



Alessandra Zarcone

Dipendenze a Distanza

(http://urts120.uni-trier.de/index.php/Dipendenze_a_distanza)

1 Introduzione

Il termine generico “dipendenze a distanza” (“unbounded distance constructions”, “long-distance dependencies” è spesso usato in letteratura per racchiudere diversi fenomeni:

topicalizzazioni

- [Nel cassetto]_i, ti ho detto che l'ho messo @_i.
- [Someone like Mary]_i, I wouldn't ask @_i anything.

frasi relative:

- Biagio mi ha restituito il libro che _i gli avevo prestato @_i.
- Biagio mi ha portato il libro che _i aveva promesso di prestarmi @_i.
- Mary read the book that _i John had lent @_i her .

wh-questions:

- [Che libro]_i hai comprato @_i ieri?
- [Che libro]_i ha detto Diego che Gabriella ha comprato @_i ieri?
- [Which book]_i did Mary buy @_i yesterday?
- John asked [which book]_i Mary bought @_i yesterday.

scissioni:

- È stato Andrea_i, che Biagio ha rimproverato @_i.
- It was John who_i Mary asked @_i something.

Tutti i fenomeni elencati comportano la dislocazione di un costituente (a volte chiamato *filler*, riempimento), che non si trova più *in situ* ma in un'altra posizione (a volte chiamata *gap*, o *trace*, traccia, che indichiamo negli esempi presenti in questo articolo con il simbolo "@").

Fenomeni di questo tipo sono quindi un problema nell'analisi sintattica delle lingue a ordine fisso, in cui si deve tener conto di cambiamenti nell'ordine dei costituenti, e in particolare sono problematici per quei formalismi fortemente basati sull'ordine lineare, come la Phrase Structure Grammar. Sono invece più elastici da questo punto di vista altri modelli, come la Dependency Grammar.

Un formalismo che descriva questi fenomeni deve quindi rendere conto di alcune cose:

1. la **posizione** del costituente non *in situ*;
2. la relazione di **dipendenza** tra il costituente non *in situ* e il suo posizionamento "non marcato";
3. il fatto che queste relazioni **non abbiano restrizioni** sulla distanza tra il costituente non *in situ* e il suo posizionamento "non marcato" (se non ovvie restrizioni legate a risorse fisiche della memoria del parlante).

2 Altri tipi di dipendenze a distanza

Altri fenomeni sono a volte compresi nella famiglia delle dipendenze a distanza, a volte considerati una "classe minore" (*weak unbounded dependencies* contro *strong unbounded dependencies*):

tough-constructions:

- Leonardo_i è sempre stato molto difficile da convincere @_i.
- Mary_i is tough to fool_i.

estrazione del soggetto:

- [Chi]_i @_i deve ancora arrivare?
- [Who]_i @_i is still to come?

(riportiamo separatamente il problema dell'estrazione del soggetto, più frequentemente ricondotto alle *Wh*-questions, perché viene considerato un problema computazionale di tipo diverso per l'approccio HPSG)

estrazione del modificatore:

- [Per quanto tempo]_i Vincenzo corre nel parco ogni mattina @_i?
- [For how many years]_i, had Mary thought that John had loved her @_i, before she found out that he was cheating on her?

vuoti parassitari:

- [Which president]_i did articles about @_i call @_i a racist? (ma: * Who_i were articles about @_i very revealing?)

3 Le dipendenze a distanza in un approccio trasformativo

Nella teoria Generativo-Trasformativa (Chomsky 1965) le dipendenze a distanza venivano spiegate con il movimento di un costituente dalla sua posizione standard nella frase non marcata: esisterebbe una **struttura profonda** ("deep structure" o "d-structure") in cui il costituente si trova *in situ*, e un processo di movimento genererebbe una **struttura superficiale** ("surface structure" o "s-structure") in cui il costituente non è più in quella posizione. Per tenere conto di questi fenomeni si postulavano operazioni di "cancellazione" e di "riscrittura". Erano tuttavia necessari maggiori vincoli sul tipo di operazioni effettuabili, perché la grammatica rischiava di ipergenerare.

Chomsky (1973) assumeva quindi a livello di struttura superficiale l'esistenza di una **traccia**, fonologicamente vuota, co-indicizzata con il costituente estratto, nella posizione da cui si è effettuata la cancellazione (o il "movimento"). Si veda l'esempio:

- [Someone like Mary]_i, I wouldn't ask @_i anything.

