



## **PSICOLINGUISTICA COMPUTAZIONALE**

RELAZIONE PROVA IN ITINERE

### **Una o più parole? Elaborazione, rappresentazione e interpretazione dei composti**

Studentessa:

Lucrezia Labardi (600163)

Professori:

Marcello Ferro

Claudia Marzi

Vito Pirrelli

ANNO ACCADEMICO 2022-23

# **Sommario**

<b>1. Introduzione</b>	<b>3</b>
<b>2. Caratteristiche dei composti</b>	<b>3</b>
<b>3. Processo di elaborazione dei composti</b>	<b>5</b>
3.1 I primi modelli	5
3.2 Un nuovo schema di composizione del significato	6
3.3 Il ruolo della trasparenza lessicale	8
3.4 Interpretazione della relazione tra costituenti	9
<b>4. Conclusioni</b>	<b>12</b>
<b>Bibliografia</b>	<b>13</b>

## 1. Introduzione

La modalità con cui le parole sono elaborate e rappresentate nel cervello umano è da sempre una questione fortemente divisiva. Numerosi linguisti e psicolinguisti si sono succeduti nel tentare di dare una spiegazione al fenomeno e solo recentemente sono emerse delle nuove evidenze scientifiche.

I composti sono una particolare tipologia di parole complesse, alla cui formazione concorrono due o più parole aventi già un senso compiuto. La natura duale che le caratterizza facilita la creazione di nuove parole e spinge ad ampliare il lessico sia personale che di intere comunità linguistiche (Libben et al., 2020). La loro importanza appare quindi indiscutibile.

Nella presente relazione vengono descritte alcune caratteristiche dei composti (sezione 2) e successivamente mostrate delle evidenze emerse tramite diversi studi ed esperimenti (sezione 3). Si effettua un confronto tra i modelli “classici” con cui è stata teorizzata la rappresentazione del nostro lessico mentale (sezione 3.1) ed uno schema di composizione del significato realizzato più recentemente (sezione 3.2). Si definisce il ruolo che ha la trasparenza lessicale nei modi e nei tempi di elaborazione dei composti (sezione 3.3) ed infine si tenta di definire l’importanza delle relazioni tra i singoli costituenti di tali parole (sezione 3.4).

Gli studi che vengono qui presi in considerazione sono solo una piccola parte della letteratura in materia, ma risultano esaustivi per la formazione di un quadro generale dei possibili sviluppi futuri.

## 2. Caratteristiche dei composti

Le parole composte sono strutture fondamentali per l'evoluzione della lingua, in quando definiscono la nostra abilità di creare nuove parole a partire da quelle esistenti. Paradossalmente, per quanto risulti facile dare una loro definizione nel dizionario - le parole composte sono parole formate dall’unione di due o più parole o parti di parole<sup>1</sup> -, per gli psicolinguisti ed i morfologi il compito è più difficile. La modalità con cui due parole si fondono per formarne una nuova è infatti tutt’altro che semplice da indagare e la relazione tra i costituenti ed il composto dipende dalle loro caratteristiche formali e semantiche.

Una prima classificazione dei composti può essere effettuata tra esocentrici ed endocentrici, in base alla derivazione delle loro proprietà linguistiche. I composti endocentrici presentano una “testa” che può essere chiaramente identificata tramite la relazione semantica ISA (“è un tipo di”): ad esempio, dato che *pesceccane* è un tipo di *pesce*, *pesce* è la sua testa. La testa determina le informazioni categoriali, i tratti sintattico-semantiche ed il genere del composto.

---

<sup>1</sup> Definizione dell’enciclopedia Treccani online (sito visitato il 29/12/2022):  
[https://www.treccani.it/enciclopedia/parole-composte\\_\(La-grammatica-italiana\)](https://www.treccani.it/enciclopedia/parole-composte_(La-grammatica-italiana))

Al contrario, negli esocentrici la testa non può essere individuata in nessuno dei costituenti, come accade ad esempio in *pellierossa*, che infatti non è un tipo di pelle.

A livello semantico la questione è ancora più interessante: i singoli costituenti del composto, dotati di un preciso significato, lo cedono in favore del significato della parola composta che contribuiscono a formare. Il significato finale del composto spesso non è riconducibile alla sola somma dei significati delle due parole semplici, ma contiene qualcosa in più.

