

Abstract

Il contributo si propone di mostrare l'applicazione della tradizionale prospettiva funzionalista alle forme tipiche delle scritture brevi del web, tra storia della linguistica (Martinet) e teoria matematica dell'informazione (Shannon).

Parole chiave: linguistica funzionale, teoria dell'informazione, Martinet, Shannon

This paper aims to show the application of the traditional functionalist perspective to the typical forms of short writings of the web, combining the history of linguistics (Martinet) and the mathematical information theory (Shannon).

Keywords: functional linguistics, information theory, Martinet, Shannon

1. Informatività, economia della lingua e grammatica funzionale delle scritture brevi²

Evidenti sono, e note, ed anche da noi considerate, le istanze sociali ed economiche che hanno determinato l'insorgere delle scritture brevi via telefono mobile, le quali attengono alla dimensione pragmatica e sociolinguistica dell'atto comunicativo: adesione alle regole del gruppo (che in termini sociolinguistici costituisce la base per lo sviluppo e per il mantenimento delle varietà gergali), costi legati alla lunghezza del messaggio, dimensioni materiali del supporto, minuscola tastiera e piccolo schermo del cellulare, limitazioni dei caratteri disponibili, stile "di fretta" o "in movimento" dello scrivente.

Il trasferimento della pratica della scrittura veloce ai contesti della CMC costituisce un osservatorio privilegiato per l'analisi dell'occorrenza di fenomeni sistematici nell'interscambio comunicativo – sincrono e asincrono - a distanza, manifestazioni che danno vita alla fondazione di varietà pertinenti alla dimensione sociolinguistica diafasica, classificate con l'etichetta di "lingue speciali" o "gerghi di Internet".

La rilevanza assunta dalle scritture brevi per la definizione delle fenomenologie in questione si concretizza nella elaborazione di dizionari e repertori gergali on-line, che valgono sostanzialmente come elenchi di forme grafiche abbreviate, acronimiche (es. tvb, LOL) o troncate (es. complex = compleanno), o genericamente sintetiche, iconiche e simboliche (es. serie degli *smileys* ed *emoticons*).

La natura compilativa delle raccolte alimenta e soprattutto giustifica le modalità non ufficiali della circolazione, evidenziando il carattere per lo più estemporaneo delle forme e in particolare la varietà degli esiti, che contrastano con la nozione comune di vocabolario, inducendo piuttosto la ricerca di una chiave di analisi che possa dare conto della multiforme (e difforme) fenomenologia.

Conseguentemente alla infinita estensione della grande rete mondiale (il World Wide Web), tanti e diversi risultano infatti i dizionari delle abbreviazioni, contenenti elenchi diversificati, spesso non coincidenti neppure al loro interno (comprendono infatti volentieri forme non univoche e dalle realizzazioni multiple, rispondenti a principi apparentemente non omogenei). L'attesa adesione al postulato normativo, convenzionale, risponde alla volontà peculiarmente umana di ridurre a regole la complessità dell'agire linguistico, ma contrasta con l'occorrenza di forme non fisse né stabili.

Per esse, tuttavia, sembra possibile recuperare una dimensione di validità e dei criteri di rigore logico che rispondono ai principi della comunicazione parlata informale/non standard, quale è la condizione concreta della cosiddetta CMC sincrona tra pari: *turn-taking*, *floor-taking*, coesione e coerenza, implicature conversazionali e convenzionali, efficacia, efficienza, informatività, ritmi discorsivi, risparmio dei tempi e dei mezzi.

¹ La concezione del lavoro è unitaria, frutto della collaborazione fra gli autori; solo ai fini di attribuzione formale, Francesca Chiusaroli è autrice del paragrafo 1 e Fabio Massimo Zanzotto è autore del paragrafo 2.

² Il presente contributo procede, sviluppando la già presentata prospettiva funzionalista, dal mio *Scritture brevi oggi: tra convenzione e sistema*, pubblicato nel I Quaderno di Linguistica Zero dedicato a Scritture brevi, a cura di F. Chiusaroli e F. M. Zanzotto (2012): http://www.unior.it/userfiles/workarea_477/Chiusaroli%20e%20Zanzotto%20-%20Scritture%20Brevi%20di%20Oggi.pdf. A tale contributo rinviamo anche per i riferimenti critici che costituiscono anche in questa sede l'orizzonte bibliografico di riferimento.

