

Innanzitutto è necessario sottolineare come ogni processo di comprensione sia complesso e si appoggi contemporaneamente su numerosi diversi fattori (De Mauro 47–62) e sia concepibile anche come una sorta di continuo processo di decrittazione di un messaggio codificato o come lo scioglimento di un gioco enigmistico (Chiari 29–37). In questo quadro la comprensione non è quindi considerata come qualcosa che è misurabile in termini assoluti (sì/no), quanto piuttosto una sorta di processo di costruzione progressivo e probabilistico che si serve di fattori di natura differente (situazionali, statistici, linguistici e metalinguistici a un tempo). Le ridondanze sono appunto alcuni (numerosi) di questi fattori di appoggio alla comprensione. Se il problema della cloze viene affrontato in questi termini, risulta scontato che in ogni processo di riempimento di un bianco entrano in gioco diverse abilità (che con spirito wittgensteiniano ci spingeremmo a definire non solo linguistiche ma anche extralinguistiche). La risposta generale alla questione della globalità/specificità sarà dunque in favore della prima.

Sostenere la posizione della globalità (posizione che implica anche il fatto che la cloze si configura come un test di lingua di livello avanzato, non adatto nella sua forma classica a livelli iniziali dell'apprendimento di una lingua) non significa tuttavia eliminare la possibilità di differenziare la prevalenza di alcuni fattori e abilità specifici. Permette invece di diversificare alcune categorie di bianchi usati nelle cloze, di valutarne la difficoltà, e soprattutto di caratterizzare su questa base, anche in modo intuitivo, la differenza tra lingue diverse.

Come esemplificazione della questione si riporta di seguito un esempio di un testo sottoposto alla cancellazione tipica dell'originale procedura cloze classica. Si è usato un programma dedicato proprio alla costruzione di passaggi con la cloze⁶ che utilizza la strategia a intervalli regolari. Si è scelto come testo di esempio un passo tratto da *L'ombra e la meridiana* di Paolo Maurensig (54). Con una cancellazione regolare a intervalli di cinque parole il risultato è stato il seguente:

Dimenticavo di dire che (1)_____ signora Teresa ha avuto (2)_____ bella idea di presentarmi (3)_____ suoi parenti, facendomi passare (4)_____ un suo nipote "ospite (5)_____ di lei per un (6)_____ periodo di convalescenza," e (7)_____, colto di sorpresa, non (8)_____ la prontezza di contraddirla, (9)_____ dato il via a (10)_____ reazione a catena di (11)_____. Ben presto mi ritrovo (12)_____ in un intrico crescente (13)_____ parentele. Incontro persone che (14)_____ di avermi visto nascere. (15)_____ cugini ormai non si (16)_____ più; il cortile è pieno (17)_____ zii al sole, la (18)_____ invasa dalle zie, e (19)_____ gabinetto sempre occupato dai "(20)_____ nipotini." Come se ciò (21)_____ bastasse, ecco che, potenza (22)_____ parte, mi ritrovo ad (23)_____ a braccia tese un (24)_____ venuto, a far da (25)_____ a qualche smarrita famigliola, (26)_____ a consolare qualche sconosciuta (27)_____ in lacrime.

Risposte:

1) la 2) la 3) ai 4) per 5) presso 6) lungo 7) io 8) trovando 9) ho 10) una 11) equivoci 12) impigliato 13) di 14) giurano 15) I 16) contano 17) di 18) cucina 19) il 20) cari 21) non 22) della 23) accogliere 24) nuovo 25) guida 26) o 27) signora

Una rapida occhiata ai bianchi permette di osservare che la procedura classica garantisce una generale varietà nei tipi di parole cancellati, confermando che la cloze originale è un test da destinarsi a studenti di livello intermedio o avanzato, se non si vogliono introdurre troppe modificazioni. In secondo luogo si osserva che in alcuni casi anche il parlante nativo incontra difficoltà e individua più 'soluzioni' per lo stesso

bianco, soprattutto in corrispondenza di verbi e sostantivi. La presenza di collocazioni e locuzioni polirematiche rende alcune risposte ancora più complesse, se affrontate da uno straniero.

