

1. Introduzione

In questo lavoro vorrei discutere brevemente come quella che appare come una metodologia molto tradizionale e considerata spesso datata possa essere in realtà utilizzata per determinare con un grado di precisione estremamente alto quali sono esattamente i tratti semantici degli elementi funzionali, che per loro natura sono spesso elusivi. Recenti sviluppi in varie branche applicate della linguistica (acquisizione della prima e della seconda lingua, disturbi del linguaggio di vario genere, psicolinguistica) hanno portato ad una sorta di contrapposizione tra la linguistica sperimentale e quella teorica. Sono del parere che non ci possa essere linguistica sperimentale senza una teoria forte e precisa alle spalle, non è nemmeno pensabile di poter costruire dei test senza aver prima una ipotesi esatta relativa al fenomeno che si sta indagando. D' altra parte è vero che la linguistica teorica si è spesso concentrata su nozioni così astratte da essere non determinabili empiricamente. Ritengo che questa divisione tra linguistica teorica e linguistica sperimentale sia estremamente deleteria in entrambi i sensi e che la disciplina sia unica e non possa prescindere dalla continua interazione tra empiria e riflessione astratta, che si compensano e integrano in una serie di passi successivi come in una spirale. Un fondamentale passo in avanti in ambito empirico è dato dai nuovi strumenti di lavoro quali data base di ogni genere che sono stati costruiti in ambito tipologico, dialettologico, diacronico e di acquisizione e che oggi possono essere utilizzati molto semplicemente come risorse in rete. Credo sia quindi necessario aprire un dibattito su come utilizzare "big data" in ambito linguistico, che tocchi ovviamente sia il nostro modo di fare teoria linguistica che i tipi di metodologie utilizzati per la raccolta e il trattamento dei dati, che devono essere il più possibile standardizzate per poter essere comparabili. E peraltro vero che la grande massa di dati ci permette di depurare le nostre generalizzazioni da eventuali deviazioni date da singoli effetti, ma è anche vero che se i dati vengono raccolti sistematicamente in modo diverso si possono avere distorsioni a largo spettro che rendono i dati non comparabili. In questa comunicazione cercherò di dimostrare che la dialettologia è sempre stata per sua natura fin dagli albori una disciplina sperimentale, e che adottare una prospettiva qualitativa e al contempo una quantitativa consente una integrazione della potenza numerica con la precisione del dato qualitativo che contribuiscono ad una teoria più precisa e completa. La dialettologia è stata vista in anni recenti da un lato come una disciplina eminentemente sociolinguistica, che descrive mutamenti sociali e il modo in cui diverse varianti si diffondono in una comunità linguistica variegata e complessa, senza tener conto che il dato microvariazionale può in effetti essere rivelatore dell'architettura dell'intero sistema tanto quanto quello tipologico. La dialettologia è anche stata vista come contrapposta alla variazione tipologica, che ci mostrerebbe uno spettro molto più ampio di variazione, mentre la dialettologia ci appare come una visione al microscopio e ci permette di vedere dettagli che non sono visibili alla variazione ad ampio spettro. Approcci formali come quello sviluppato in Roberts (2012) codificano questa visione che i parametri di variazione che osserviamo nella variazione dialettologica siano di ordine fondamentalmente diverso da quelli che osserviamo nella variazione tipologica. Un esempio di questo tipo è l'ordine basico delle parole, che può variare a livello di macrovariazione, ma non di microvariazione. In Poletto (2012) ho cercato di mostrare in dettaglio che a ben guardare la variazione che troviamo all'interno di una singola area dialettale contiene lo stesso tipo di possibili varianti che ritroviamo in tipologia, la differenza tra i due tipi di variazione è soltanto quantitativa, nel senso che in una stessa area linguistica alcune varianti sono considerevolmente meno

frequenti di altre, ma comunque non escluse a priori. I confini della micro e macrovariazione sono quindi fondamentalmente gli stessi. In Poletto (2012) ho cercato di dimostrare empiricamente questa assunzione sulla base di esempi, che riassumo brevemente qui. Sappiamo che con tutta probabilità non esistono dialetti romanzi con un ordine non marcato delle parole SOV, tutte le varietà romanze investigate fin qui mostrano un ordine basico SVO. Come nota Roberts (2012) l'ordine di base delle parole dimostra inoltre un grado di stabilità temporale molto alto, dato che è mutato dal latino alle lingue romanze nel corso di un numero consistente di secoli ed è poi rimasto lo stesso. Tuttavia, se consideriamo invece specifiche strutture, notiamo che un ordine SOV è possibile anche nelle varietà romanze: senza voler citare il caso ovvio della posizione dei pronomi clitici, che mantiene appunto l'ordine SOV del latino, si noti che ad esempio nelle varietà friulane ci sono casi di ordini che sembrano quelli tipici delle lingue SOV in particolari contesti che corrispondono semanticamente ad un aspetto compiuto:¹

- (1) a. O ai lis sigaretis desmenteadis
 b. O ai desmenteadis/desmenteat lis sigaretis

Costruzioni che mostrano un ordine diverso dei costituenti con regole precise, come quella che si riscontra in friulano, dove l'anteposizione dell'oggetto al participio comporta accordo obbligatorio del participio passato, mentre la postposizione dell'oggetto permette entrambe le opzioni, (cioè accordo con l'oggetto o accordo non marcato al maschile singolare) sono abbastanza frequenti nelle varietà italo-romanze, anche il fiorentino antico mostra lo stesso tipo di regola di accordo (vedi Egerland 1996 e Poletto 2014). Quando definiamo i tipi linguistici non dobbiamo dimenticare che regole appartenenti ad altri tipi non sono frequenti, ma non impossibili a priori. Ad esempio, tutti i dialetti italiani possono essere definiti come lingue flessive. Nessun dialetto italiano funziona dunque come il turco, in cui i morfemi sono chiaramente identificabili ed ogni morfema ha un'unica valenza semantica¹¹¹:

- (2) oku-ya-ma-yabil-ir-im
 read-Abil-Neg-Abil-Aor-1sg
 'potrei non riuscire a leggere'

Questo però non vuol dire che non ci siano esempi di morfologia agglutinante nelle varietà italiane: un caso di questo tipo riguarda la flessione del congiuntivo in alcune varietà di lombardo alpino descritta in origine da Salvioni ed analizzata da Benincà (1996), che mostra come quello che etimologicamente è un avverbio derivante da 'bene'² è stato rianalizzato come un morfema modale che marca il congiuntivo, e si aggiunge alla destra del morfema di accordo di persona e numero, che in genere nelle lingue romanze "chiude" la parola e non permette l'aggiunta di altri morfemi alla propria destra. Tuttavia, come mostra Benincà (1996:3), è possibile trattare il morfema *-be* come un caso di morfologia agglutinante, perché si aggiunge alla parola già completa del morfema di persona che corrisponde alla forma del presente indicativo per formare il congiuntivo presente, ad esempio :

¹ Si veda il lavoro recente di Cinque (2016) per una analisi microvariazionale dell'ordine di base OV/VO.