Una seconda classificazione delle parole composte può essere effettuata in trasparenti ed opachi, in base a quanto il significato individuale dei costituenti è presente e riconoscibile all'interno del significato del composto. E' necessario ricordare che tale suddivisione non è mai netta ma viene definita su un gradiente di trasparenza semantica. In Libben et al. (2020) viene dimostrato come la trasparenza semantica influenzi l'elaborazione, la comprensione e la produzione lessicale (vedi sezione 3.3).

Tra una lingua e l'altra la forma dei composti può variare. Si individuano composti *head-initial*, che hanno la testa come primo elemento, o *head-final*, che hanno la testa come elemento finale. Nelle lingue germaniche tutti i composti sono *head-final*, nell'ebreo *head-initial*, mentre nelle lingue romanze è possibile trovare entrambe le tipologie. Nessun parlante nativo ha difficoltà a decifrare la struttura interna dei composti, mentre sono comuni gli errori di interpretazione di parole flesse o derivate, motivo per cui le parole composte sono molto produttive.

Generalmente i composti sono formati da due elementi lessicali concatenati l'uno all'altro, ma tale struttura può subire delle variazioni: è possibile che contengano delle preposizioni che esplicitano la relazione tra i costituenti, oppure degli infissi (spesso derivati da vocali tematiche greche o latine) che ne rendono più lenta l'elaborazione (Dressler et al., 2001)<sup>2</sup>. E' inoltre possibile riscontrare diverse tipologie di separazione tra i costituenti, come uno spazio o un trattino.

La fruttuosità di uno schema di creazione dei composti in termini di *type* e *token frequency* influenza la sua produttività. Uno schema produttivo può generare famiglie di composti molto ampie, ad esempio *pesce palla*, *pesce spada*, *pesce cane*, ma anche *pescetariano*, *pesce d'aprile*, ecc. In tutti questi esempi è presente la parola *pesce* come testa del composto e come prima componente (cioè si trova sempre nella stessa posizione). Data l'ampiezza della famiglia di composti, sarà facile per i parlanti coniare nuovi termini utilizzando la stessa regola *pesce* + parola (Libben et al., 2020). Dressler (2007)<sup>3</sup> evidenzia che gli schemi più produttivi vengono applicati anche a nuove parole di origine straniera la cui forma viene adattata alle condizioni già valide per le parole esistenti. Egli riporta anche che gli schemi meno produttivi hanno più

---

<sup>2</sup> Citato in Libben et al. (2020)

<sup>3</sup> Citato in Libben et al. (2020)

probabilità di essere legati a restrizioni grammaticali e pragmatiche, quindi vengono applicati soltanto a parole native.

Inoltre, la ricchezza dei composti in una lingua è molto legata al numero di schemi produttivi che possiede (Libben et al., 2020). Tale caratteristica si riflette sulla ricchezza dell'uso dei composti nei bambini e nell'età in cui essi diventano produttivi a livello lessicale. In Dressler et al. (2017)<sup>4</sup> viene riportato, a tal proposito, come questi fenomeni accadano prima nei bambini che stanno imparando il tedesco o il danese, rispetto a quelli che imparano il francese o il russo, dove il fenomeno della composizione è meno produttivo. Mattiello e Dressler (2018)<sup>5</sup> notano inoltre che nei bambini, emergono in un primo momento, solamente composti semanticamente trasparenti e privi di infissi.

### 3. Processo di elaborazione dei composti

A partire dagli anni Settanta si sono susseguite diverse proposte di modelli per descrivere l'elaborazione e la rappresentazione mentale delle forme lessicali complesse. Tra esse, i composti si prestano particolarmente agli studi (psico)-linguistici in quanto si tratta di parole generalmente a bassa frequenza costituite da termini appartenenti a classi aperte, che per questo richiedono un costo di scomposizione maggiore.

#### 3.1 I primi modelli

Ad un primo estremo delle diverse teorie proposte si ha il modello di *full listing*, teoria per cui le parole vengono memorizzate come un tutto ed il ruolo della loro struttura morfologica interna è marginale nell'elaborazione.