In particolare questo ultimo riferimento all'economia implicata nella conversazione sul web fornisce un utile inquadramento per l'interpretazione delle forme grafiche qui sotto indagine all'interno di un'ottica che salvaguarda il rendimento funzionale rispetto alla disposizione normativa, che è poi la caratteristica che maggiormente distanzia i nostri testi dalla scrittura tradizionale.

La corrispondenza biunivoca tra suono e segno, inizialmente essenziale prerogativa della scrittura alfabetica, storicamente svigorita a motivo della prevalente logica convenzionale, ha conosciuto nuova occasione di manifestazione e di tenuta nella scrittura veloce sul web, lì dove l'assenza di controllo dall'alto ha come effetto lo scardinamento della norma grafica, ma la necessità (ancor meglio l'urgenza) della comunicazione comporta l'istituzione di regole di funzionamento interne al messaggio (il *tweet*, la stringa della chat), agli attanti (chi scrive e chi legge), alla comunità social (il popolo della rete).

Di qui l'idea di ciò che abbiamo già chiamato una "grammatica delle scritture brevi", che si esplicita nell'equilibrio tra norma e sistema e che soprattutto conduce al massimo grado l'interpretazione funzionalista della lingua e delle sue manifestazioni concrete (*langue e parole*).

La scrittura qui in oggetto può a questo punto fornire uno speciale contesto di applicazione e di verifica delle ipotesi strutturaliste sull'economia della lingua, ciò che cercheremo in questa sede di provare sperimentando l'applicazione di tali presupposti a partire da una delle fonti prime e più autorevoli al riguardo, che è la tesi funzionalista di André Martinet.

La varietà delle produzioni e delle opzioni manifestata nei messaggi della CMC sincrona, la ricorrenza contestuale, anche in uno stesso utente, di forme abbreviate o estese, o anche variamente abbreviate, varianti per altro registrate nel repertori on-line, fanno osservare infatti un comune principio organizzativo fondato sul correlato teorico dell'economia linguistica.

Economia, efficacia, potenza, valore funzionale

Si tratta di un concetto rispetto al quale la linguistica si è posta non certo recentemente.

Risparmiare energia ottenendo il massimo dell'effetto nella espressione e nella comunicazione. Si tratta di un principio che domina la produzione linguistica innanzi tutto verbale, a partire dalla considerazione per cui gli organi adibiti a parlare non sono esclusivamente o originariamente destinati a tale scopo, bensì, più sostanzialmente, sono riservati ad attività connesse all'esistenza stessa, come respirare e nutrirsi.

Se respirare e parlare sono operazioni complementari, escludentisi reciprocamente, per di più vitali, occorrerà organizzarsi per ottimizzare al massimo entrambe le attività, primariamente economizzando le concatenazioni di volta in volta coinvolte.

Sono entrate nel novero dei possibili rilievi alla tesi martinettiana le intraviste posizioni "estremiste", talora evidentemente scollegate dall'osservazione dei reali fatti di lingua, dai quali si evincono non raramente spinte anti-economiciste, tensioni equilibratrici dirette dalla resistenza all'annullamento delle differenze (cfr. Saussure: "nella lingua non ci sono che differenze").

Tali orientamenti paiono massimamente accentuati dal bilanciamento tra norma e uso, tra competenza e esecuzione, tra i piani della grammatica prescrittiva e descrittiva, che contrappongono la tendenza alla cristallizzazione linguistica con la pressione variazionale in sincronia e in diacronia.

In aggiunta a tali complessi divari, ancora maggiore distanziamento si intravede nel rapporto tra la lingua e la scrittura, lì dove le vicende storico-culturali annesse alla diffusione della norma scrittoria e alla standardizzazione grafica hanno determinato un progressivo e definitivo disconoscimento della funzione di trascrizione fonetica originaria delle tradizioni alfabetiche alla loro fondazione.