² L'avverbio *bene* ha in italiano vari valori, non è solo un avverbio di maniera, ma anche un avverbio aspettuale nella sua variante *ben* e può avere in alcuni casi un valore modale. Si noti inoltre che *ben* entra nella composizione del complementatore complesso *benché*, quindi possiamo dire che è un elemento polifunzionale.

(3) *avéj* “avere”

a. PRES INDIC	b. PRES SUBJUN
<i>e</i>	<i>ke ebe</i>
<i>t as</i>	<i>ke t asbe</i>
<i>l a</i>	<i>ke l abe</i>
<i>on</i>	<i>ke onbe</i>
<i>ei</i>	<i>ke eibe</i>
<i>i a</i>	<i>ke i abe</i>

Quindi la morfologia agglutinante non è di per sé un' opzione esclusa, è semplicemente molto rara nelle varietà italo-romanze. Potrei continuare con una serie di altri esempi di costruzioni che apparentemente violano il tipo linguistico italo-romanzo, ma che sono, anche se non frequenti, sicuramente attestate, ma credo che il punto sia chiaro. I principi che portano a caratteristiche tipologiche diverse da quelle tipiche del dominio romanzo non sono bloccati di per sé, sono semplicemente molto più rari o non sistematici. L'idea che la microvariazione che osserviamo tra dialetti sia un tipo più superficiale e implichi categorie di un ordine più piccolo di variazione rispetto a quella tipologica è fondamentalmente errata non solo sulla base del rasoio di Occam, che ci impone di non aggiungere categorie superflue se possiamo farne a meno, ma anche sbagliata dal punto di vista empirico. L'osservazione che la variazione dialettale e quella tipologica sono fondamentalmente dello stesso tipo che ho cercato di provare sistematicamente in Poletto (2012) elimina immediatamente la possibilità che la variazione possibile si possa suddividere sulla base di tipi di parametri diversi, alcuni più fondamentali e stabili di altri ed elimina anche l'idea, che è stata elaborata in origine da Baker (1997) e più recentemente formalizzata da Roberts (2012) che un determinato valore parametrico ne implichi automaticamente altri a cascata. Credo piuttosto che la variazione non vada organizzata in maniera gerarchica, ma per campi, in cui varie proprietà sono connesse tra di loro in una rete ma non in senso necessariamente implicazionale. Anche gli universali implicazionali, come la scuola tipologica di Greenberg ha ben mostrato, sono formalizzabili come tendenze e non come implicazioni matematiche assolute, il che mostra che ci sono eccezioni. Quello che dobbiamo scoprire quindi, non è perché alcuni parametri ne implicano automaticamente altri, ma perché alcune combinazioni siano significativamente più rare di altre e cosa significano le “tendenze” che osserviamo. Ovviamente la primacosa che ci resta da spiegare adottando una prospettiva che non distingue tra micro e macro-variazione è il motivo per cui lingue geneticamente affini “tendono a convergere” verso grammatiche simili e soprattutto le mantengano per un arco di tempo molto lungo. In realtà quello che sorprende nella variazione dialettale è quindi un fatto quantitativo, non qualitativo, cioè la rarità di alcuni fenomeni. Che cosa significa in termini di acquisizione dire che un fenomeno è raro? Che spazio ha un'osservazione di tipo statistico in una teoria che adotta regole senza eccezioni? Significa che dobbiamo pensare che il sistema di acquisizione del linguaggio lavori su base statistica? Si potrebbe perfino pensare che tutti i dialetti siano in un certo senso la stessa lingua e che le differenze, che sono più “superficiali” vengono apprese più tardi. Credo invece che l'uniformità di alcuni fenomeni in ambito dialettologico ma non tipologico sia da spiegare in altro modo. Credo che la spiegazione sia molto semplice, e si basi sull'osservazione direi quasi banale, che lingue geneticamente simili hanno ereditato lo stesso lessico. Dato che ammettiamo che la variazione areale e il mutamento diacronico siano dovuti a caratteristiche degli elementi lessicali funzionali, è chiaro che avendo i lessemi che esprimono tratti funzionali una semantica simile, il mutamento produrrà percorsi simili di sviluppo, dato

che il mutamento linguistico in genere non stravolge le proprietà iniziali dell' elemento funzionale che lo realizza lessicalmente ma al contrario si basa proprio su queste. Ammettere che esiste un ordine solo di parametri che sono collegati tra loro tramite una rete e non tramite una gerarchia consente di mantenere una teoria lessicalista forte dell'acquisizione, che si rivela essere molto semplice: se pensiamo che la variazione linguistica dipenda dalle caratteristiche degli elementi lessicali ed in particolare dalle proprietà del lessico funzionale, tutto quello che il bambino deve fare è acquisire con le parole anche le loro proprietà formali definite da contesto in cui l' elemento funzionale viene usato. Le solidarietà tra parametri verranno spiegate tramite la presenza di tratti fondamentali simili nel sottoinsieme di parametri che adottano lo stesso valore.

Se la risposta alla domanda relativa al fatto che la variazione dialettale appare statisticamente più uniforme di quella tipologica è banale ed è data dal fatto che lingue geneticamente affini hanno un lessico analogo, le conseguenze di questa risposta non sono affatto banali. Visto che gli elementi funzionali sono stati grammaticalizzati da elementi con le stesse proprietà, anche la sintassi che ne è derivata è molto simile. Quindi, il lessico ha un ruolo fondamentale nella variazione sintattica, che finora è stato sottovalutato, quindi non si può costruire una teoria della variazione sintattica senza considerare il lessico e la sua storia. L'intento di questo lavoro è mostrare che, se il lessico è stato usato finora come campo di indagine per descrivere regole di mutamento diacronico fonetico, fonologico e morfologico, possiamo usarlo in maniera nuova anche per rendere conto della variazione sintattica. Mostrerò qui due esempi di utilizzo del lessico come indizio per capire quali siano i meccanismi interni della variazione sintattica: il primo di ordine qualitativo, il secondo di ordine quantitativo. Entrambi i tipi di approccio, quantitativo e qualitativo, sono in grado di mostrare in maniera interessante delle proprietà del sistema linguistico, e quindi non è necessario escludere l'uno o l'altro ma i due tipi di analisi vanno utilizzate in parallelo. Mostrerò come indizi lessicali ci possano condurre verso un'analisi più precisa dei fenomeni sia utilizzando una prospettiva qualitative, come nella sezione 2, sia utilizzando una prospettiva quantitativa, come nella sezione 3.