All'estremo opposto si trova il *full parsing*, di cui fanno parte diversi modelli di decomposizione che teorizzano che le parole siano memorizzate ed elaborate suddividendole in costituenti morfologici.

Negli anni Novanta è poi emerso il cosiddetto *race model*, modello per cui nel cervello sono presenti entrambe le rappresentazioni e viene preferita di volta in volta quella che impiega minor tempo e risorse.

Come viene riportato in Ji et al. (2011), era comune agli studiosi ritenere che la scomposizione delle parole richiedesse più tempo ed un maggiore costo computazionale perché ogni singola parola doveva essere suddivisa in sotto-costituenti e poi ricomposta per trovarne il significato, e che l'accesso diretto fosse sempre conveniente. Le evidenze erano però discordi: alcune attribuivano alle parole complesse una maggiore difficoltà di elaborazione rispetto a quelle

---

<sup>4</sup> Citato in Libben et al. (2020)

<sup>5</sup> Citato in Libben et al. (2020)

monomorfemiche (Bertram et al., 1999)<sup>6</sup>, altre ne mostravano i vantaggi (Burani & Thornton, 2003)<sup>7</sup>. E' pertanto possibile sostenere che "le parole complesse non siano necessariamente più complicate da elaborare rispetto alle monomorfemiche" (Ji et al., 2011).

Il principio comune, che tutti i modelli devono seguire, è quello del bilanciamento tra la memorizzazione del lessico ed il calcolo morfologico. E' possibile intuire, infatti, che memorizzare per intero tutte le parole possa essere dispendioso e ridondante, come dover computare ogni volta ciascuna parte dei diversi simboli a cui si è esposti può richiedere un tempo eccessivo. Il bilanciamento di questi due fattori rappresenta la strada più breve e semplice che il parlante potrà seguire per elaborare e rappresentare il proprio lessico mentale.

### 3.2 Un nuovo schema di composizione del significato

In Ji et al. (2011) viene riportato uno studio del ruolo delle diverse componenti (lessicale e semantica) nelle tempistiche di elaborazione dei composti. Attraverso sei esperimenti di decisione lessicale attuati confrontando parole composte e parole monomorfemiche aventi la stessa frequenza, è emerso che i composti vengono processati meglio e più velocemente rispetto alle parole monomorfemiche, soprattutto se hanno una trasparenza semantica maggiore.

Il primo esperimento consisteva in un compito di decisione lessicale durante il quale i partecipanti dovevano stabilire se la sequenza di caratteri mostrata fosse effettivamente una parola. Era comune in letteratura ritenere che le parole complesse ad alta frequenza fossero più probabilmente riconosciute tramite l'accesso diretto, mentre quelle a bassa frequenza attraverso la scomposizione (Alegre & Gordon, 1999)<sup>8</sup>, pertanto lo studio era volto ad indagare anche se il vantaggio delle parole composte (probabilmente dovuto all'accesso ai costituenti) fosse limitato ai composti a bassa frequenza.

Tale esperimento ha smentito sia la teoria di *full listing*, dato che il tempo di elaborazione dei composti si è rivelato minore di quello delle parole monomorfemiche<sup>9</sup>, sia alcune teorie di scomposizione automatica, che attribuivano un vantaggio alle parole monomorfemiche rispetto ai composti trasparenti (Taft, 1994)<sup>10</sup>. E' emerso pertanto che il costo della scomposizione viene compensato dal vantaggio della disponibilità dei costituenti, che facilita la comprensione della parola intera.

---

<sup>6</sup> Citato in Ji et al. (2011)

<sup>7</sup> Citato in Ji et al. (2011)

<sup>8</sup> Citato in Ji et al. (2011)

<sup>9</sup> Secondo la teoria di *full listing* tutte le tipologie di parole vengono accedute per intero. Una differenza nel tempo di *processing* tra parole monomorfemiche e composti suggerisce invece che ci siano delle differenze nella loro elaborazione.