Pertanto i sistemi grafici nazionali appaiono poco funzionali e spesso non economici, nella misura in cui non riproducono la pronuncia e soprattutto non aderiscono al principio ideale della corrispondenza biunivoca tra suono e segno.

Se il fine della scrittura non è la trascrizione dei suoni (come fatalmente dimostrano i fallimenti di ogni riforma ortografica), lo scopo della riproduzione del parlato è acclarato invece nel caso della scrittura veloce della CMC sincrona, dato il legame essenziale di tale forma comunicativa con il livello orale della lingua, ma ancora più interessante nella nostra prospettiva è la considerazione della rilevanza del principio economico che condiziona la produzione delle forme, parole, frasi, messaggi.

Nel contesto del singolo atto linguistico prodotto, infatti, entrano in gioco variabili sistematiche che, in termini economici e funzionali, sembrano raffigurare la dimensione espressiva maggiormente rispondente all'ideale struttura linguistica: si potrà in tal senso delineare la nostra grammatica delle scritture brevi.

L'argomento è tanto più cogente, ed anche denso di contenuti di attualità, in quanto il postulato dell'economia appare fondamentale connotato della teoria dell'informazione che regola la rappresentazione della conoscenza sul web.

Le modalità di produzione e di diffusione delle informazioni attraverso la rete sono regolate dalle esigenze del risparmio dei mezzi, della riduzione delle ridondanze, e contemporaneamente dal rispetto dei livelli di comprensibilità e dalla corretta veicolazione dei contenuti.

Le antesignane teorie dell'informazione dell'età moderna – dalla *ars combinatoria* di marca lullista, alla *mathesis universalis* leibniziana, attraverso il "carattere reale" testato nella baconiana Royal Society – hanno posto basi sicure per la fondazione di un linguaggio essenziale, scevro dei tratti superficiali e direttamente rinvianti alle cose e alle nozioni, cosa che in termini di scrittura si è tradotta nella elaborazione di grafie sintetiche, spesso logografiche e ideografiche, iconiche e direttamente espressive di concetti.

Ma è soprattutto l'era di Internet ad aver provveduto allo stato di maggiore necessità (in termini socioculturali) per l'applicazione e per il funzionamento dell'assunto economico rispetto al mezzo e al fine della comunicazione implicata.

La velocizzazione richiesta dal mercato per l'utenza va sempre più perfezionando le operazioni di selezione dei contenuti trasmessi e, significativamente per la prospettiva qui assunta, tale operazione, insieme tecnica ed ideologica, trova concretizzazione nei termini delle "scritture brevi".

I sistemi predittivi dei testi, come il T9 per il telefono mobile, Google Instant e Google Suggest per il motore di ricerca, dimensioni con-testuali quali il *tweet* o la stringa fisicamente visualizzata nella finestra delle chat per i social network, inducono e quasi impongono lo sviluppo e la diffusione di fenomenologie precipuamente basate su pratiche di accorciamento e di sintesi grafica per le quali la dimensione informativa costituisce il confine entro cui le tendenze centrifughe alla infinità variabilità sono contenute e comprese.

Per una efficace verifica del postulato economia/informazione qui inteso, citeremo in questa sede le affermazioni contenute nel paragrafo "Economia della lingua" dal volume *Elementi di linguistica generale* (sigla ELG) di Martinet (1960; trad. it. 1966).

Si può concepire l'evoluzione linguistica come governata da un'antinomia permanente fra le necessità di comunicazione dell'uomo e la tendenza che egli ha a ridurre al minimo la sua attività mentale e fisica. Qui, come altrove, il comportamento umano è soggetto alla legge del minimo sforzo, secondo cui l'uomo consuma energia solo nei limiti necessari a raggiungere i fini che si è proposto. (ELG 197)

L'interpretazione funzionalista del fenomeno coglie, cristallizzandoli, i fattori coinvolti di efficienza e di utilità e gli apporti alla teoria della economia del linguaggio e della lingua.

Ma gli stessi postulati paiono applicarsi all'organizzazione della pratica grafica, lì dove essa sia da considerarsi come una rappresentazione visiva della catena parlata, ciò che è il presupposto della scrittura istantanea nel web.