Nella vasta letteratura sulla cloze, praticamente interamente dedicata alla lingua inglese, si risolve la questione proponendo una classificazione basata sulle differenze nella valutazione della 'facilità' o 'difficoltà' delle ricostruzioni. In inglese (Nuccorini) è stato sottolineato come, per esempio gli ausiliari, i pronomi, le congiunzioni e gli articoli siano considerati facili da ricostruire (*easy to fill*); mentre i nomi, gli avverbi e i verbi sono considerati difficili (*hard to fill*), mentre a mezza strada si trovano gli aggettivi e le preposizioni (*medium to fill*).⁷

L'osservazione del passaggio esemplificativo sottoposto alla cloze e della struttura della lingua italiana permettono immediatamente di osservare che la situazione della lingua italiana non risulta nient'affatto speculare a quella dell'inglese. Gli ausiliari, i pronomi e gli articoli in italiano non possono essere certamente inseriti in una categoria *easy to fill*, poiché tutte queste categorie in italiano richiedono una flessione (gli ausiliari seguono la flessione verbale complessa per tempo, numero e persona; e i pronomi presentano la difficoltà di aggiungere al numero e al genere anche alcune variazioni di caso), mentre in inglese esigono molte meno scelte da parte del discente. Quindi andranno inseriti in una categoria *hard to fill*, con l'eccezione degli articoli che prevedono una flessione più semplificata e saranno *medium to fill*. Le congiunzioni sono invariabili dunque rimangono nella categoria *easy to fill* come in inglese, e con loro potranno essere inseriti gli avverbi sempre perché invariabili. Le preposizioni sono notoriamente uno dei capitoli più complessi della lingua italiana per uno studente straniero, e dunque anch'essi andranno inseriti nella categoria *hard to fill*: questo succede poiché non vi sono regole strette per il loro uso e vi è una grande variabilità nei diversi contesti, e dunque saranno inserite nella categoria *hard to fill* per la loro poca prevedibilità. Mentre i sostantivi, così come gli aggettivi (che subiscono flessione di numero e in alcuni casi di genere) potrebbero essere inseriti nella classe *hard to fill*, il caso dei verbi è sicuramente peculiare in italiano, poiché oltre alla difficoltà in comune con l'inglese (e in comune con i sostantivi) di possedere numerosi sinonimi e quindi di generare perplessità nel soggetto che esegue il test, vanno anche adeguatamente coniugati, e dunque saranno inseriti in una categoria *very hard to fill*. È scontata l'osservazione che la difficoltà dei bianchi sarà differente da quella appena suggerita, quando il test venga usato con soggetti di lingua materna italiana come misurazione di leggibilità.

Riassumendo la situazione comparativa dell'italiano e dell'inglese si osservano queste categorie intuitive (che ovviamente necessiterebbero di test e di elaborazioni quantitative più precise):

Classi di difficoltà	Lingua inglese	Lingua italiana
<i>Easy to fill</i>	Ausiliari, pronomi, congiunzioni e articoli	Congiunzioni e avverbi
<i>Medium to fill</i>	Aggettivi e preposizioni	Articoli
<i>Hard to fill</i>	Verbi e sostantivi	Ausiliari, pronomi, preposizioni, sostantivi e aggettivi
<i>Very hard to fill</i>	–	Verbi

Sarebbe auspicabile un'analisi sperimentale per determinare per la lingua italiana come vengono riempite in maniera più precisa le tre o quattro classi *easy*, *medium*, *hard*, e *very hard to fill*, è comunque evidente che la complessità del compito di riempire i bianchi della cloze differisce secondo le caratteristiche strutturali della lingua presa in esame.

Riassumendo le principali osservazioni qui presentate si sono messe in luce le principali relazioni tra la procedura cloze e la proprietà delle lingue naturali di essere ridondanti a ogni livello di analisi linguistica. Si è presentata, proprio attraverso la nozione di ridondanza, la procedura cloze come un test essenzialmente globale che misura le abilità linguistiche di livello generalmente intermedio-alto. Si è inoltre osservato che la globalità della cloze non impedisce di concentrare l'attenzione su abilità linguistiche specifiche, proprio perché anche la ridondanza può essere suddivisa in diverse forme e tipologie. E inoltre si sono osservate le principali differenze tra le difficoltà generali che si incontrano nella risoluzione di un test cloze in inglese e in italiano, difficoltà che dipendono dalla diversa codificazione delle differenti classi grammaticali nelle due lingue. Si auspica inoltre che le osservazioni qui proposte vengano messe alla prova mediante test quantitativi appropriati di verifica delle diverse classi di difficoltà individuate e delle attribuzioni delle categorie grammaticali a tali classi.