<sup>10</sup> Citato in Ji et al. (2011)

In Ji et al. (2011) vengono inoltre riportate due teorie che spiegano tale vantaggio: secondo la teoria della *dual-route* il vantaggio era dovuto ad una facilitazione statistica per cui la via della scomposizione impiega meno tempo dell'accesso diretto (Schreuder & Baayen, 1995)<sup>11</sup>, mentre secondo la teoria di Libben (1998)<sup>12</sup> a livello lessicale si verifica una determinante pre-attivazione della rappresentazione dei costituenti. Entrambe queste teorie però prevedevano che il vantaggio delle parole composte a bassa frequenza dovesse essere maggiore, mentre dall'esperimento è emerso chiaramente come la frequenza di una parola composta non sia legata al suo grado di trasparenza.

Il terzo esperimento presentato nello studio di Ji et al. (2011) era volto a indagare se il vantaggio dei composti si verificasse solo a livello lessicale o anche a livello semantico. Mettendo a confronto delle triplette formate da un composto opaco, uno trasparente ed una parola monomorfemica, tutti con la stessa frequenza, è stato ripetuto il compito di decisione lessicale.

E' emerso che tutte le tipologie di parole composte (trasparenti o opache) hanno un vantaggio sulle parole monomorfemiche in termini di tempo. Per quanto riguarda invece l'accuratezza, soltanto per i composti trasparenti si sono riscontrati dati migliori rispetto alle parole monomorfemiche. Inoltre, la presenza di un primo costituente ad alta frequenza si è rivelato vantaggioso sia per i composti opachi che per i trasparenti. Da questo si può dedurre che alla facilitazione dei composti contribuiscono sia il livello lessicale che quello semantico, con una maggiore rilevanza del primo. Se l'attivazione a livello lessicale e a livello semantico avessero la stessa importanza, infatti, la presenza del primo costituente ad alta frequenza sarebbe un beneficio soltanto per le parole composte trasparenti, mentre costituirebbe un'interferenza per quelle opache.

Negli esperimenti numero 4, 5 e 6 di Ji et al. (2011) sono state inserite delle manipolazioni nelle parole scelte per la decisione lessicale, con l'obiettivo di facilitare per i partecipanti la distinzione dei costituenti, ad esempio inserendo uno spazio tra loro o riportandoli con colori diversi. Si è osservato che soltanto i composti trasparenti ne hanno tratto beneficio. Inoltre, la presenza nei composti del primo costituente ad alta frequenza è risultato positivo per i trasparenti e negativo per gli opachi.

Possiamo dedurre da questo come l'accesso alle informazioni lessicali, semantiche e concettuali dei costituenti avvenga sempre, in ogni caso, perché da esso il parlante può costruirsi una prima idea del significato della parola. Rendere più evidente la rappresentazione lessicale migliora la connessione tra i costituenti e la parola composta intera, aumentando il costo di elaborazione per i composti opachi, nei quali tale connessione non è presente.

---

<sup>11</sup> Citato in Ji et al. (2011)

<sup>12</sup> Citato in Ji et al. (2011)

Il nuovo schema che viene suggerito in Ji et al. (2011), per arrivare a comprendere il significato finale di una parola composta, si basa sulla convinzione che il cervello umano ricerchi sempre l'esistenza di un significato "lessicale" a partire da quello dei singoli costituenti, il quale entra in competizione con uno o più significati "costruiti" tramite analisi successive. La competizione avviene a livello concettuale e la valutazione viene fatta anche in base al contesto in cui la parola appare (se presente). Nel caso in cui il composto sia trasparente la competizione è minore rispetto ad un composto opaco, pertanto anche il tempo di elaborazione diminuisce.

Le evidenze emerse portano a ritenere che l'accesso a tipologie di parole diverse avvenga nello stesso modo ma con tempistiche differenti: rispetto alle parole monomorfemiche, le parole che presentano dei sotto-costituenti ben definiti traggono vantaggio da tale caratteristica dipendentemente dal proprio grado di trasparenza lessicale.

### **3.3 Il ruolo della trasparenza lessicale**

Come riportato in Libben et al. (2020), la trasparenza lessicale è un fattore che influenza fortemente l'elaborazione dei composti, sia in compiti di comprensione che di produzione lessicale. In generale, i composti con testa trasparente richiedono minor tempo di elaborazione nel compito di decisione lessicale rispetto ai composti opachi. E' inoltre noto come la frequenza relativa del primo costituente sia determinante nel favorire una maggiore o minore rapidità di elaborazione.