La storia degli alfabeti ci illustra come all'origine essi nascano con la finalità di riprodurre la pronuncia delle lingue, ma, dopo la fissazione della norma scritta, solitamente con l'introduzione della stampa, lingue e scritture separano le loro esistenze, la lingua evolvendo senza considerazione della grafia che la rappresenta, la scrittura fissandosi in forme sclerotizzate che valgono di per sé, praticamente senza più appiglio al riflesso fonetico.

La nuova scrittura che è qui sotto indagine, proprio in quanto non vincolata alla norma, ha il vantaggio della flessibilità e garantisce l'adesione alla dimensione orale, ciò che fa valere per la comunicazione in rete le nozioni di economia, funzionalità, informatività che sono proprie della lingua parlata.

Abbiamo già osservato la dominanza del principio fonetico nella costituzione della forma grafica breve, secondo l'immaginata ideale corrispondenza uno a uno fra gli elementi della scrittura e della lingua:

ke fai? = che fai?

L'uso di lettere/numeri/simboli per sostituire sequenze foniche e intere parole comporta rispetto della biunivocità forma/funzione e solitamente risparmio di movimenti (numero di *click* sui tasti) con positivo effetto di immediatezza che ben si accorda con la limitatezza dello spazio concesso (il piccolo schermo del cellulare, il rigo visibile all'emittente e al destinatario del messaggio nella chat).

L'adozione di un simbolo (aritmetico, numerale) comporta un livello di sinteticità dell'espressione più elevato rispetto al principio alfabetico:

inglese:

2b or not 2b

italiano:

c 6? = ci sei?

+ o - = più o meno

francese:

Cb1 - c'est bien

GspR b1 - j'espère bien

tedesco:

8ung - achtung (attention)

Gn8 - gute nacht

spagnolo:

T2 - todos

Ma nelle tecniche di abbreviazione per la scrittura istantanea l'adeguatezza del sistema non è decretata dalla simmetria delle forme singolarmente considerate, bensì dall'equilibrio funzionale del sistema. Ciò fa ravvisare il superamento della rigida prospettiva fonetica in favore di soluzioni che tengono in considerazione la duplice componente del codice e del messaggio implicata nei fatti di lingua:

1. dimensione della *langue* – saussurianamente intesa, come sistema di elementi interconnessi, a tutti i livelli, fonologico, morfologico, sintattico.

2. dimensione della *parole* - processi di selezione e combinazione, meccanismi di organizzazione ordinata delle unità, ricerca del massimo valore funzionale degli elementi relativamente allo specifico atto linguistico individuale o atto comunicativo coinvolto.

Il "costo" di una parola non corrisponde al numero dei fonemi, o, nel nostro caso, al numero delle lettere che la compongono, bensì piuttosto alla sua capacità informativa all'interno della catena sintattica e alla capacità informativa dei suoi componenti.

Più un'unità è frequente, meno essa è informativa. (ELG 206)

C'è un rapporto costante e inverso fra la frequenza di un'unità e l'informazione che essa dà, cioè in un certo senso la sua efficacia. (ELG 206)

Ovvero:

Nel corpo della parola gli elementi fono-grafici manifestano peso informativo in maniera inversamente proporzionale alla loro frequenza.

Nella pratica di sopprimere elementi della parola o della frase si opta volentieri per l'eliminazione delle vocali, più frequenti e meno informative, nel corpo fonetico dei lessemi:

txt, tnx, thx, ... = <thanks>

dmn, gg, vd, nn, sn, grz... = <domani, oggi, vado/vedo, non, sono, grazie>
gg nn vd = <oggi non vado>

Parimenti è abituale (nelle lingue a componente dominante flessiva) la cancellazione dei costituenti morfologici desinenziali della parola, con la parallela conservazione del lessema, che contiene maggiore informazione sostanziale:

and = <andare/andiamo/andate>

Nel processo di accorciamento, la tendenza alla proliferazione di forme omografiche, con plurime possibili letture (*and* = <andare?/andiamo?/andate?>) trova efficace momento di compensazione nella dimensione sintagmatica.