ISABELLA CHIARI

Università di Roma, "La Sapienza"

NOTE

¹Caulfield e Smith (58–83) al fine della loro comparazione tra la procedura cloze e il *reduced redundancy test* preferiscono per motivi di facilità di costruzione e somministrazione del test la prima. Tuttavia è utile sottolineare che le abilità misurate non sono esattamente le stesse. La capacità di ascolto e di ricostruzione di testi ipoarticolati uditi attraverso disturbi sonori di tutti i tipi costituisce uno dei primi ostacoli che colui che intraprende lo studio di una lingua straniera incontra quando entra a contatto con parlanti nativi. La sua rilevanza non andrebbe quindi assolutamente sottovalutata, per quanto vi si riconosca una maggiore complessità nella preparazione del materiale da testare in laboratorio. Vero è che la cloze non misura globalmente solo abilità di scrittura o lettura ma abilità genericamente linguistiche, tuttavia la componente strettamente acustica verrebbe totalmente eliminata pur costituendo un elemento sostanziale nell'approccio con le lingue seconde.

²Si consideri che uno dei principali motori dello sviluppo della teoria dell'informazione fu proprio la decrittazione di messaggi nemici durante la Seconda Guerra Mondiale. In un saggio reso pubblico nel 1949, subito dopo la pubblicazione della *Teoria matematica della comunicazione*, ma compilato durante la guerra in un rapporto riservato al governo degli Stati Uniti, Shannon indaga proprio sulla struttura matematica dei codici di sicurezza, trattandoli in modo parallelo al suo esame dei sistemi di trasmissione (*Teoria*).

³L'uso di questo sistema implica una serie di considerazioni specifiche sugli utenti della lingua. In particolare si assume che “nella coscienza linguistica del predicente sono compresi dati sulle qualità statistiche della lingua” (Novak e Piotrovskij 332). Inoltre si considera che quando il parlante soggetto all'esperimento realizza una predizione ottimale si possa opportunamente utilizzare questa predizione come una “misura dell'organizzazione della lingua” (Novak e Piotrovskij 332).

⁴Un'indagine più approfondita della storia, teoria e delle manifestazioni della ridondanza nelle lingue naturali, con una ridefinizione puramente linguistica della questione è il risultato della tesi di dottorato in filosofia del linguaggio dell'autrice, intitolata appunto “La ridondanza nelle lingue storico-naturali,” diretta dal prof. Tullio De Mauro e discussa presso l'Università di Palermo nel gennaio 2001, e di cui questo articolo affronta un aspetto specifico.

⁵Una questione centrale è naturalmente anche il fatto stesso che la teoria dell'informazione prenda in considerazione esclusivamente i *grafemi* nelle analisi linguistiche. Innegabilmente i risultati per i fonemi sarebbero diversi, soprattutto per alcune lingue: mentre però per alcuni questo problema sembra poco rilevante, come per Miller e Chomsky (419–491) che affermano che i risultati sarebbero poco modificati da un'analisi condotta sui fonemi piuttosto che sui grafemi, per altri risulterebbe quanto mai determinante (Moles *Percezione*). È chiaro però che il problema richiede soluzioni ad hoc in dipendenza delle operazioni che si intendono fare sul testo, dei metodi, e anche delle lingue prese in considerazione. Un esempio può essere costituito dalle analisi condotte sulla traccia del lavoro di Markov su testi russi al fine di individuare le dipendenze sintagmatiche (probabilità condizionali) delle unità (cfr. Petruszewycz 37–39). Nei casi analizzati il confronto (sempre basato sullo scritto, mai sul parlato) tra i valori ottenuti per le lettere alfabetiche e quello condotto sulla base di una trascrizione fonematica non ha

presentato particolare divergenza nei valori, permettendo di utilizzare, *per quel tipo di esperimento*, anche soltanto i grafemi.

⁶Il programma si può trovare gratuitamente al sito <http://drott.cis.drexel.edu/clozeproze.htm> in versione DOS e Mac.

⁷Per considerazioni più dettagliate sulle parole *easy or hard to replace* cfr. Taylor, *Developments*.