Diversi esperimenti di *priming* hanno fatto emergere come tutti i tipi di composti, sia trasparenti che opachi, traggano beneficio dall'esposizione ad un prime corrispondente ad uno dei due costituenti, in quanto esso preattiva in parte la parola target. Come è facile immaginare, soltanto i composti trasparenti beneficiano invece dell'esposizione ad un prime che sia una parola semanticamente correlata ad uno dei due costituenti. Questo si può attribuire all'interferenza che genera l'attivazione di un certo significato del costituente quando non è quello che esso ha nella parola target.

Dagli esperimenti descritti in Ji et al. (2011) sappiamo inoltre che le manipolazioni che facilitano l'individuazione dei costituenti danno beneficio ai composti trasparenti e danneggiano la comprensione di quelli opachi. Anche questo può essere motivato con la pre-attivazione dei singoli costituenti, che nel caso dei composti opachi va ad interferire, almeno semanticamente, con il significato della parola intera, e richiede più tempo di elaborazione per "scartare" il significato inopportuno.

In conclusione, facendo nuovamente riferimento a Ji et al. (2011), è necessario ricordare che l'opacità di un composto può essere totale o parziale. Se un composto ha la testa trasparente, il suo significato farà parte della stessa area semantica della testa, quindi sarà vicino al significato "lessicale". Al contrario, se il composto è del tutto opaco, nessuno dei costituenti sarà legato al significato della parola intera, che quindi



dovrà essere “costruito”. La competizione tra significato lessicale e costruito sarà quindi più complessa se i due appartengono ad aree semantiche diverse, con due possibili conseguenze: la loro lontananza semantica rende più facile scartare il significato inopportuno quando esso non si adatta al contesto, ma allo stesso tempo rende i due significati difficilmente confrontabili in caso di incertezza.

### **3.4 Interpretazione della relazione tra costituenti**

Agli anni Settanta risalgono i primi tentativi di individuare una relazione tra i costituenti delle parole composte, a partire dalla teoria di Gleitman e Gleitman (1970)<sup>13</sup> per la quale i composti sarebbero derivati da proposizioni relative alla cui struttura è possibile risalire. Il maggior risultato in questo campo è stato ottenuto da Levi (1978)<sup>14</sup>, la quale ha individuato sedici categorie relazionali con le quali è possibile classificare i composti della lingua inglese, che riteneva come inizialmente costituiti anche da un verbo che poi è andato perduto<sup>15</sup>.

Nel testo di Schmidtke et al. (2016) si fa riferimento ad uno studio di Libben et al. (1999), che ha indagato attraverso diversi esperimenti se le relazioni tra i costituenti di un composto contribuiscono al suo riconoscimento. La prima prova era volta a determinare se la maggioranza dei partecipanti tendeva ad assegnare ad un composto la prima scomposizione possibile oppure no. Ai partecipanti sono stati pertanto somministrati alcuni composti inediti chiedendo di individuarne la divisione in costituenti. I risultati erano equamente divisi tra la prima scomposizione possibile e la seconda, quindi possiamo ritenere che l'individuazione dei costituenti di una parola composta avvenga in seguito ad una valutazione della plausibilità di ciascuna opzione. Per tale valutazione appare determinante l'accesso semantico ai costituenti e non solo alla parola intera. Inoltre, altre evidenze hanno suggerito che più appariva facile costruire il significato del composto basandosi sui costituenti, e più era difficile ritenere che quel composto non fosse una parola esistente. A tal proposito, un altro riferimento viene fatto all'esperimento di Coolen et al. (1991), durante il quale i partecipanti dovevano fornire una parafrasi del significato delle parole composte somministrate. Le loro risposte venivano poi classificate in base alle relazioni proposte da Levi (1978).