A livello morfofattico, infatti, troncamenti ed eliminazioni di desinenze possono attuarsi grazie al contestuale recupero della funzione grammaticale attraverso la collocazione nella frase. (Che è poi l'interpretazione funzionalista della storica evoluzione tipologica delle lingue dalle fasi flessive a quelle analitiche, attestata storicamente nel gruppo indoeuropeo.)

È dotato di informazione tutto quello che ha come effetto la riduzione dell'incertezza. (ELG 203)

Che significa:

La creazione della frase procede progressivamente alla "compitazione" della stringa frasale, con lo scopo di eliminare le possibili ambiguità. L'opzione per la forma abbreviata rispetto a quella estesa o la scelta del tipo di abbreviazione fra quelle possibili è determinata dall'intera sequenza delle parole coinvolte o dalle parole immediatamente contigue alla forma segnalata.

La scrittura breve ammette la coesistenza di omografi, in quanto l'ambiguità può essere sciolta dal contesto.

Se una forma grafica può prestarsi a molteplici letture, sarà la catena sintattica – saussurariamente le relazioni paradigmatiche e sintagmatiche - a farci decidere per la corretta interpretazione, ad esempio:

qll è mia sorella = <quella è mia sorella>
qll è mio fratello = <quello è mio fratello>

cs 6 vnt = <così sei venuto>
cs 6 vnt a fare = <cosa sei venuto a fare>

Anche la dimensione co-testuale linguistica scioglie e disambigua il senso:

it. *nn* = <non>
ingl. *nn* = <nice night>

L'informazione non è dunque un attributo della significazione, poiché vi partecipano unità non significanti. (ELG 204)

Nella scrittura elementi singoli, presi isolatamente, non avranno significato, ma saranno altamente informativi ed economici con l'aggiunta di forme brevi che funzioneranno da disambiguatori linguistici:

c = <c/ci/ce>
c 6? = <ci sei?>

k = <chi> [ki] / <che> [ke]
k 6? = <chi sei?>
k fai? = <che fai?>
k ft? = <che fate?>

La simultaneità e la compresenza degli omografi determinano la differenziazione delle soluzioni, così che la decodifica procede attraverso comparazione e secondo la tecnica della progressiva esclusione:

k dice ke nick ama sel e ki dice ke nick ama miley

il mio amico nm ha dtt sl k è carina, ma ki è qst ragazza

No,m spiace,nm t so aiutare...Cmq buongiorno raaaa!!! Ieri m sn addormentata

Ad ogni stadio dell'evoluzione si realizza un equilibrio fra i bisogni della comunicazione, che esigono unità più numerose e più specifiche, ciascuna delle quali appaia meno frequentemente nell'enunciato, e l'inerzia dell'uomo, la quale porta a usare un numero ristretto di unità di valore più generale e di impiego più frequente. (ELG 198)

L'accorciamento delle forme lessicali di cui aumenta la frequenza è un fenomeno ben attestato. (ELG 209)

Ovvero:

Il principio della frequenza delle occorrenze nel particolare contesto induce con maggiore facilità la scelta della forma abbreviata rispetto a quella estesa.

Il lessico dell'sms e della chat non è lo stesso di quello della lingua comune, per quantità e per qualità delle parole. La selezione avviene, come sempre, per motivazioni legate all'uso della lingua e ai contesti, secondo le dinamiche proprie della variabilità diafasica.

Il vocabolario dei social network copre preferibilmente le aree semantiche relative alla vita quotidiana, per lo più relative all'attività scolastica e di gruppo nell'età dell'adolescenza.

La facilità dell'occorrenza nel contesto conversazionale, cioè a dire il successo, in termini sociolinguistici il prestigio del referente, spinge in favore dell'accorciamento delle forme:

raga = ragazzi
comple = compleanno

L'impatto sociale della pratica abbreviativa si osserva per altro nella confluenza di alcune forme rese popolari nel margine della varietà speciale giovanile/familiare/informale nel lessico comune:

metro = metropolitana
auto = autobus
tele = televisione

Quando la frequenza di un'unità cresce la sua forma tende a ridursi. (ELG 209)

