**PRINCIPIO DI COMPOSIZIONALITA’**

Che cosa è il principio di composizionalità? L’ipotesi di base è che il significato di espressioni complesse sia una funzione dei significati dei suoi costituenti (parti più piccole) e dall’applicazione di un certo numero di regole sintattiche usate nella composizione di significati atomici o semplici. A chi è venuto in mente?

Il perchè? Esistono un numero potenzialmente infinito di frasi nel linguaggio naturale che contrasta con la capacità di memorizzazione umana quindi viene su capire da dove viene la capacità di produrre strutture infinite usando risorse limitate. Quanta conoscenza deve essere memorizzata in modo più o meno esplicito nei nostri sistemi? (anche quelli semplici come un parser che può avere un insieme di regole e di espressioni imbagagliate)

Le persone capiscono delle frasi che non abbiamo mai sentito in precedenza. Anche nel momento in cui ascoltiamo il linguaggio combiniamo significati relativi ad elementi atomici della frase.

Il principio di composizionalità è motivato da una esigenza di efficienza. L’ambiguità è uno **strumento economico**, con pochi termini definiamo più significati. Qui cerchiamo di capire il livello minimo di risorse che ci permettono di analizzare e capire il linguaggio. Dal punto di vista dello sviluppo dei sistemi, il principio di composizionalità serve per gestire il divario capacità computazionali e capacità di memorizzazione.

**Principio della produttività:** Da un numero finito di parole e un numero finito di regole di composizione sintattiche generiamo frasi infinite. Siccome il linguaggio è infinito questo è un modo economico per generale frasi. Quello della produttività è un principio che postula l’esistenza di un principio di composizionalità dei significati.

**Obiezione 1 all’argomento della produttività:** Ci sono frasi idiomatiche come “**flat tyre” o “flat bee**r” per cui il significato di una parola non è sempre lo stesso. Se fosse semplice che da significati atomici e da regole uno si spiega il significato della frase evidentemente non riusciremo a dare un senso a queste 2 espressioni. Il nome di fronte ad un aggettivo cambia significato. (problema del buon amico, buon coltello). Se cerchiamo si salvare la composizionalità ammettendo che queste espressioni siano già note andremo contro l’argomento della produttività, quindi non vale la composizionalità.

**Principio della sistematicità:** Un altro argomento che prova l’esistenza della produttività è l’argomento della sistematicità. L’idea è che chi capisce una frase come “Jonh ama mary” tendenzialmente capisce frasi come “Mary ama Jonh”.

**L’obiezione alla sistematicità** è data dal fatto che chi capisce frasi come: “**within an hour**” e “**without a watch**” non capisce il senso di frasi come “**within a watch”** e “**without an hour”**. Quindi il solo fatto di prendere pezzi di frase capite e combinarle danno luogo a frasi che non sono facili da capire. I linguaggi non sono sistematici in questo senso.

**Principio della modularità:** esiste un approccio modulare alla costruzione di frasi del linguaggio naturale. Si parla di **incapsulamento informazionale** abbastanza simile al concetto di incapsulamento nella programmazione ad oggetti. Molto rilevante! Quali sono le caratteristiche di questo principio? I sistemi percettivi tra cui quelli deputati alla elaborazione linguistica non accedono meglio all’insieme delle conoscenze che sono rappresentate all’interno della totalità dell’organismo. I moduli procedono in maniera separata e senza una regia complessiva. Quali implicazioni ha questo principio?

**Teoria del singolo modulo**: Una grammatica che ha un certo potere generativo può essere catturata all’interno di un singolo modulo. Ogni modulo è caratterizzato da 2 elementi:

* un insieme finito di significati lessicali
* un insieme finito di regole sintattiche per manipolare i termini lessicali

L’output del modulo, ovvero il significato di una espressione complessa sarebbe una funzione della conoscenza accessibile al modulo ovvero unicamente lessico e sintassi.

**Teoria dei 2 moduli:** Un modulo presiede alla elaborazione dell’analisi sintattica il cui output diventa l’input di un altro modulo che contiene l’insieme dei significati per gli elementi analizzati e le procedure di combinazione del significato dell’espressione complessa. La composizionalità è il vincolo sul tipo di connessione e comunicazioni tra moduli. La modularità in 2 passi è stata importante perchè ha elaborato **la composizionalità come un principio di elaborazione linguistica**.

Quali sono le caratteristiche di questa forma di elaborazione linguistica?