E' emerso che i composti considerati ad alta interpretabilità venivano più facilmente parafrasati con le relazioni di Levi rispetto a quelli a bassa interpretabilità. Studi successivi (Gagné & Shoben, 1997)<sup>16</sup> hanno evidenziato come l'interpretazione dei composti sia influenzata dalla disponibilità della loro struttura relazionale, a sua volta

---

<sup>13</sup> Citato in Schmidtke et al. (2016)

<sup>14</sup> Citato in Schmidtke et al. (2016)

<sup>15</sup> Alcuni esempi delle categorie relazionali di Levi sono: MADE OF (“è fatto di”), FOR (“per”), HAS (“ha”, “presenta”), ISA (“è un tipo di”)

<sup>16</sup> Citato in Schmidtke et al. (2016)

determinata dall'uso sia generale che recente. L'uso generale si riferisce alla consapevolezza di quanto una certa relazione sia utilizzata in associazione ad un certo costituente. In italiano spesso queste relazioni emergono tramite l'utilizzo di preposizioni all'interno della parola composta, ad esempio *palla* è spesso associata allo sport in cui viene utilizzata, come in *palla da calcio*, *palla da basket*, *palla da pallavolo*. In inglese la relazione è invece implicita: ad esempio la parola *magazine* è spesso associata alla relazione ABOUT ("riguardo a"): *fashion magazine*, *travel magazine*, *sport magazine*, ecc. La conoscenza dell'uso generale di una relazione dipende quindi dal grado di conoscenza della lingua.

L'uso recente invece può essere sperimentato tramite il cosiddetto *priming* relazionale. In Gagné e Spalding (2004) e Gagné et al. (2009)<sup>17</sup> sono stati effettuati a questo scopo degli esperimenti di decisione lessicale: data una parola composta come target, il tempo di latenza era minore quando essa era preceduta da un prime che presentava la stessa relazione e lo stesso modificatore del target. Al contrario, se il prime presentava lo stesso modificatore ma una diversa relazione, il tempo di latenza aumentava. Questo accadeva soltanto se il modificatore coinvolto aveva la stessa posizione sia nel prime che nel target.

Comprendere la relazione che lega i costituenti di un composto rende già chiara una possibile spiegazione del significato della parola intera. In Gagné & Shoben (1997) e Spalding et al. (2010)<sup>18</sup> vengono presentate tre fasi, non necessariamente consecutive, che loro ritengono fondamentali per interpretare un composto:

- La competizione tra le relazioni associate ad un modificatore, per essere scelte come possibile interpretazione,
- La competizione tra le relazioni associate sia al modificatore che alla testa, usate insieme alle loro informazioni semantiche per verificare una prima interpretazione del composto,
- La valutazione e l'elaborazione della prima interpretazione ottenuta, per arrivare ad un significato più preciso della parola composta

La competizione appare un elemento fondamentale nella selezione del significato delle parole composte, ed è ciò che fa aumentare la difficoltà di elaborazione.

Il grado di competizione tra le relazioni può essere calcolato tramite l'entropia, una misura che quantifica l'informazione attesa nella distribuzione di probabilità delle relazioni semantiche e calcola la media di incertezza nello scegliere una delle relazioni associate con un certo costituente. L'entropia cresce quando ci sono più relazioni semantiche associate ad uno stesso composto e quando le probabilità di tali relazioni sono valori molto vicini tra loro. Essendo una misura dell'incertezza, un suo valore maggiore farà crescere anche il tempo necessario all'elaborazione.

---

<sup>17</sup> Citati in Schmidtke et al. (2016)

<sup>18</sup> Citati in Schmidtke et al. (2016)

L'entropia può essere calcolata sulla distribuzione di tutte le relazioni concettuali associate alla famiglia di un certo modificatore, ma anche sulla distribuzione delle relazioni concettuali ottenute tramite il compito delle possibili relazioni. Tale compito è stato oggetto dello studio di Gagnè e Shoben (1997)<sup>19</sup>, durante il quale sono stati somministrati ai partecipanti alcuni composti formati da due parole, chiedendo loro di scegliere l'interpretazione relazionale più probabile tra le 16 suggerite, ricavate a partire dallo studio di Levi (1978). Con le risposte è stata generata una distribuzione di probabilità delle possibili interpretazioni del composto. Sui dati è stata calcolata la Relational Diversity, ovvero il numero di interpretazioni diverse scelte per il composto, la Relational Diversity Ratio, ottenuta dividendo il numero di relazioni attestate da più di 10 partecipanti per il numero totale di relazioni scelte da almeno 1 partecipante, e l'entropia. Quest'ultima è risultata essere la misura più diretta della competizione tra le relazioni, perché permette di prendere in considerazione non solo il numero delle possibili interpretazioni di un composto o la sua interpretazione più frequente, ma anche il loro bilanciamento di probabilità. Questo fattore rende più evidenti i casi di incertezza: quando competono molte interpretazioni ma una di esse è di gran lunga più probabile si avrà un'incertezza minore rispetto ad un caso in cui il numero di interpretazioni in competizione è minore ma la loro probabilità è simile.