La frequenza nel parlato dei connettivi logici, grammaticali e temporali, favorisce l'accorciamento pressoché sistematico delle relative parole ed espressioni, rispetto al maggiore rigore normativo riservato alla rappresentazione grafica degli elementi lessicali:

gg = <oggi>
dmm = <domani>
cmq = <comunque>
qnd = <quando>
qnt = <quanto/-a/-i/-e>

L'unità di informazione è definita come la quantità fornita da un'unità di un sistema che comporti due unità equiprobabili. (ELG 204)

[...] è del tutto eccezionale che le diverse unità di un sistema siano, in una situazione data, o a un punto determinato dell'enunciato, tutte ugualmente probabili. (ELG 205)

Ciò che accade in maniera pressoché estemporanea nella conversazione spontanea ritorna dunque come principio nella costituzione di sistemi predittivi nel trattamento automatico del linguaggio naturale.

Il sistema del dizionario intuitivo o T9 si è progressivamente perfezionato acquisendo procedimenti di completamento automatico delle parole a partire dalla digitazione delle iniziali: l'aggiunta di ogni lettera modifica la proposta degli esiti; l'elenco delle possibilità segue un ordine non fisso, bensì che costantemente si auto-organizza a partire dalla frequenza d'uso nel registro dell'utente.

Analogo procedimento si osserva nell'organizzazione dei motori di ricerca.

Le tecniche della buona ricerca prevedono che sia più opportuno inserire gli elementi lessicali o i morfemi lessicali delle forme cercate rispetto a quelli grammaticali, poiché maggiormente informativi. La digitazione della forma breve velocizza la ricerca e la rende più efficace.

È abbastanza recente la funzionalità di ricerca nota come Google Instant, la quale ci risparmia di scrivere per intero una parola o una sequenza di parole proponendoci, durante la digitazione delle lettere invece che all'invio, le possibili corrispondenze ricercate. Gli esiti proposti sono le risultanze delle forme più frequenti o "popolari".

Come anche nel T9 la ricerca si perfeziona gradualmente man mano che si procede con la scrittura della parola o della sequenza. L'esito più popolare può essere garantito dalla scrittura più breve:

fb = <Facebook>

La massima popolarità dimostra come sia sufficiente digitare la lettera iniziale per avere l'esito voluto in cima alla lista dei risultati

f = <Facebook>

Nel processo comunicativo intervengono elementi diversi dal bisogno di equilibrio informativo. (ELG 210)

Si potrebbe obiettare che l'attività umana in generale e quella linguistica in particolare possono essere fini a se stesse, un gioco: chiacchierare è spesso un'attività gratuita che non mira realmente alla comunicazione, ma piuttosto a una sorta di comunione, ciò che è ben diverso. Ma questo non vuol dire che l'evoluzione linguistica non sia governata dalla legge del minimo sforzo; un gioco non soddisfa un giocatore che nei limiti in cui ne rispetti le regole, e per il linguaggio le regole sono quelle dettate dagli usi comunicativi dello strumento linguistico (ELG 198)

Il dubbio di poter non essere correttamente interpretati e il rischio del fraintendimento operano nella direzione centripeta della tensione per il rispetto della norma, che può manifestarsi nella opzione per la forma estesa:

qnd pensi di vnr = <quando/quindi pensi di venire>

qnd pensi di vnr prx sett!! = <quindi pensi di venire la prossima settimana>

quindi pensi di vnr = <quindi pensi di venire>

Scrivere sms è primariamente attività ludica e comporta alcune volte la combinazione del principio economico con l'attenzione all'aspetto estetico, ovvero all'effetto figurativo, questo ravvisabile in certi casi di estensione analogica degli impieghi.