* Composizione incrementale
* Composizione semplice

**Composizione incrementale:** Cosa vuol dire? Vuol dire che esistono dei passi della computazione, che proseguono linearmente seguendo il flusso di ascolto della frase. Il significato della frase al passo sigma è calcolato sulla base delle sotto espressioni processate fino allo stato sigma e alla struttura sintattica costruita fino a sigma. La composizione incrementale deve essere integrata nella definizione di composizione semplice

**Composizione semplice:** Forma più rigorosa di composizionalità, riesce a spiegare meno fenomeni. Consiste nel fatto che i significati delle espressioni elementari sono gli unici vincoli sul contenuto nel corso della elaborazione. Dall’altra la sintassi delle espressioni elementari è l’unico vincolo sulla struttura. Quindi si ipotizza che si elaborano le frasi mantenendo una ipotesi corrente sulla struttura e sul possibile significato fin li ma non spiega altri fenomeni su come è possibile rivedere queste elaborazioni quando si arriva ad elementi che contrastano l’analisi fino al passo sigma.

2 elementi caratteristici dalla composizione semplice:

1. **Indipendenza dal contesto**: Il significato dell’espressione complessa non dipende dal contesto specifico in cui occorre
2. **Inside out principio:** la formazione del significato dell’enunciato complesso procede dal basso verso l’alto accorpando elementi per costituire costituienti via via più articolati all’interno.

Quindi l’elaborazione semantica di espressioni complesse è una forma di applicazione di funzioni, quindi i signficati delle espressioni elementari non si modificano dal contesto. **Esempio:** Se spiego una frase S qualunque, posso elaborare

frasi come “**penso che S**”, quindi complico il modello costruendo uno da cui S dipende. La composizione semplice permette di gestire ed elaborare espressioni di questa forma.

**Obiezioni all’Indipendenza dal contesto:** La composizione semplice non gestisce le ”**progressive construction”.** Mentre capiamo una frase come “lei assomiglia a sua madre”, frasi come “lei sta assomigliando a sua madre” non lo capiamo, mentre viceversa si capisce “lei assomiglia sempre di più a sua madre”. Il significato di assomigliare è in qualche modo funzione del contesto in cui occorre, contraddicendo l’ipotesi di indipendenza del contesto.

**Obiezioni all’Inside out:** L’attribuzione di significato è un processo che avviene anche in senso contrario all’inside out ovvero **Top-down**. Se il significato di “Mary sta assomigliando a sua madre” non può essere determinato perchè procediamo in maniera context-dependend allora significa che non è vero il principio inside out.

Esistono 2 forme di elaborazione top-down consistenti con l’ipotesi dell’incapsulamento informazionale:

1. L’informazione restituita rimane all’interno del modulo in modo analogo a quanto viene memorizzata all’interno del lessico
2. L’informazione arriva direttamente da un altro modulo o da un sistema generale di controllo dei moduli

**Principio del contesto:** Da frege, le espressioni elementari non ammettono un significato isolato. Principio opposto alla composizionalità semplice, ma solo come elementi di frasi complesse. Funzione di interpretazione di frege! Esempi di frase in cui il contesto è importante: “he is intoxicated”, “he is pissed” significano entrambi “lui è sbronzo”. Però non in tutti i contesti le due espressioni sono accettabili. C’è un principio di significato contestuale che viola la composizionalità nella sua forma più forte.

Siamo partiti dal fatto che la composizionalità può spiegare la costruzione del significato complesso usando l’economicità del linguaggio. Economico vuol dire cercare di bilanciare la quantità di memoria e di capacità di calcolo dall’altra. Il principio del contesto comporta che la memoria data necessaria contiene parecchia informazione sulle frasi che possono essere costruite da espressioni atomiche e nei contesti accettabili. Se devo immaginare una base di conoscenza complessa ovviamente stiamo spendendo tantissima memoria. Questo è contrario alla ipotesi di economicità e di modularità dell’architettura preposta all’elaborazione di questi fenomeni.

**Classi di fenomeni linguistici contro la composizionalità:**

1. Illusione semantica
2. Garden path
3. Even related brain potentials
4. Attrazione semantica
5. Coercion

**Illusione semantica:** La prima classe è l’illusione semantica. L’esempio è “**L’illusione di Mosè**”: Quanti animali di ogni specie Mosè ha caricato sulla sua arca?” “Due per ogni specie”. Il problema è che non fu “Mosè” ma “Noè”. L’illusione semantica ricorre largamente in larga scala. Se si trattasse semplicemente di attribuire significati elementari e regole sintattiche di composizionalità nessuno si sbaglierebbe. Siamo fuorviati dal fatto che selezioniamo solo alcuni tratti significativi. L’illusione è un altro controesempio all’ipotesi di composizionalità perchè esistono di meccanismi di selezionamento di tratti significativi che sono contestuali e incrementali durante il recupero di significati. La composizionalità può essere raffinata ammettendo rappresentazioni lessico semantiche predefinite ma siamo nel contesto in cui stiamo ipotizzando moduli esterni o elementi aggiuntivi oltre quelli della composizionalità