2. Un' analisi quantitativa: la distribuzione di *che* nelle interrogative indirette

Un fenomeno ben noto nella letteratura recente sulle frasi interrogative riguarda la presenza di un complementatore dopo il pronome interrogativo nelle domande indirette (alcuni dialetti hanno esteso questa struttura anche alla interrogative dirette, ma non mi occuperò qui di questa estensione, vedi Poletto (2000) per una descrizione della distribuzione del fenomeno):

- (4) a. I sai nen **anté che** mama a l abia catà le fior (Torino)
b. Dimi **ce c** al fas Giuan (Cesarolo, Friulano)

Il fenomeno è esteso a macchia di leopardo in tutta l' Italia settentrionale, dal Piemonte al Trentino, ed è presente ma spesso facoltativo anche in alcune varietà emiliane, lombarde,

- (5) a. An so minga **cus (c)** al faga Giani (Ferrarese Dialekt)
b. Dime **chi (c)** a l à piat el quader (Druento, Turin)

Il Veneto è probabilmente la regione in cui si ha la maggior uniformità del fenomeno, che risulta obbligatorio nelle interrogative indirette con tutti i *wh* in tutte le varietà (a parte l' area veronese).

(6) No so chi *(che) ze partio Veneziano

Lo stesso fenomeno si riscontra nei dialetti tedeschi, in particolare in area bavarese ed alemannica, come mostrano Bayer e Brandner (2008), che notano che anche nei dialetti alemannici parlati nell' area intorno al lago di Costanza abbiamo lo stesso fenomeno di introduzione del complementatore dopo il pronome interrogativo:

(7) I woass it **wieviel dass** er für des Auto zahlt hät
 lo so non quanto che lui per la macchina pagato ha

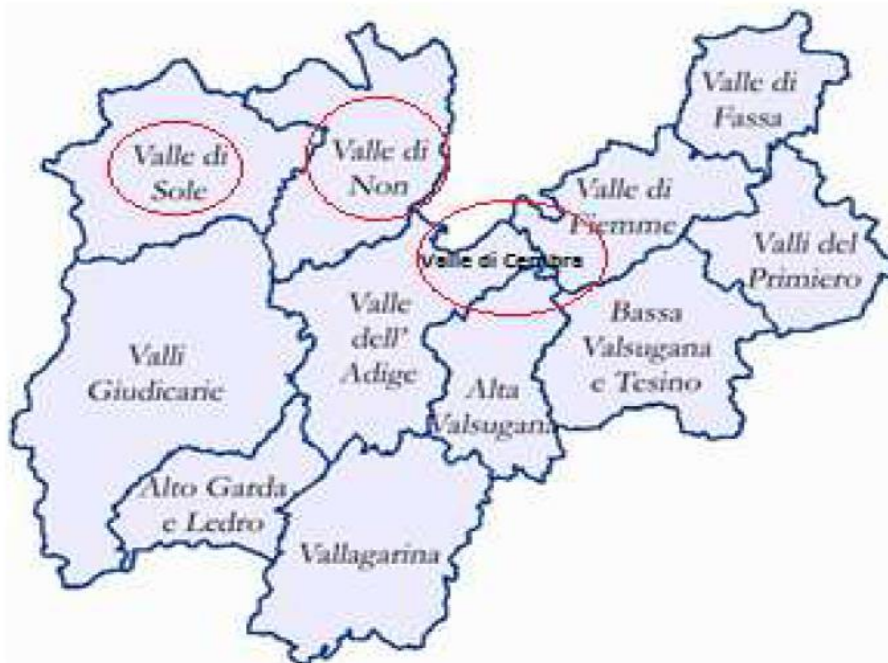
Si potrebbe quindi pensare che si tratti dello stesso fenomeno di raddoppiamento, magari proprio di un fenomeno di Sprachbund, visto che le due aree hanno in comune anche altri tratti, quali la perdita del passato remoto, l'uso di surcomposti (in recessione nell'Italia settentrionale), l'uso di verbi composti con particelle preposizionali, eccetera. Se però guardiamo al percorso di estensione del fenomeno dobbiamo ricrederci: nei dialetti alemannici il fenomeno comincia ad estendersi a partire dai wh complessi e si estende solo in alcune varietà a pronomi interrogativi semplici. Riporto qui di seguito la tabella di estensione della distribuzione del complementatore (abbreviato come DFC) dal lavoro di Bayer e Brandner (2008), che notano espressamente che i pronomi wh-monosillabici come *wie* "come", *wo* "dove" *was* "cosa" ecc. non occorrono con un complementatore.

Tabella 1

X-bar Status	Subtype	DFC-restriction
wh-phrase	Wh-DPs, Wh-PPs	best with overt C
wh-word I	<i>warum, wieviel, wem</i>	
wh-word II	<i>wer, wen, was, wie, wo</i>	worst with overt C



Un' indicazione etimologica che ci fa sospettare che non si tratti dello stesso fenomeno viene dal fatto che il complementatore germanico non è una forma di pronome interrogativo, ma una forma complessa che include un elemento di tipo d- (la radice dei pronomi dimostrativi e dell' articolo definito) mentre nelle lingue romanze la corrispondenza tra complementatore e l'elemento wh che corrisponde a una delle forme di "che cosa" è immediata. Osservando più da vicino tramite un'indagine sul campo la variazione dialettale in un ambito ristretto si arriva alla conclusione che in effetti l'indicazione che ci proviene dall'etimologia è corretta: il fenomeno osservato nei dialetti tedeschi non sembra essere lo stesso osservato nei dialetti italiani settentrionali. Per dare la dimostrazione circoscriverò il campo di indagine ai dialetti trentini, varietà in cui il fenomeno non è ancora grammaticalizzato per tutti gli elementi wh-, come ad esempio in Veneto, e che quindi ci consente di osservare in vivo il cambiamento linguistico. Riporto nella cartina qui sotto l' area investigata.



Nelle tre valli in cui è stata condotta l'indagine si possono individuare tre fattori indipendenti che influenzano la presenza del complementatore:

- A) Il tipo di elemento interrogativo
- B) La posizione pre/postverbale del soggetto
- C) La differenza tra interrogative dirette e indirette

Questi fattori erano già stati individuati in parte in Poletto e Vanelli (1993) e in Garzonio (2007) e mostrano effettivamente che la struttura compare prima nelle interrogative indirette, come è plausibile, visto che si tratta di una frase secondaria, anche se ci si potrebbe in realtà aspettare un complementatore di tipo *se*, che marca le interrogative polari e non il complementatore *che*, che marca le frasi dipendenti dichiarative. È interessante notare che anche nelle varietà francesi in cui il fenomeno è riportato e in inglese medio, che mostra almeno in maniera opzionale il fenomeno, il complementatore è sempre quello delle frasi dichiarative e mai quello delle interrogative polari. Nelle tre valli prese in esame, le interrogative dirette non mostrano la presenza di *che* ma forme di inversione del clitico soggetto, la struttura più conservativa dei dialetti alto-italiani.