L'elemento in parentesi quadre non è *in situ*, ma il suo movimento lascia un gap (un vuoto) o, più precisamente, una categoria vuota (la traccia).

L'assunzione di una traccia, e di un meccanismo di associazione tra la traccia e l'elemento estratto dovrebbero trovare prova psicolinguistica. Inoltre, assumere l'esistenza di processi di movimento richiederebbe una distinzione tra classi di frasi senza movimento (e quindi di più rapida e facile elaborazione) e classi di frasi con movimento (e quindi di più complessa e lenta elaborazione). Una frase come "I wouldn't ask someone like Mary anything" sarebbe quindi meno complessa di "Someone like Mary, I wouldn't ask anything".

Traccia e componente trasformazionale sono quindi stati spesso bersaglio di critiche da parte di altre correnti teoriche, che hanno proposto modelli di estrazione e dipendenze a distanza differenti.

4 Il problema della trasformazionalità

Le prime ricerche psicolinguistiche volte a dimostrare l'esistenza di meccanismi analoghi alle trasformazioni sintattiche nella comprensione di frasi da parte di parlanti (operazioni di **detrasformazione**, o trasformazione inversa) non hanno avuto un buon esito (Fodor 1978). Frasi per cui si postulava un numero di operazioni di trasformazione maggiore non si rivelavano più complesse da interpretare di altre più vicine a una "struttura profonda". Piuttosto, sembra che in comprensione i parlanti interpretino **incrementalmente** almeno alcune parti di ciò che sentono, integrandole in una prima struttura sintattica, senza aspettare ulteriori input; semmai analisi sintattiche sbagliate sono rivedute in tempo reale secondo meccanismi di rianalisi.

Esistono diversi formalismi sintattici non trasformazionali, come la Dependency Grammar (e la sua versione Word Grammar di Hudson), la Phrase Structure Grammar di Gazdar, la Tree-Adjoining Grammar e la Head-driven Phrase Structure Grammar (HPSG).

4.1 La PSG di Gazdar e le dipendenze a distanza

Gazdar (1981) propone di eliminare la componente trasformazionale della Grammatica Generativa, assumendo che una PSG ([Phrase Structure Grammar](#)), con l'aggiunta di alcuni **simboli complessi** possa rendere conto di fenomeni come le dipendenze a distanza.

I nuovi simboli sono del tipo S/S, S/NP, NP/NP e NP/S. La notazione è così spiegata: un nodo A/B governerà un nodo di tipo A, che a sua volta governerà un sottoalbero contenente un "buco" (*hole*, traccia) di tipo B.

A queste nuove categorie si aggiunge un insieme di **regole derivate** dalle regole base. Ad esempio ("Q" sta per "Question"):

- S/NP -> NP VP/NP

- Q -> NP S/NP

Un nodo di tipo S/NP governerà un nodo di tipo S (una frase dichiarativa), che a sua volta governerà un sottoalbero contenente un "buco" di tipo NP (e il sottoalbero sarà dunque più o meno ramificato, senza vincoli di distanza tra il nodo radice e il nodo complesso che incide la traccia). Ad esempio, si vedano i seguenti passaggi di riscrittura:

- S -> NP_i S/NP_i
- S -> NP_i NP VP/NP_i
- S -> NP_i NP ha comprato @_i
- S -> NP_i Chiara ha comprato @_i
- S -> Dei libri_i Chiara ha comprato @_i

VP/NP potrà dunque rendere conto di un sintagma che ha al suo interno ha un "buco", e il meccanismo di riscrittura permette al nodo radice S/NP di ereditarne la traccia fino al riempimento del costituente non *in situ*.

5 Il problema del collegamento *filler-gap*

In una frase con tre possibili gap ('gap' è in un certo senso il correlato teorico della traccia in modelli psicolinguistici) come in (1) è necessario collegare il primo costituente all'ultimo dei tre possibili gap:

- (1) [What]_i do you want (@) Mary to sing (@) to the baby about @_i?

Un modello della comprensione delle dipendenze a distanza che assuma l'esistenza di gap deve tener conto del problema dell'associazione del costituente "spostato" con una e una sola posizione nella frase (**identificazione del gap** o *gap-filling process*, Fodor 1978). Anche in un modello detrasformazionale questo problema deve essere risolto, perché è comunque indispensabile collegare il costituente a una posizione nella struttura profonda, prima di riportarlo a quella posizione.