**Garden path:** Un altro elemento che non può essere colto dal principio di composizionalità. Rappresenta il meccanismo che rompe l’analisi sintattica di una certa frase. Esempio: “While Anna dressed the baby played in the crib”. In ogni caso abbiamo il “the baby” che dopo aver ascoltato il verbo, diventa l’oggetto diretto del verbo. Abbiamo un sintagma “the baby” che gioca 2 ruoli sintattici che sono incompatibili. Non ci può essere un nodo dell’albero etichettato con 2 etichette. Siamo costretti ad aggiungere la virgola dopo “dressed”. Queste interpretazioni erronee può essere usato per provare che l’interpretazione di queste frasi non è composizionale.

L’interpretazione è non monotona. Consente la revisione di strutture costruite in precedenza. Ma non viene detto come viene creata la non monotonia. Backtracking? programmazione dinamica?

**Event related brain potentials:** sono strumenti che consistono in rilevazione di impulsi elettrici e onde. Tra le varie onde che il cervello emette in presenza di immagini e suoni ne consideriamo 2:

* **N400**
* **P600**

Sono onde importanti negli studi sull’impatto della composizionalità. La prima N400, è un’ onda che viene prodotta dalla ricezione di ogni content word ma non sempre con la stessa ampiezza a seconda della semantic reletedness che c’è tra gli elementi della frase. La prendiamo come spia della complessità che l’ascoltatore affronta nel momento in cui tende a combinare il significato dei termini che composizional

mente non stanno bene. Scatta.

P600 è un’onda che viene prodotta in presenza di violazioni a costrutti sintattici abituali. Ad esempio la struttura della locuzione, o strutture come “garden path”. In generale ad alta complessità.

Queste 2 onde rilevano la reletedness dei significati (N400 come possono essere combinati i risultati atomici) e una spia che si accende in presenza di violazioni sintattiche. N400 è un indice di complessità di composizione, P600 è un indicatore di tempo e risorse richieste per inserire un termine in un albero sintattico in costruzione.

Esempio: “The peanut was salted”. E’ insolita perchè peanut potrebbero essere i cartoni animato. P600 alta. Violazione della composizionalità: i sensi sono context indipendent. (la composizionalità semplice è indipendente dal contesto. Non può essere così. E’ il contesto che determina l’ampiezza della N400, violiamo il principio dell’inside out, ci sono degli elementi che vanno dall’alto al basso, dal contesto fino alla determinazione dei significati dei contesti.)

**Attrazione semantica**: Problema sintattico, osserviamo una violazione sintattica, studiamo la P600

* The hearty meal was devouring the kids. (P600 alta)
* The hearty meal was devoured by the kids.

La spiegazione di questi effetti è che occorre un fenomeno di attrazione semantica quindi viene scelta l’interpretazione più plausibile. Questo è alla base dei vincoli durante l’elaborazione linguistica. I vincoli sovrascrivono delle evidenze sintattiche.

**Coercion**: Significa conversione di tipo semantico (parlato durante la teoria del significato). Esempio: “ho iniziato un articolo” contro “Ho scritto un articolo”. Iniziare non è un evento ma viene forzato ad essere un evento. E’ uno dei potenziali fenomeni difficili da trattare. Non viene gestito dalla composizionalità. Iniziare lascia non esplicitata l’azione. Ancora una volta l’interpretazione deve essere ricostruita da altri indizi semantici e memorizzati a parte. Coercion è una forzatura di tipi. La costruzione del significato non è solo sintattica. Abbiamo bisogno di altra conoscenza al di fuori della composizionalità.

**Conclusione**

Per salvare la composizionalità e la modularità è necessario ipotizzare delle basi di conoscenza esterne al di fuori di quelle messe a disposizione. Problema di fondo. L’ipotesi della composizionalità è stata formulata con il principio di produttività. Ipotizzavamo un principio che spieghi l’economia. Adesso invece sappiamo che abbiamo bisogno di molte informazioni. (contesto, meccanismi dall’alto al basso, ipotizzare blocchi di informazioni)

**MA l’ipotesi di composizionalità è importante. E’ una ipotesi che non è stata mai abbandonata.**