Si noti inoltre che la posizione del soggetto nell'interrogativa indiretta è rilevante: la posizione postverbale favorisce la presenza del complementatore come mostra la seguente coppia di frasi:

- (8) a. Me domandi *quando che i nirà tuti i operai.*
- b. Me domandi *quando tut el laor el sarà fat*

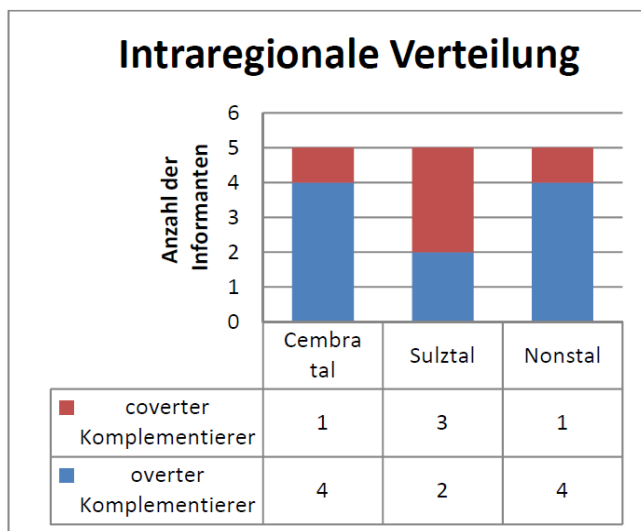
Il terzo fattore rilevante è il *wh-* stesso, che richiede, permette o proibisce la presenza del *che*.

- (9) a. Voi saver ndo che i è nadi. Trento
- b. No so cosa l'abia comprà per zena.

Visto che alcuni wh sono compatibili con *che* più facilmente di altri, è possibile costruire una scala implicazionale tra elementi wh tenendo fermi gli altri due fattori (il fatto che la frase sia una secondaria e la posizione del soggetto). La scala rilevata proviene da un'indagine condotta in loco da Polonio (2014) per mezzo di un test di elicitazione di frasi dialettali tramite un contesto presentato nella varietà dialettale investigata da un collaboratore madrelingua della varietà. L'inchiesta è stata condotta con cinque parlanti per ognuna delle tre località investigate nelle tre valli (Segonzano, S. Bernardo e Cles)³³. Le scale implicazionali di wh-che si ricavano dai dati sono le seguenti:

- (10) a. Dove → Chi → Quando
 b. Quanti → Quanti X
 c. Come → Perché

I contesti utilizzati nel test sono stati bilanciati per verificare se la presenza del complementatore fosse determinata dall'interpretazione del tipo di interrogativa indiretta, come ad esempio le interrogative retoriche. L'indagine ha potuto escludere che i fattori rilevanti fossero: a) d-linking dell'elemento wh-; b) la distinzione tra pronomi e aggettivi interrogativi; c) la distinzione tra argomenti e aggiunti. Un'analisi statistica dei dati presentata in Polonio (2014), di cui riporto qui la tabella 2 (pagina 29) non ha portato ad un quadro preciso della distribuzione, come si può notare nel seguente grafico, la distribuzione del complementatore appare opzionale in tutte le tre valli, seppur con frequenze diverse.



Siamo davanti ad un caso in cui la variazione statistica presa nel suo complesso non ci dice niente sulle sue cause. Quindi in questo caso un'analisi puramente statistica ancora non ci illumina su cosa stia succedendo veramente. Se invece proviamo a fare un'analisi qualitativa, e osserviamo nel dettaglio le forme dei pronomi wh- vediamo che la loro distribuzione ci suggerisce un'ipotesi precisa del percorso di estensione del complementatore. Si osservi che nel caso dell'elemento corrispondente a "che cosa" la realizzazione di *che* dipende dalla forma lessicale del pronome interrogativo, che può avere tre forme: *kwel*, che etimologicamente corrisponde al dimostrativo distale,³ *kosa*, che etimologicamente è un

³ Per un'analisi dettagliata dello sviluppo dei dimostrativi come elementi wh- si veda Munaro (2001).

‘nome leggero/classificatore che si accompagna al wh vero e proprio e si ritrova in varietà italiane anche con altri quantificatori (si veda ad esempio il siciliano *tutticosi tutto*’), e *ke*, cioè la forma che etimologicamente è un puro wh-⁴. Se scindiamo le tre forme otteniamo un quadro netto: il complementatore *che* è sempre obbligatoriamente presente nel caso in cui il pronome sia in origine un dimostrativo. Sia nell’indagine sul campo di Polonio (2014) che nel data base ASIt non sono attestati casi di *kwel* senza *che*. Possiamo quindi formulare la seguente generalizzazione descrittiva:

- (11) Se un dialetto usa una forma del dimostrativo distale come pronome interrogativo, *che* compare obbligatoriamente

Riporto qui di seguito tre esempi dai tre punti di inchiesta delle tre valli considerate:

- (12) a. Voi saver kel che fa Giovanni tut el dì ay (Segonzano)
b. Carmen l’ha domandà kel che venen far (San Bernardo)
c. La me ha domandà kel che fan le me sorele (Cles)

Il fatto che sia proprio il dimostrativo a cooccorrere obbligatoriamente con il complementatore ci induce a pensare che queste costruzioni abbiano in origine la stessa struttura delle relative libere e in particolare la struttura di una relativa libera con una testa cosiddetta leggera, cioè di tipo pronominale, che in molte lingue romanze è obbligatoria a differenza dalle lingue germaniche. Già Parry (2003) aveva notato che il complementatore compare nella storia del Piemontese primariamente in frasi relative libere e poi si estende alle interrogative.

L’ipotesi che queste strutture siano in origine relative libere spiega:

A) perché il fenomeno comincia nelle frasi secondarie e non nelle principali, dato che le relative libere sono solo strutture secondarie, non ci aspettiamo che il fenomeno cominci nelle frasi principali.

B) Perché i soggetti postverbalmente sono preferiti. Le frasi relative hanno la tendenza ad avere il soggetto postverbale nelle varietà italiane, e in alcuni casi la posposizione del soggetto è obbligatoria.

C) Il fatto che strutture con il *che* comincino proprio con i pronomi che possono introdurre relative libere, cioè elementi come *dove*, *come* ecc. mentre sono molto più rare con elementi wh che sono più difficilmente inseribili in una relativa libera come *perché* o *quanti + N*.