L'impiego della <x> come doppia <ss> si riscontra, ad esempio, nella forma *adexo*, attestata anche con raddoppiamento della *x* (*adexxo*), con evidente annullando dell'effetto del risparmio (*xx* = <ss>). La resa funzionale va in tal caso piuttosto riconosciuta nel richiamo alla potenza evocativa di un carattere grafico non originariamente italiano. Di qui l'uso sperequativo della <x>, raddoppiata, triplicata, senza motivazione razionale:

belliximo = *bellissimo*

bnxximo = *benissimo*

complex = *compleanno*

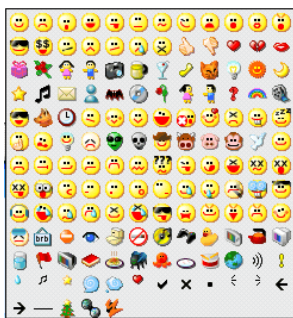
Resta certamente nella dimensione ludica l'espansione oltre misura delle forme brevi, che attesta l'ineludibile presenza del fenomeno della ridondanza nella lingua, evidentemente anche nelle sue espressioni veloci:

- tvb* = <ti voglio bene>
- tvtb* = <ti voglio tanto bene>
- tvttb* = <ti voglio tantissimo bene>
- tvtttttttb* = <ti voglio tantissimissimo bene>
- tvtttttttxs* = <ti voglio tantissimissimo bene per sempre>
- tb1kdbxseo4e* = <ti voglio un casino di bene per sempre e oltre for ever>

Non potremmo concludere questo elenco di forme senza citare l'impiego dei segni di punteggiatura o paragrafematici, l'uso delle maiuscole e altri fenomeni di sovrabbondanza grafica, la moltiplicazione dei simboli come gli *smileys* e gli *emoticons*, le icone come il cuore in tutte le forme, il "pollice verso" introdotto da Zuckerberg:

auguri di vero cuore tvbbbbbbbbbb

LAIII DETTO CIAOOOOOOOOOOOOOOOOOOOOOOOOOOOO TVBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB
VITAA



"J'aime", "I like", "Mi piace", "Me gusta", "Gefällt mir"

Dirette espressioni di tonalità emotive e riflessi affettivi, elementi grafici dalla più efficace capacità di "sentimento", e dalla eccezionale valenza extranazionale ed universale, nel contesto pratico conversazionale tali segni non raramente divengono, allo stesso modo delle scritture brevi alfabetiche, eccedenze grafiche, elementi ripetuti senza risparmio, certamente contravvenenti le proprietà o le potenzialità informative, in una parola forme vive dell'agire linguistico.

Tale versatilità degli impieghi sulla base dei contesti, degli umori, delle personalità, decreta il successo planetario delle scritture brevi, la loro eccezionale spendibilità, l'adattabilità alle condizioni comunicative più variamente intese, la produttività per il singolo e per la comunità, l'inarrestata supremazia nel villaggio globale:



2. Dalla teoria dell'informazione allo scrivere brevemente

Il processo di risparmio nella scrittura delle parole e delle frasi descritto precedentemente può essere tecnicamente riesposto attraverso la teoria dell'informazione di C. E. Shannon ("A Mathematical Theory of Communication", Bell System Technical Journal, 27, July, October, 1948, 379-423, 623-656).

Tale teoria vuole modellare matematicamente lo scambio di informazioni tra un produttore ed un destinatario allo scopo di comprendere come il messaggio possa passare attraverso un canale ed arrivare completamente alla sua destinazione e come un messaggio possa essere codificato efficacemente per usare al meglio lo stesso canale di comunicazione.

Lunghezza del messaggio ed entropia della sorgente

Un notevole risultato della teoria dell'informazione, già presente nella sua introduzione (Shannon 1948), è la stima di quale deve essere la lunghezza media di un messaggio in bit se i messaggi che vengono mandati su di un canale appartengono ad un insieme e hanno probabilità diverse di essere mandati. Questa stima prende il nome di *entropia delle sorgenti dei messaggi* $H(X)$ e ha la formula seguente:

$$H(X) = \sum_{i=1}^n P(X = x_i) \log_2 \frac{1}{P(X = x_i)}$$

dove X è l'insieme dei possibili messaggi e $P(X = x_i)$ è la probabilità che il messaggio x_i si presenti.

La formula dell'entropia delle sorgenti di messaggi è molto interessante perché racconta di come un insieme di messaggi possa essere codificato.