Altre teorie come la Word Grammar di Hudson (1984) non richiedono invece la postulazione di tracce o gap, ma stabiliscono una relazione di dipendenza tra l'elemento "estratto" e il costituente che lo sottocategorizza. L'esempio in (1) verrebbe quindi rappresentato in modo diverso:

- (2) [What]_i do you want Mary to sing to the baby about_i?

Notare che questa volta sono coindicizzati non più il costituente estratto e la traccia ma il costituente e i diversi elementi di cui può essere argomento. Il problema è quindi non più collegare *filler* e *gap*, ma predicato e argomento.

5.1 La HPSG e le dipendenze a distanza

Nella storia della Head-driven Phrase Structure Grammar si distinguono tre approcci diversi, sviluppati in momenti diversi:

1. **struttura filler-gap** (Pollard e Sag, 1994, capitolo 4): il primo approccio fa uso di una traccia all'interno del lessico, e di schemi di formazione delle frasi per collegarla con il componente estratto (*head-filler schema*); la traccia è un segno linguistico vero e proprio, ma privo di realizzazione fonetica, e si combina con gli altri costituenti creando una posizione di *filler* (TO-BIND|SLASH), coindicizzata con la posizione della traccia, e questa posizione dovrà necessariamente essere riempita perché la frase sia considerata grammaticale; il vincolo del riempimento del filler può risalire l'albero sintattico, senza limiti espliciti sulla distanza della dipendenza;

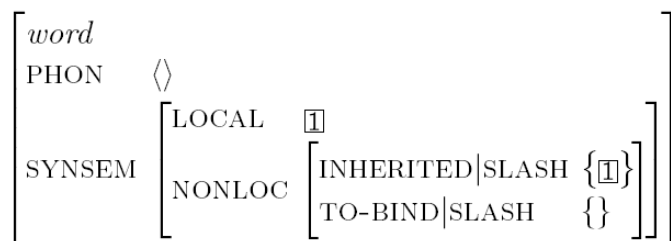


Figura 1: La traccia nella prima versione della HPSG

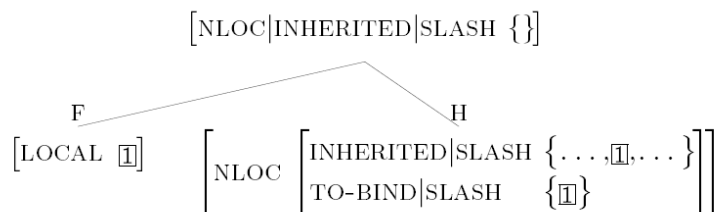


Figura 2: La saturazione del TO-BIND|SLASH nella prima versione della HPSG

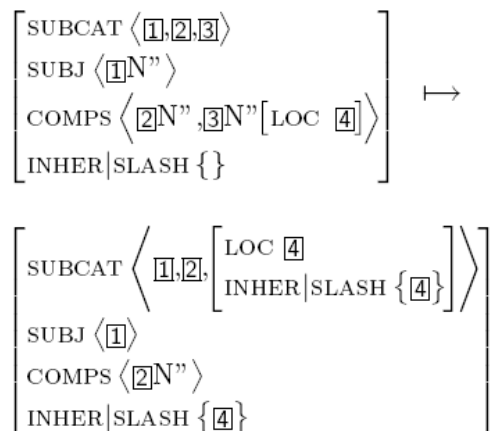


Figura 3: Esempio di regola lessicale della seconda versione della HPSG

2. **analisi senza traccia** (Pollard e Sag, 1994, capitolo 9): il secondo approccio riconosce come problematica la presenza di un elemento fonologicamente vuoto nel lessico (sia da un punto di vista computazionale sia da un punto di vista psicolinguistico), nega la presenza di tracce e fa invece uso di regole lessicali, per creare predicati più complessi e integrarli con il componente estratto appena possibile, creando la dipendenza prima di un eventuale gap. Questa seconda soluzione permette a ogni struttura sintattica di "generare" una struttura con dipendenza a distanza, creando una sorta di "meccanismo trasformatore" all'interno di un formalismo che vuole essere fortemente anti-derivazionale; l'approccio è molto vicino a quello di Gazdar (1981), perché le regole lessicali generano categorie "slashed", che possono includere al loro interno i gap;
3. **analisi senza regole lessicali** (Bouma et al. 1998): Bouma et al. cercano di unificare i fenomeni di dipendenze a distanza senza fare uso di tracce o di regole lessicali, reinserendo la traccia nella gerarchia dei tipi di segni previsti dal formalismo (non più dunque nel lessico); mentre prima la traccia, essendo inserita nel lessico, era solo uno dei possibili costituenti, in questo approccio essa non è un segno linguistico, ma un "contenuto" per uno slot argomentale non soddisfatto localmente. Se un argomento non viene soddisfatto localmente, un *filler* viene cercato non *in situ* nel seguente modo: il costituente fa ereditare lo slot *filler* vuoto ai nodi superiori, di cui è testa. Questa posizione "vuota" può risalire quindi l'albero sintattico, senza limiti espliciti sulla distanza della dipendenza. L'albero è incompleto fino a quando, con il raggiungimento del filler, la posizione viene saturata.