RIFERIMENTI BIBLIOGRAFICI

- Alderson, J. Charles. "Native and Non-Native Speaker Performance on Cloze Tests." *Language Learning* 30 (1980): 59–76.
- Asher, R. E., ed. *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon, 1994.
- Bachman, Lyle F. "Performance on Cloze Tests with Fixed-Ratio and Rational Deletions." *TESOL Quarterly* 19.3 (1985): 535–55.
- Bateson, Gregory. *Steps to an Ecology of Mind*. New York: Ballantine, 1972. 411–25.
- Bensoussan, Marsha. "Redundancy and the Cohesion Cloze." *Journal of Research in Reading* 13.1 (1990): 18–37.
- Brown, J. D. "What Are the Characteristics of Natural Cloze?" *Language Testing* 10 (1993): 93–116.
- Caulfield, J., and W. C. Smith. "The Reduced Redundancy Test and the Cloze Procedure as Measures of Global Language Proficiency." *Modern Language Journal* 65 (1981): 58–83.
- Chapanis, A. "The Reconstruction of Abbreviated Printed Messages." *Journal of Experimental Psychology* 48 (1954): 496–510.
- Chavez-Oller, M. A., T. Chihara, K. A. Weaver, and J. W. Oller. "When Are Cloze Items Sensitive to Constraints Across Sentences?" *Language Learning* 35.2 (1985): 181–206.
- Chiari, Isabella. "La grammatica del rebus: l'immagine e la parola." *Il Gioco: segni e strategie*. Edl Alessandro Perissinotto. Torino: Paravia Scriptorium, 1997. 29–37.
- De Mauro, Tullio. "La comprensione del linguaggio come problema." *Capire le parole*. 1993. Bari: Laterza, 1999. 47–62.
- Frick, F. C., and W. H. Sumbly. "Control Tower Language." *Journal of the Acoustical Society of America* 24 (1952): 595–97.
- Fritz, E. L., and G. W. Grier. "Pragmatic Communication: A Study of Information Flow in Air Traffic Control." *Information Theory and Psychology*. Ed. H. Quastler. Glencoe, IL: The Free Press, 1955. 232–43.
- Fry, Dennis B. "Communication Theory and Linguistic Theory." *IEEE Transactions on Information Theory* 1 (1953): 120–24.
- Hinofotis, F. "Cloze Testing: An Overview." *Methodology in TESOL*. Ed. M. Richards and J. Long. Rowley, MA: Newbury House, 1987.
- Hughes, Arthur. "Conversational Cloze as a Measure of Oral Ability." *ELT Journal* 35.2 (1981): 161–68.
- Jamison, D., and K. Jamison. "A Note on the Entropy of Partially Known Languages." *Information and Control* 12 (1968): 164–67.
- Jonz, John. "Another Turn in the Conversation: What Does Cloze Measure?" *TESOL Quarterly* 24.1 (1990): 61–83.

La procedura cloze 481

- Kokkoto, V. "Letter-Deletion Procedure: A Flexible Way of Reducing Text Redundancy." *Language Testing* 5 (1988): 115–19.
- Mandelbrot, Benoit. "Linguistique statistique macroscopique." *Logique, langage et théorie de l'information*. Ed. Leo Apostel, Benoit Mandelbrot, and A. Morf. Paris: Presses Universitaires de France, 1957. 1–80.
- Manfrino, R. L. "The Entropy of the Italian Language and Its Computation." *Alta Frequenza* 29 (1960): 4–29.
- Maurensig, Paolo. *L'ombra e la meridiana*. Milano: Mondadori, 1998.
- Miller, George A. "Communication." *Ann. Rev. Psychology* 5 (1954): 401–20.
- Miller, George A., and Noam Chomsky. "Finitary Models of Language Users." *Handbook of Mathematical Psychology*. Ed. R. R. Bush, R. D. Luce, and E. Galanter. New York: Wiley & Sons, 1963. 419–91.
- Moles, Abraham. *Teoria dell'informazione e percezione estetica*. 1958. Milano: Lerici, 1972.
- Novak, L. A., and R. Piotrovskij. "Esperimento di predizione ed entropia della lingua rumena." *Accademia delle scienze URSS, Statistica linguistica*. 1968. Bologna: Patron, 1971. 324–63.
- Nuccorini, Stefania. "Il cloze test in inglese." *Ricerca, metodologia, didattica*. Roma: Carocci, 2001.
- Petruszewycz, Micheline. *Les chaines de Markov dans le domaine linguistique*. Genève: Editions Slatkine, 1981.
- Shannon, Claude. "Communication Theory and Secrecy Systems." *Bell System Technical Journal* (1949): 656–715.
- _____. "The Prediction and Entropy of Printed English." *Bell System Technical Journal* 30 (1951): 50–64.
- Taylor, Wilson L. "'Cloze' Procedure: A New Tool for Measuring Readability." *Journalism Quarterly* 30 (1953): 43–64.
- _____. "Recent Developments in the Use of 'Cloze Procedure.'" *Journalism Quarterly* 33 (1956), 42–99.
- Weaver, Wendell. "Theoretical Aspects of the Cloze Procedure." *Toward a Psychology of Reading and Language*. Ed. Wendell Weaver and Albert Kingston. Athens, GA: U of Georgia P, 1977. 16–33.