Una analisi dettagliata della distribuzione del complementatore nelle interrogative indirette in tre valli trentine ci mostra quindi che l’originaria indicazione etimologica dalla quale siamo partiti, cioè che il complementatore nelle lingue romanze è etimologicamente diverso dal corrispondente elemento delle lingue germaniche è in effetti corretta. Quello che vediamo in alemannico e dialetti italiani settentrionali non è lo stesso fenomeno, visto che nelle varietà alto-italiane parte probabilmente da una struttura simile alle frasi relative libere e con i wh- che possono trovarsi in questi contesti, mentre nelle varietà alemanniche il contesto da cui parte la struttura sono i wh-complessi.

Possiamo inoltre ipotizzare che probabilmente il secondo stadio in cui il fenomeno si estende ai wh che non compaiono frequentemente in frasi relative libere (come ad esempio *quanti*) corrisponda al momento in cui viene rianalizzato come una sorta di raddoppiamento interno del wh (del tipo [cosa che]), e che nel suo ultimo stadio di estensione, come ad esempio nei dialetti veneti, diventa una sorta

⁴ E possibile che in realtà l’elemento wh sia solo *k-*, e che la *e* sia un word marker del tipo che si osserva in nomi della classe di *volpe*, *amante*, ecc. Non credo che la *e* sia solo una vocale epentetica, perché questa in genere in italiano è una /i/.

di raddoppiamento di tutti i pronomi wh. Casi di raddoppiamento dei pronomi wh- esistono anche nelle frasi interrogative principali nei dialetti alto-italiani. Riporto qui alcuni casi come esempio di una strategia che va ulteriormente investigata, sia per capire come si estenda la struttura con il complementatore dalla sua originaria struttura analoga ad una relativa libera come abbiamo individuato qui, sia per capire in generale come analizzare il fenomeno del raddoppiamento degli elementi wh:

- (13) a. Che 'ncontre-t chi? Malonno
 Wh meet-you who
 'Whom are you meeting?'
- b. Che mànge-t 'nséma chi?
 Wh eat-you with who
 Whom are you going to eat with?
- c. Che l'è-t fat con che ròba?
 Wh it have-you done with which thing
 'With what did you do it?'
- d. Ch' ö-t qual?
 Wh want-you which
 'Which one do you want?'

Il fatto che il fenomeno del raddoppiamento *wh-che* esista indipendentemente anche nelle interrogative dirette nei dialetti italiani rafforza l'idea che la diffusione della struttura wh-che possa passare attraverso uno stadio in cui l'originaria struttura di una relativa libera venga rianalizzata come un caso di reduplicazione interna al wh-. Possiamo quindi concludere questa prima parte notando che non sempre una semplice valutazione statistica della distribuzione del fenomeno in esame ci illumina sulla via da percorrere, mentre invece considerando l'origine etimologica del complementatore da un lato e del pronome wh-dall' altro, la soluzione ci appare in maniera molto più chiara.

3. Un caso di analisi quantitativa

Nel paragrafo precedente ho mostrato un caso in cui una valutazione quantitativa non si rileva lo strumento corretto per identificare attraverso quali stadi sia passata una certa costruzione sintattica. Tuttavia, un'analisi quantitativa, anche in termini apparentemente molto rozzi, può essere molto rivelatrice in particolare per la costituzione in termini di tratti degli elementi funzionali. Se si adotta anche per gli elementi funzionali la strategia tradizionale usata dai dialettologi per costruire le varie etimologie per un unico nome (si vedano ad esempio i bei lavori di Pellegrini sul ladino dolomitico e sulle varietà venete)⁵ che permette di tipizzare le varianti etimologiche di un lessema identificando delle strategie usate nei vari dialetti. Se ad esempio si fa una tipizzazione dei morfemi di negazione nelle varietà dialettali italiane si giunge ad una lista di possibilità che assomiglia molto a quella già riscontrata dai tipologi (vedi van der Auwera (2011)) e che contiene i seguenti tipi: a) un elemento che indica una minima quantità (veneto *miga/mina*, emiliano *brisa*, fiorentino *punto*, lombardo alpino (*buc*)*ca* ecc.) b) l'elemento corrispondente al pronome indefinito *niente* (piemontese *nen*, provenzale *res*) c) l'elemento che viene usato come negazione pro-frase, trentino e lombardo /no/ d) un elemento che corrisponde ad

⁵ Cito qui solamente due tra i tanti volumi di studi e saggi di Pellegrini, (1977) e (1999).1

un avverbio correlato ad una forma del verbo *mancare*, come ad esempio nel lucano *manco* e) una forma *neca* (siciliano) che deriva con tutta probabilità da una frase scissa (*n*)*un-è-ca*, ‘non è che’, che ormai funziona come un avverbio. Ci si può chiedere come mai proprio questi elementi possano andare a sostituire tramite il ciclo di Jespersen l’originaria negazione *non* derivata dal latino e non altri: ad esempio non si trovano casi di negazioni frasali che derivino dall’elemento corrispondente a *nessuno*, quindi *niente* deve avere qualcosa di speciale rispetto ad altri pronomi/aggettivi indefiniti. Credo che tipizzare etimologicamente elementi funzionali come la negazione ci possa rivelare molto delle operazioni semantiche che vengono adottate dai parlanti nel negare una frase. In Poletto (2017) ho avanzato l’ipotesi che l’operazione di negazione nelle lingue naturali non corrisponda a quella della logica formale, cioè che non si abbia un operatore unico come $\neg P$ nelle lingue naturali. Se osserviamo le lingue del mondo, solo una esigua minoranza mostra la negazione davanti a tutta la frase, il cosiddetto principio del Negation-first espresso da Corblin e Tovenà (2003) non sembra applicarsi nella maggioranza delle lingue. Lingue che hanno la negazione nella periferia sinistra della frase la mostrano in genere nelle frasi secondarie, per cui si osservano complementatori negativi sul genere del latino *ne*, ma nelle frasi principali la negazione non compare davanti a tutta la frase. Nemmeno l’italiano ha una struttura di questo tipo, visto che la negazione di frase compare prima del verbo ma sempre dopo il soggetto, i pronomi interrogativi, gli elementi dislocati, focalizzati e tutta una serie di avverbi modali. Osserviamo invece che in molte lingue per negare il predicato si nega in genere l’oggetto (si pensi all’articolo negativo *kein-* del tedesco ad esempio) o si hanno comunque negazioni di frase che compaiono molto in basso nell’albero strutturale come avverbi nell’ambito strutturale dove in genere si trovano avverbi di tipo aspettuale o di maniera. Il motivo per cui credo che la negazione di frase possa essere rappresentata dal pronome indefinito *niente* ma non da elementi quali *nessuno* o *mai* è dovuta probabilmente proprio al fatto che *niente* rappresenta nella stragrande maggioranza dei casi proprio un oggetto che viene negato. Molti linguisti, notando questa discrepanza tra la logica formale e le lingue naturali hanno ipotizzato che in molte lingue ci sia un operatore nullo collocato all’inizio della frase che corrisponde a quello della logica formale. Credo invece che si possa spiegare questo fatto in una maniera alternativa, e cioè non dando per scontato che le lingue naturali debbano necessariamente realizzare formule della logica formale. Supponiamo invece che quando si nega una frase non abbiamo una formula come $\neg P$, ma una serie complessa di operazioni, ognuna delle quali può essere derivata sulla base degli elementi che servono ad esprimerla. Ad esempio, se molte lingue usano degli elementi che indicano in origine una piccola quantità (il fenomeno viene riportato da van der Auwera (2011) anche per le lingue bantu, quindi non si tratta di una peculiarità delle lingue romanze) è possibile ipotizzare che una delle operazioni interpretative che servono a negare una frase abbia a che fare con un elemento scalare. Il fatto che si utilizzi in varie lingue un elemento derivato dalla struttura copulare di una frase scissa, che viene utilizzata in genere anche per marcare la focalizzazione di un costituente⁶ indica probabilmente che per negare una frase, una delle operazioni semantiche che vengono fatte è la stessa che viene fatta anche nella focalizzazione, cioè probabilmente l’apertura di un set chiuso di elementi e la selezione di uno di questi elementi rispetto agli altri, su cui poi viene applicato il predicato nel caso del focus ed invece viene escluso dal predicato nel caso della negazione, e questo valore di esclusione dovrebbe essere quello realizzato morfologicamente da elementi come *manco*. Non intendo dare qui una analisi dettagliata di tutte le possibili operazioni semantiche della negazione frasale (si veda a questo proposito Poletto (2017)), ma semplicemente illustrare come una analisi a largo spettro delle