6 Il problema delle categorie vuote e la creazione della dipendenza

6.1 A favore delle categorie vuote: l'effetto gap-filler

Swinney et al. (1988), Crain e Fodor (1985) e Stowe (1986) riportano fenomeni interpretabili come effetti *gap-filler*. Swinney et al. registrano un effetto di *priming* di parole associate al costituente estratto dopo la posizione del presunto gap ma non prima (*gap-filler effect*). Ad esempio:

- (3) The policeman saw the boy that the crowd at the party accused of the crime

In (3) gli associati di *boy* sarebbero attivati dopo *accused* ma non prima. Stowe (1986) mette a confronto coppie di frasi come (4) e (5):

- (4) My brother wanted to know if Ruth will bring **us** home to Mum at Christmas
 (5) My brother wanted to know who Ruth will bring **us** home at Christmas

I tempi di lettura di *us* sono più lunghi in (5), perché *bring* verrebbe associato a un gap, ma questa prima analisi verrebbe poi smentita da *us*, che farebbe scattare un meccanismo di reanalisi. Tuttavia, come notano Pickering e Barry (1991), esempi di

questo tipo sono costruiti in modo tale che il presunto gap seguirebbe il verbo, quindi tali dati non decidono necessariamente in favore dell'esistenza di categorie vuote, ma possono anche essere spiegati da meccanismi di dipendenza come in (2).

6.2 Contro le categorie vuote: l'ipotesi di attivazione immediata

Pickering e Barry (1991) forniscono esempi in cui un approccio con gap e un approccio senza gap porterebbero a previsioni diverse, come:

- (6) [In which box]_i did you put the cake @_i?
- (7) [In which box]_i did you put_i the cake?

Sarebbe inefficiente e non incrementale un modello in cui si attende la fine della frase per un gap, piuttosto che stabilire la dipendenza a livello del predicato. La differenza in termini di efficienza sarebbe ulteriormente evidente in frasi particolarmente lunghe come:

- (8) We gave every student capable of answering every single tricky question on the details of the extremely complicated theory a prize.
- (9) That's the prize_i we gave every student capable of answering every single tricky question on the details of the extremely complicated theory @_i.
- (10) [Which box]_i did you put the very large and beautifully decorated wedding cake bought from the expensive bakery in_i?
- (11) [In which box]_i did you put_i the very large and beautifully decorated wedding cake bought from the expensive bakery?

La presenza di categorie vuote richiederebbe che la frase (9) sia più difficile della frase (8), perché il gap richiederebbe di essere mantenuto in memoria fino alla fine della frase, mentre le frasi (10) e (11) dovrebbero essere comparabili. Il giudizio di naturalità del parlante (Pickering e Barry 1991) è invece esattamente l'opposto: Pickering e Barry giudicano (8) e (10) significativamente più complesse da analizzare rispetto a (9) e (11). Sembrerebbe quindi che in (9) e (11) l'elemento non in loco sia integrato in un'analisi sintattica incrementale quando viene analizzato il predicato, e non alla fine della frase.

6.3 L'ipotesi di doppia attivazione

I dati raccolti da Nicol (1993) sembrano essere compatibili sia con l'idea, più vicina alle Grammatiche della dipendenza, che il costituente estratto venga subito associato al predicato, sia con l'idea del "filler-gap effect", in quanto si osserverebbe una duplice attivazione del costituente, sia dopo il predicato, sia in posizione di gap. Nicol propone quindi un modello in cui il predicato attivi un ruolo tematico per il costituente, e che questa prima ipotesi venga poi consegnata alla posizione del gap. Sag e Pollard (1994) suggeriscono un modello di dipendenze a più livelli, "semantiche" (collegamento ruolo tematico-predicato) e sintattiche (collegamento filler-gap), non necessariamente sovrapponibili (è tipico del formalismo HPSG distinguere, seppur in un approccio non-derivazionale, incrementale e integrato, questi livelli).