Ad esempio, abbiamo 8 messaggi diversi:

$$X_8 = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$$

Se questi sono tutti equiprobabili, ovvero hanno tutti la stessa probabilità di essere pronunciati, occorrono mediamente 3 bit per trasmetterli. Gli 8 messaggi si devono chiamare 000, 001, 010, 011, 100, 101, 110, 111 e, per capire di quale messaggio si tratta, occorre sempre arrivare fino alla fine della ricezione.

L'entropia dell'insieme X_8 di messaggi equiprobabili può essere calcolata tenuto conto che la probabilità di ogni singolo messaggio è:

$$P(X = x_i) = \frac{1}{8}$$

e, quindi, l'entropia $H(X_8)$ è:

$$H(X_8) = \sum_{i=1}^8 \frac{1}{8} \log_2 8 = \log_2 8 = 3$$

Un caso più interessante è invece quando questi messaggi non sono tutti equiprobabili ma hanno probabilità diverse.

Si prende l'esempio precedente e si associano probabilità come quelle contenute nella tabella sottostante:

messag	x	x	x	x	x	x	x	x
gio	1	2	3	4	5	6	7	8
$P(x_i)$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Ciò significa che il primo messaggio (x_1) ha una altissima probabilità di essere trasmesso, il secondo messaggio (x_2) ha una probabilità minore, ma comunque sostenuta, di essere trasmesso e tutti gli altri messaggi hanno una equi-probabilità di essere trasmessi.

Si potrebbero codificare i messaggi nel seguente modo:

messag gio	x 1	x 2	x_3	x_4	x_5	x_6	x_7	x_8
Codifica	0	1	110	110	111	111	101	101
		00	0	1	0	1	0	1

In questo modo, una sequenza qualsiasi sarebbe sempre decodificabile in quanto se si osserva uno 0, il messaggio è costituito da un solo carattere se si osserva un 1, bisogna guardare due caratteri dopo. Se tutti e due sono 0 allora il messaggio è terminato altrimenti bisogna guardare un altro caratteri per avere il messaggio completo.

Ad esempio, la sequenza:

11101010001000100

corrisponde univocamente alla sequenza:

(1110)(1010)(0)(0)(100)(0)(100) = $x_5 x_7 x_1 x_1 x_1 x_2 x_1 x_2$

La lunghezza media del messaggio è dunque:

$$H(X_8) = \frac{1}{2} \log 2 + \frac{1}{8} \log 8 + \frac{6}{16} \log 16 = \frac{1}{2} + \frac{1}{8} \cdot 3 + \frac{6}{16} \cdot 4 = \frac{1}{2} + \frac{3}{8} + \frac{3}{2} = 2,375$$

Si può dunque osservare che per tale teoria il messaggio più probabile è quello che può essere codificato con una sequenza più corta.

Le parole accorciate e i messaggi brevi

I risultati della teoria dell'informazione sono immediatamente applicabili alla questione delle scritte brevi in quanto modellano in maniera interessante il fenomeno. In questo caso, ci si potrebbe chiedere quante lettere mediamente ed al minimo occorrono per scrivere una determinata parola che appartiene ad un lessico con una certa frequenza.

Abbiamo un lessico L di parole a cui è associata la frequenza in un corpus e , dunque, è stimabile la probabilità di apparizione di ciascuna parola e utilizziamo un alfabeto di 26 lettere.

È possibile scrivere la formula dell'entropia lessicale $HL(L)$ come:

$$HL(L) = \sum_{i=1}^n P(L = w_i) \log_{26} \frac{1}{P(L = w_i)}$$

dove w_i è una delle possibili parole e $P(L = w_i)$ è la probabilità che tale parola si manifesti.

Date le osservazioni fatte nel paragrafo precedente, appare naturale osservare che parole come:

$gg = \langle \text{oggi} \rangle$

$dmm = \langle \text{domani} \rangle$

$cmq = \langle \text{comunque} \rangle$

siano contratte in parole corte essendo esse molto frequenti.

Un interessante studio sarebbe valutare lo scostamento del HL(L) che rappresenta la lunghezza minima dalla lunghezza effettiva in un corpus di scritture brevi.

Francesca Chiusaroli e Fabio Massimo Zanzotto

f.chiusaroli@unimc.it
francesca.chiusaroli@gmail.com

fabio.massimo.zanzotto@uniroma2.it