⁶ il caso che ho riportato per il siciliano *neca* si osserva anche in aramaico, quindi anche qui non siamo di fronte ad una peculiarità dei dialetti italo-romanzi

basi etimologiche di morfemi funzionali possa essere d'aiuto nel capire quali operazioni semantiche stiano alla base di fenomeni universali, quali la negazione di frase, i pronomi interrogativi, vari tipi di quantificatori ecc.. Quindi, utilizzare il metodo della tipizzazione tipico della dialettologia tradizionale può rivelare insospettite prospettive sulle operazioni che il sistema computazionale della lingua compie. Un altro tipo di analisi quantitativa che si può svolgere sugli elementi funzionali sempre utilizzando i tipi etimologici consiste nel comparare quanti possibili tipi di morfemi vengono utilizzati in un determinato ambito linguistici microvariazionale per esprimere un determinato valore funzionale. Porterò qui due esempi che mostrano come l'indice di variazione lessicale di elementi funzionali come quantificatori e pronomi interrogativi vada di pari passo con la complessità strutturale e semantica che questi mostrano. Prendiamo prima il caso dei quantificatori: nell'ambito del sistema dei quantificatori vediamo che quello che esprime il valore universale risulta estremamente stabile: la carta 664 dell' AIS "sedetevi tutti quanti" riporta il caso dei quantificatori *tutti* che risulta etimologicamente estremamente stabile nelle varietà italo-romanze. Non ho potuto trovare alcun punto AIS in cui la forma di questo quantificatore non fosse riportata all'etimo *tutti* tramite l'applicazione di regole fonologiche che interessano la vocale /u/, ad esempio dittongazioni nell' area pugliese, o la geminata /t:/ che viene palatalizzata in molte varietà, molto spesso al nord il morfema di numero /i/ cade, come cadono in genere tutte le vocali finali diverse da /a/ ecc. Non sembrano esserci casi di una radice etimologica diversa, quindi possiamo concludere che il quantificatore universale ha un grado di uniformità lessicale massimo. Prendiamo invece in esame il quantificatore corrispondente a *molto*: si vede immediatamente da una ricerca nel data base AISIt che gli elementi che esprimono questo quantificatore sono di varia natura: presento qui di seguito una lista non esaustiva che serve ad esemplificare il grado di variazione che si riscontra nelle varietà italo-romanze: Il primo tipo etimologico che troviamo corrisponde a *tanto* e ai suoi composti: Alassio (Liguria) *tantu*, Andreis (Friuli) *mitant*, Altare (Liguria) *atant*, Cimolais (Friulia) *betant*, Erto (Friulia) *matant*. Il secondo tipo è formato dall'avverbio *bene* e dai suoi composti: Rodoretto (Provenzale) *bien*, Livorno Ferraris (Piemonte) *motoben*, Poirino (Piemonte) *motobin*, Vallecrosia (Liguria) *benben* Sondalo, Grosio (Lombardia) *begè*. Il terzo tipo corrisponde ad un articolo indefinito seguito da un nome di quantità di vario genere, spesso non numerabile: Albosaggia Bormio (Lombard) *un bel po'*, Lecco (Lombard) *un bel pitt*, Chioggia (Veneto) *mondo*, Carpi (Emilian) *dimondi*, Forlì (Emilian) *namasa*, Cesena (Emilian) *una masa*, S. Michela (Friulian) *una vora*. Un quarto tipo meno comune corrisponde ad aggettivi che hanno anche il valore di 'troppo': *assé* (very), Ferrara (Romagna) *puras*, Corvara (Ladin) *trop*, Monno (Eastern Lombard). Il quinto tipo è formato da aggettivi che indicano probabilmente una intensità come: *fes*, La Spezia (Liguria) *gran* Casarza (Liguria) *dubelu*. Credo che questa molteplicità di etimi sia dovuta al fatto che un avverbio come *molto* è semanticamente più complesso di un quantificatore universale: l'operazione semantica che ci serve per interpretare un quantificatore universale è relativamente semplice, si tratta a) dell'apertura di un set e poi b) di considerare il set nel suo insieme o eventualmente nella sua valenza distributiva, cioè prendendo sempre come base l'insieme ma visto dalla prospettiva della sua interezza o considerando ogni singolo elemento (quindi in una lettura analoga al quantificatore *ogni*). Se vogliamo invece interpretare un quantificatore come *molto* dobbiamo non solo avere l'operazione di a) apertura di un set, probabilmente lessicalizzato da elementi tipo *mondo*, *massa*, del tipo tre, che si ritrova peraltro anche nel quantificatore universale *tout le monde* del francese, ma in aggiunta non viene considerato tutto l'insieme, ma dobbiamo anche b) fare una partizione nel set e in aggiunta c) il set deve essere inserito su una scala e d) il valore considerato deve essere valutato come alto rispetto ad un valore medio preso a modello. Avremo quindi non solo etimi che corrispondono all'apertura del set, ma anche etimi che corrispondono ad aggettivi intensificatori