Lo studio di Schmidtke et al. (2016) è stato svolto proprio con lo scopo di esaminare ulteriormente l'entropia come misura della competizione tra relazioni. Utilizzando un dataset di giudizi raccolti tramite altri esperimenti sulle possibili relazioni per un certo numero di composti, hanno richiesto il compito di decisione lessicale somministrando ai partecipanti tali composti in forma concatenata.

L'entropia della distribuzione di probabilità delle relazioni concettuali scelte da più di un partecipante ha rivelato, coerentemente con quanto detto prima, un effetto inibitorio: un maggiore grado di incertezza porta quindi ad un maggior tempo di latenza per la decisione.

E' risultato evidente inoltre che i composti con famiglie morfologiche più ampie vengono processati più velocemente di quelli con famiglie ristrette, e composti più frequenti e con costituenti più frequenti vengono processati più velocemente di quelli a bassa frequenza. Inoltre, valutando la correlazione tra l'entropia e la trasparenza semantica, calcolata tramite Latent Semantic Analysis (LSA)<sup>20</sup>, non è emersa nessuna significatività, per questo è stato stabilito che la trasparenza semantica e l'entropia delle relazioni concettuali non sono fattori che si influenzano nel riconoscimento di un composto.

---

<sup>19</sup> Citato in Schmidtke et al. (2016)

<sup>20</sup> La Latent Semantic Analysis è una misura statistica del grado di distanza tra due parole, basata sui contesti in cui due parole appaiono vicine all'interno di un corpus. Può essere calcolata come indicatore di trasparenza semantica confrontando i costituenti di un composto sia con la parola intera che tra loro.

## 4. Conclusioni

La natura duale delle parole composte - sono parole uniche ma anche il risultato di una combinazione di parole - le rende un ottimo soggetto di studio sui sistemi di elaborazione, rappresentazione ed interpretazione delle parole complesse.

A partire dagli studi di Ji et al. (2011), Schmidtke et al. (2016) e Libben et al. (2020) è risultato evidente che, indipendentemente dal tipo di composto, durante l'elaborazione vengono attivate sia la parola intera che i suoi costituenti, sia a livello lessicale che semantico. Per comprendere il significato del composto in un primo momento il parlante attiva il significato lessicale (basato sulle sole caratteristiche dei costituenti) ed in seguito genera altri possibili significati costruiti. Tutti i significati proposti entrano in competizione tra loro a livello concettuale.

La trasparenza lessicale è un fattore fondamentale: dal momento che una maggiore opacità corrisponde ad una minore connessione tra composto e costituenti, se il composto è molto opaco la competizione tra significati richiede un maggiore costo di elaborazione. Facilitare l'individuazione dei costituenti, quindi evidenziare la struttura lessicale, non sempre è d'aiuto: soltanto i composti trasparenti beneficiano della pre-attivazione del significato dei propri costituenti, mentre la stessa operazione negli opachi crea interferenza.

Infine, la comprensione del significato di un composto può essere considerata come la comprensione della relazione tra i costituenti che lo compongono. L'entropia è una misura fondamentale dell'informazione per quanto riguarda il grado di incertezza relativo alla probabilità di una certa relazione. In base a quanto una relazione è associata a ciascun costituente e all'uso che è stato fatto di tale relazione, è possibile calcolare la distribuzione di probabilità relativa a ciascuna relazione per ogni coppia di costituenti. Se un composto ha molte interpretazioni relazionali simili tra loro necessita di un tempo maggiore per essere decifrato.

In conclusione, i composti che vengono elaborati più facilmente sono quelli appartenenti a schemi molto produttivi, morfologicamente semplici e semanticamente trasparenti, con un numero ridotto di possibili interpretazioni relazionali. Essendo presi in considerazione sempre sia i costituenti che la parola per intero, il processo richiede un maggiore costo computazionale dovuto alla decomposizione rispetto alle parole monomorfemiche, ma esso viene compensato in termini di tempo e correttezza dal vantaggio che la disponibilità dei costituenti fornisce. Nel caso dei composti con un alto grado di opacità, la pre-attivazione dei costituenti non dà beneficio e risulta particolarmente dannosa la presenza del primo costituente ad alta frequenza, pertanto il loro tempo di elaborazione risulta maggiore.