6.4 Eye-tracking e Potenziali evocati

Gli studi di eye-tracking di Traxler et al. (1996) confermano il fatto che i parlanti stabiliscano subito una relazione di dipendenza tra il predicato e il costituente estratto (ed eventualmente ricorrono in seguito a meccanismi di reanalisi quando questa relazione dovesse risultare errata), ma non trovano prova sperimentale della "seconda attivazione" dovuta al gap prevista da Nicol (1993).

Gli studi con i potenziali evocati di Kaan et al. (2000) rilevano una componente P600 quando un soggetto riceve come input un verbo che completa una relazione di dipendenza con un precedente elemento *wh* della frase, e suggeriscono che la sua ampiezza dipenda proporzionalmente dalla lunghezza della dipendenza (ricordiamo che la componente N400 è genericamente associata alla rianalisi semantica, la P600 a quella sintattica). Questa previsione è verificata da Philips et al. (2005), che trovano i seguenti risultati:

1. una negatività anteriore prolungata (*sustained anterior negativity*, da non confondere con la N400), che interpretano come il risultato di un mantenimento in memoria di lavoro del componente "semantico" estratto
2. l'ampiezza della P600 non è invece influenzata dalla lunghezza della dipendenza, mentre invece il suo onset è ritardato in dipendenze a lunga distanza. Philips et al. ricollegano questo dato all'ipotesi che la P600 non si manifesti solo nella rianalisi di anomalie, ma anche quando non si riscontrano ambiguità sintattiche.

7 Riferimenti Bibliografici

- Bouma, G., Malouf, R., Sag, I. A. 1998. "A unified theory of complement, adjunct, and subject extraction. In: Proceedings of the Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar, and Categorical Grammar. Saarbrücken. Pp. 83-97.
- Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.
- Chomsky, Noam. 1973. *Conditions on Transformations*. In: S.R. Anderson and P. Kiparsky (eds), *A Festschrift for Morris Halle*. London: Holt, Rinehart & Winston.
- Chomsky, Noam. 1981. Lectures on Government and Binding. Dordrecht: Foris.
- Crain, S., Fodor, J.D. 1985. How can grammars help parsers?. In: D.R. Dowty, L. Karttunen, A.M.Zwicky (Eds.), *Natural Language parsing: Psychological, computational and theoretical perspectives*. Cambridge: Cambridge University Press.
- Fodor, Janet Dean, 1978. Parsing strategies and constraints on transformations. In: *Linguistic Inquiry*, 8, 427-473.
- Gazdar, Gerald. 1981. Unbounded Dependencies and Coordinate Structure. In: *Linguistic Inquiry*, 12, 155-184.
- Kaan, E., Harris, A., Gibson, E., Holcomb, P. 2000. The P600 as an index of syntactic integration difficulty, *Language and Cognitive Processes* 15, 159– 201.
- Lasnik, Howard & Saito, Mamoru. 1992. *Move alpha: Conditions on Its Application and Output*. Cambridge, Mass.: MIT Press.

- Nicol, J.L. 1993. Reconsidering activation. In: Gerry T. M. Altmann, Richard Shillcock, Cognitive Models of Speech Processing: The Second Sperlonga Meeting. Psychology Press.
- Pickering, Martin & Barry, Guy. 1991. Sentence processing without empty categories. In: Language and Cognitive Processes, 6, 229-259.
- Phillips, C., Kazanina, N., Abada, S. 2005. ERP Effects of the Processing of Syntactic Long-distance Dependencies. In: Cognitive Brain Research, 22, 407-428.
- Pollard, C., Sag, Ivan A. 1994. Head-Driven Phrase Structure Grammar.
- Sag, Ivan A., Pollard, C. 1994. Extraction without traces. In: West Coast Conference on Formal Linguistics, 365-384. Stanford Linguistics Association.
- Stowe, L.A. 1986. Parsing WH-constructions: Evidence for on-line gap location. In: Language and Cognitive Processes, 1, 227-245.
- Swinney, D., Ford, M., Frauenfelder, U., Bresnan, J. 1988, *On the temporal course of gap-filling and antecedent assignment during sentence comprehension*. In: B. Grosz, R. Kaplan, M. Macken, I. Sag (Eds.), *Language structure and processing*, Stanford. Calif.: CSLI.
- Traxler, M.J., and Pickering, M.J. 1996. *Plausibility and the processing of unbounded dependencies: An eye-tracking study*. In: *Journal of Memory & Language*, 35, 454-475.