come grande, bello⁷, ecc. che indicano il valore alto su una scala. Il fatto che in alcune varietà l'elemento *molto* sia lo stesso di *troppo* mostra che entrambi hanno in comune il fatto che il valore deve essere al limite estremo in alto sulla scala di valutazione. Lo stesso vale per etimi del tipo di *bene*, anche raddoppiato in *benben* per esprimere il valore di intensità sulla scala. Si vede dunque che la tipizzazione etimologica di un ambito dialettale coerente può dare indicazioni preziose riguardo al tipo di operazioni semantiche, che in ultima analisi determinano anche i tratti formali sintattici che a loro volta innescano operazioni di movimento, accordo ecc.. Dobbiamo però fare attenzione al modo in cui si tipizza la carta, cioè a non far semplicemente corrispondere un tipo di etimo ad una operazione semantica. Questo tipo di ragionamento sarebbe troppo semplicistico: ad esempio elementi come *ben* o come *bello, grande* ecc. sono etimi decisamente diversi, ma esprimono lo stesso valore semantico di intensità alta su una scala. La tipizzazione di elementi funzionali va fatta quindi cercando di non aumentare in maniera esponenziale i tratti semantici che questi elementi rappresentano. Un caso eclatante di questo tipo potrebbe essere quello del pronome *wh* corrispondente a *dove*, che ha solo nell'Italia settentrionale sulla base della carta dell' AIS 821 volume IV «Dove vai?» ben 56 forme diverse, alcune delle quali sono evidentemente riportabili a mutamenti fonologici regolari in quest'area quali appunto la caduta della vocale finale diversa da /a/, o inserzioni di vocali (soprattutto in area emiliana, dove sappiamo essere frequente il fenomeno di prostesi ed epentesi), ma nella maggioranza dei casi all'interno dell'elemento *dove* si possono individuare vari formativi morfologici, che riporto qui di seguito. Il primo formativo corrisponde alla preposizione *in*, in forme come le seguenti *endo* (Aldeno), *indulà* (Aquileia), *indunda* (Borgomanero), *indove* (Venezia), *engleca* (Albosaggia). Il secondo formativo è anch'esso preposizionale, e corrisponde alla preposizione *da* (cioè probabilmente *di+a*, che può essere scisso per cui compare a volte solo la *d-* a volte solo la *a-*), che si osserva in forme come l'italiano antico *donde*, il friulano *daulà* (Barcis) *andé* (Cirvoi) *anté* (Poirino). Il terzo morfema è un formativo *o/u* che corrisponde forse alla *u* del latino *ubi* (cfr. *ulà* (Andreis) *dulà* (S. Michele), *indulà* (Aquileia), *undi* (Cairo Montenotte), *onde* (Calalzo). Un altro formativo chiaramente identificabile corrisponde ad un pronome locativo distale *là* e si riscontra soprattutto nelle varietà friulane e ladino centrali, si veda *ulà* (Andreis) *dulà* (S. Michele, Claut) *Indulà* Aquileia, *Olache* (Campitello di Fassa). Parecchie varietà mostrano anche una sorta di formativo *n/d* o *nd* come in *ndonde* (Novi Ligure) *undi* (Cairo Montenotte) *onde* (Calalzo). E interessante notare che facendo una tabella delle forme, mentre i vari formativi cooccorrono tra di loro in maniera libera o possono mancare per cui abbiamo coppie minime del tipo *dulà/ulà*, in cui compare o meno il formativo preposizionale *d-* a inizio parola, oppure casi come *indove/dove* in cui compare o meno il formativo preposizionale *in-* non abbiamo nessun caso in cui i formativi *là* e *n/d* occorrono, il che ci potrebbe suggerire che entrambi esprimano lo stesso valore di deissi locativa che evidentemente deve essere semanticamente presente nel *wh*-locativo. La distribuzione complementare tra formativi all'interno della parola funzionale in un ambito dialettale ci fornisce quindi una prima indicazione della tipizzazione stessa, cioè di quali formativi esprimono lo stesso valore semantico, il che ci permette di non moltiplicare i formativi delle parole funzionali e quindi mantenere bassi i loro valori semantici. In Munaro e Poletto (2014) viene proposta una lista esaustiva dei formativi che compongono l'elemento *dove* e dei loro valori semantici, ma non è questo lo scopo di questo lavoro, che invece intende mostrare come una metodologia tradizionale possa essere applicata al contenuto degli elementi funzionali per

⁷ Si noti che in tante varietà alto italiane l'elemento bello ha assunto il valore aspettuale dell'avverbio *già*, si veda il trentino *bel*, ladino *belche* o il veneto *belo che*. Questo dimostra che l'aggettivo bello può essere considerato puramente nel suo valore di elemento massimo su una scala senza necessariamente portare con sé anche il valore lessicale della scala a cui è in genere associato in italiano.

determinare in ultima analisi i valori di base dei tratti formali, che corrispondono ad operazioni semantiche che sono già stati determinati in fonologia da lungo tempo, ma che per quanto riguarda la sintassi formale sono ancora in larga parte ignoti. Quello che in sintassi vengono oggi considerati come tratti, quali ad esempio il tratto di [persona], sono chiaramente un complesso di tratti semantici più semplici che comprendono, ad esempio nel caso della persona, come minimo il parlante, l'ascoltatore e il numero. Osservare i formativi degli elementi funzionali ci può aiutare a, diciamo così, fare la lista delle operazioni semantiche basilari che dobbiamo postulare e che si riflettono nelle operazioni sintattiche formali, che spesso considerano blocchi di operazioni come unità allo stesso modo in cui i tratti fonologici vengono espressi in fasci nei fonemi. Ad esempio nel caso dei formativi dell'elemento *wh-dove*, il fatto che ci siano con tutta probabilità tre formativi preposizionali (*in, di, a*) ci potrebbe pensare ad ipotizzare che la semantica vettoriale delle preposizioni come concepita ad esempio da Svenonius (2010) o Cinque (2010) sia un elemento contenuto anche nella semantica del pronome *wh*-corrispondente a *dove*, il quale deve avere anche un locativo deittico al suo interno. Elementi così sfuggenti come i pronomi, che per loro natura sono formati da un'unica parola e elementi funzionali come la negazione, possono essere investigati nei loro tratti costitutivi solo tramite una metodologia come quella che ho delineato qui.