## Bibliografia

1. Alegre, M., & Gordon, P. (1999). *Frequency effects and the representational status of regular inflections*. Journal of Memory and Language, 40, 41–61.
2. Bertram, R., Laine, M., & Karvinen, K. (1999). *The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language*. Journal of Psycholinguistic Research, 28, 213–226.
3. Burani, C., & Thornton, A. M. (2003). *The interplay of root, suffix and whole-word frequency in processing derived words*. In H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 157–208). Berlin: Mouton de Gruyter.
4. Coolen, R., Van Jaarsveld, H. J., & Schreuder, R. (1991). *The interpretation of isolated novel nominal compounds*. Memory & Cognition, 19(4), 341–352.
5. Dressler, W.U., Libben, G., Stark, J., Pons, C. & Jarema, G. (2001) *The processing of interfixed German compounds*. In Geert Booij and Jap van Marle (eds.), *Yearbook of Morphology*, 185–220. Dordrecht: Kluwer Academic.
6. Dressler, W.U. (2007). *Productivity in word formation*. In Gonia Jarema & Gary Libben (eds.), *The Mental Lexicon: Core Perspectives*, 159–183. Amsterdam: Elsevier.
7. Dressler, W.U., Ketrez, F.N. & Kilani-Schoch, M (eds.) (2017) *Nominal Compound Acquisition*. Amsterdam: Benjamins.
8. Gagné, C. L., & Shoben, E. J. (1997). *Influence of thematic relations on the comprehension of modifier-noun combinations*. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23(1), 71.
9. Gagné, C. L., & Spalding, T. L. (2004). *Effect of relation availability on the interpretation and access of familiar noun-noun compounds*. Brain and Language, 90(1), 478–486.
10. Gagné, C. L., Spalding, T. L., Figueredo, L., & Mullaly, A. C. (2009). *Does snow man prime plastic snow? The effect of constituent position in using relational information during the interpretation of modifier–noun phrases*. The Mental Lexicon, 4(1), 41–76.
11. Gleitman, L.R. e Gleitman, H. (1970). *Phrase and paraphrase: some innovative uses of language*.
12. Ji, H., Gagné, C.L. e Spalding, T.L. (2011). *Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds*. Journal of Memory & Language. 65. 406–430
13. Levi, J.N. (1978). *The syntax and semantics of complex nominals*. New York. Academic Press.
14. Libben, G., Gagné, C. e Dressler, W.U. (2020). *The representation and processing of compound words*. In V. Pirrelli, I. Plag & W.U. Dressler (eds.) *Word knowledge and word usage: a cross-disciplinary guide to the mental lexicon*, 336-352
15. Libben, G., Derwing, B. L., & Almeida, R. G. D. (1999). *Ambiguous novel compounds and models of morphological parsing*. Brain and Language, 68(1), 378.386.
16. Mattiello, E. & Dressler, W.U. (2018) *The Morphosemantic Transparency/ Opacity of Novel English Analogical Compounds and Compound Families*. Studia Anglica Posnaniensia 53. 67–114. doi: 10.2478/stap-2018-0004.
17. Schmidtke, D., Kuperman, V., Gagné, C.L. e Spalding, T.L. (2016). *Competition between conceptual relations affects compound recognition: The role of entropy*. Psychonomic Bulletin & Review 23 (2). 556–570
18. Schreuder, R., & Baayen, R. H. (1995). *Modeling morphological processing*. Hillsdale, NJ, England: Lawrence Erlbaum.
19. Spalding, T. L., Gagné, C. L., Mullaly, A. C., & Ji, H. (2010). *Relation-based interpretation of noun–noun phrases: A new theoretical approach*. Linguistische Berichte Sonderheft, 17, 283–315.
20. Taft, M. (1994). *Interactive-activation as a framework for understanding morphological processing*. Language and Cognitive Processes, 9, 271–294.