Conclusioni

In questo lavoro ho cercato di mostrare come una prospettiva di indagine qualitativa sia a volte più adatta a capire fenomeni sintattici complessi (quali ad esempio l'estensione della struttura *wh*-complementatore in lingue diverse) di una prospettiva quantitativa. D'altro canto, una prospettiva quantitativa ci fornisce un punto prospettico irrinunciabile in alcuni ambiti, come ad esempio quello della determinazione dei tratti formali degli elementi funzionali, che in ultima analisi derivano da operazioni semantiche/cognitive che vengono espresse tramite la scelta di elementi lessicali il cui valore segnala queste operazioni cognitive. Utilizzare quindi una prospettiva etimologica o meglio di tipizzazione di morfemi o formativi del lessico funzionale ci aiuta a identificare queste operazioni basilari, senza le quali una teoria generale dei tratti formali in sintassi non è possibile, e in effetti ancora manca a questo stadio della ricerca. Se si compara il punto a cui è già arrivata la fonologia in questo ambito ci si rende conto che in ambito sintattico ci sia ancora moltissimo lavoro da fare, che probabilmente non è stato fatto finora proprio per mancanza di strumenti di indagine adeguati a identificare con precisione operazioni così elusive, dato che per il componente cognitivo non possiamo usare l'osservazione diretta come può fare la fonologia con la fonetica sperimentale. Qui ho cercato di mostrare che recuperare i "big data" che la dialettologia tradizionale ci offre può essere un possibile metodo proprio per aggirare questo ostacolo e pervenire finalmente ad una teoria dei tratti formali della sintassi sulla quale basare le nostre analisi empiriche.

Riferimenti bibliografici

AIS Sprach-und Sachatlas Italiens und der Südschweiz, K. Jaberg and J. Jud, Zöfingen, 1928-1940.

ASIT (Atlante Sintattico d'Italia) <http://asis-cnr.unipd.it/>

Baker (1997) Baker, M. (1997). *The atoms of languages*. New York: Basic Books.

Bayer, J. & E. Brandner (2008a), "On Wh-Head-Movement and the Doubly-Filled-Comp Filter". In *Proceedings of the 26th West Coast Conference on Formal Linguistics*. A cura di Chang C. B. & H. J. Haynie. Somerville: Cascadilla Proceedings Project, 87-95.

Benincà (1996) Agglutination and Inflection in Northern Italian Dialects, in *Aspects of Romance Linguistics*, Selected Papers from the Linguistic Symposium on Romance Languages XXIV (10-13 marzo 1994), a cura di C. Parodi, C. Quicoli, M. Saltarelli, M.L. Zubizzareta. Washington, Georgetown University Press, pp. 59-72.

Cinque, G. (2010) 'Mapping spatial PPs: an introduction'. In *Mapping spatial PPs*, a cura di G. Cinque and L. Rizzi. Oxford: Oxford University Press. 3-25.

Cinque, G. 2016. A microparametric approach to word order typology. Di prossima pubblicazione in *Parameters: what are they? Where are they?* a cura di S.Karimi and M.Piattelli Palmarini. Special issue of *Linguistic Analysis*.

Corblin, F. & L. Tovenà (2003) L'expression de la négation dans les langues romanes. In *Les langues romanes: problèmes de la phrase simple*. A cura di D. Godard. Paris: CNRS Edition 279-34

Egerland, Verner. 1996. *The Syntax of Past Participles*. A Generative Study of Nonfinite Constructions in Ancient and Modern Italian. Lund: Lund University Press

Garzonio (2007) Complementatori nelle interrogative delle varietà trentine: variazione diatopica e generazionale. In *Dialecto, Memoria e Fantasia*. Atti del Convegno, Sappada 28 giugno-2 luglio 2006. A cura di G. Marcato. Padova:Unipress. 179-183.

Munaro, N. (2001) 'Free relatives as defective wh-elements'. In *Romance languages and linguistic theory*. a cura di Yves D'Hulst et alii. Amsterdam: John Benjamins. 281-306.

Munaro and Poletto (2014) "Synchronic and diachronic clues on the internal structure of 'where' in Italo-Romance", in *Festschrift in honor of M Parry*. a cura di P. Benincà et al., Oxford & New York: Oxford University Press, 2014, pp. 279-300.

Parry, M, 2003, 'Cosa ch'a l'é sta storia? The interaction of pragmatics and syntax in the development of WH-interrogatives with overt complementizer in Piedmontese'. a cura di Ch. Tortora. *The Syntax of Italian Dialects*. OUP, pp. 152 – 174

Pellegrini (1977) *Studi di dialettologia e filologia veneta*, Pisa, Pacini, 1977

Pellegrini (1999) *Comparazioni lessicali "retoromanze"*. *Complemento ai "Saggi ladini"* di G. I. Ascoli, Venezia, Istituto Veneto di scienze Lettere ed Arti.

Poletto, C. (2000) *The higher functional field: evidence from Northern Italian Dialects*. Oxford: Oxford University Press.

Poletto C. (2012). Comparative Linguistics and Microvariation: the role of Dialectology. *Language in Context*. 12.1, 47-68.

Poletto (2014) *Word Order in Old Italian*, Oxford University Press, Oxford & New York, 2014.

Poletto (2017) Negative doubling: in favor of a Bing NegP analysis." in *Studies on Negation: Syntax, Semantics, and Variation* [Wiener Arbeiten zur Linguistik], a cura di Cruschina, Silvio, Katharina Hartmann & Eva-Maria Remberger (in press). Göttingen: V&R unipress / Vienna University Press.

Poletto and Vanelli (1993) "Gli introduttori delle frasi interrogative nei dialetti italiani", in *Atti del Convegno Italia Settentrionale: Crocevia di Idiomi Romanzi*, a cura di E. Banfi, G. Bonfadini e P. Cordin, Niemeyer, Tübingen, 1993, pp. 145-158.

Polonio, K. (2014) Die komplexe Struktur einer Interrogativstrategie am Beispiel der Trienter Varietäten. Tesi di Laurea Goethe Universität Frankfurt

Roberts, Ian (2012) 'On the nature of syntactic parameters: a programme for research', In *Parameter Theory and Linguistic Change*. C. Galves, S. Cyrino, R. Lopez & J. Avelar. Oxford: OUP, 319-334.

Salvioni, Carlo. 1886. "Saggi Intorno ai Dialetti di Alcune Vallate dell'Estremità Settentrionale del Lago Maggiore". *Archivio Glottologico Italiano* 9.188-260.

Svenonius, P. (2010) 'Spatial P in English', in G. Cinque and L. Rizzi, (eds.) *Mapping spatial PPs*, Oxford: Oxford University Press. 127-160.

van der Auwera, J. (2010) On the diachrony of negation. In *The Expression of Negation*. a cura di Horn, L. Berlin: Mouton De Gruyter, pp. 73-109.