

L'AMBIGUITÀ SINTATTICO-SEMANTICA: INDAGINE SPERIMENTALE SUL RUOLO DELLA PROSODIA IN PRODUZIONE E IN PERCEZIONE

SILVIA SPERTI

1. Presupposti teorici

Nell'ambito di questa ricerca rientrano considerazioni su aspetti prosodici che, riguardando segmenti superiori al fono, sono strettamente collegati a fattori studiati da altre discipline quali la psicolinguistica e la psicologia del linguaggio. L'attuale ricerca si occupa, infatti, dei meccanismi e delle strategie adoperate dal parlante nella produzione di frasi potenzialmente ambigue.

L'ambiguità è un elemento tipico della lingua e non sempre passa inosservato, soprattutto nei casi in cui si pone ai limiti dell'incomunicabilità. Alla disambiguazione prendono parte elementi linguistici e soprasegmentali ma soprattutto situazionali, di contesto e di referenza. La letteratura sull'ambiguità e la disambiguazione è molto vasta ed ha investito ormai da decenni campi diversi, con studi largamente interdisciplinari. La ricerca è stata prevalentemente condotta sull'inglese, ma vi sono numerosi studi anche sul tedesco e le lingue germaniche, sulle lingue romanze (pochi sull'italiano) e sul giapponese (cfr., fra gli altri, Frenck-Mestre [1997], Mazuka *et al.* [1997] e Zagar *et al.* [1997]).

La prosodia è un fenomeno indagato dalla scienza linguistica solo da pochi anni in quanto difficile da studiare per molti motivi. Il primo è certamente il fatto che tutte le variabili fisiche che la determinano (il tempo e l'intensità per il ritmo, la frequenza per l'intonazione) si dispongono lungo un *continuum* non segmentato, come è invece, grazie alla scrittura, nel caso dei tratti segmentali, quali fonologia, morfologia, lessico e, in parte, sintassi (cfr. Bertinetto [1981], D'Imperio e Gili Fivela [2004], Gili Fivela [1999], Watson e Gibson [2004]).

In secondo luogo, un segmento intonativo di prosodia ritagliato dal suo contesto è in sé totalmente privo di significato e di funzione, perché i valori che determinano la prosodia, e dunque anche l'intonazione, sono sempre e tutti radicalmente relativi, valutabili e interpretabili solo in rapporto a ciò che segue e a ciò che precede all'interno dell'intera unità prosodica considerata. Quindi, di un segmento vocalico in sé, del quale possiamo misurare l'altezza, la durata e l'intensità (oltre che il timbro), non possiamo dire se sia linguisticamente acuto o grave, lungo o breve, intenso o debole (cfr. Bertinetto e Magno Caldognetto [1993]).

Inoltre, la natura particolare della variabilità prosodica, anche a parità di variabili diastratiche, diatopiche e diafasiche, è sempre il riflesso di una differenza nelle intenzioni comunicative del parlante e nella sua stessa natura comunicativa: ad una variazione prosodica corrisponde sempre una variazione pragmatica dell'enunciato. Questa importantissima componente della comunicazione orale si realizza attraverso un meccanismo articolatorio estremamente economico, basato su una variazione di poche decine di Hz rispetto alla frequenza fondamentale propria di ciascuno, di pochi millesecodi o decibel nella durata e nell'intensità dei segmenti, e di piccole variazioni nella pressione dell'aria nel canale epilaringeo. L'importanza della componente prosodica

si spiega, allora, nel ruolo indispensabile che essa svolge nel processo di analisi delle produzioni linguistiche in generale e dei casi di ambiguità in particolare.

L'*eye tracking*, metodo utilizzato in questa ricerca, permette allo sperimentatore di evitare analisi metalinguistiche e di monitorare il contesto visivo e comunicativo di un enunciato e l'atteggiamento del mittente/ricevente nei suoi confronti (cfr. Duchowski [2007]). Alla luce di questi vantaggi operativi esso può essere impiegato, come vedremo per il nostro caso, in compiti sperimentali interattivi e associati dando origine a paradigmi e metodi innovativi e di straordinario impatto scientifico. Il metodo di *eye tracking* ha il vantaggio di fornire un quadro sostanzialmente chiaro di ciò che avviene a livello cognitivo alla presentazione di uno stimolo visivo e di uno uditivo. Si tratta di uno strumento preciso in grado di registrare ogni variazione che avviene nel sistema attenzionale e trova un riflesso nei movimenti oculari.

È evidente il campo sempre mutevole e variegato in cui si muovono studi come questi. Non è possibile definire un unico e definitivo paradigma teorico, ma è auspicabile integrare e confrontare risultati e conclusioni generali a cui ogni lavoro giunge, creando in questo modo un terreno sempre fertile di confronto e di stimolo per ulteriori ricerche e ampliamenti sperimentali, nuove teorie e moduli di ricerca.

2. L'indagine sperimentale

L'attuale studio rappresenta l'unione di due indagini relative al ruolo svolto dalle informazioni segmentali e prosodiche nella produzione e nell'interpretazione di frasi potenzialmente ambigue da un punto di vista sintattico-semantic¹. La prima indagine rappresenta uno studio pilota, utilizzato come punto di partenza per una seconda fase di ricerca, su cui è incentrato il presente studio. In particolare, la ricerca è stata focalizzata sulla prosodia, elemento indispensabile e caratterizzante ogni enunciato, che svolge l'importante compito di personalizzare e rendere comprensibile il significato in qualsiasi processo comunicativo, attraverso la modifica di alcuni parametri, quali ritmo, intonazione, durate sillabiche, durate pausali e intensità.

I materiali strumentali all'indagine sui processi di produzione di frasi ambigue sono stati rappresentati dalle registrazioni di parlato effettuate su dieci parlanti salentini. Le frasi selezionate rappresentano due casi di ambiguità sintattico-semantic, in cui sono possibili tanto una doppia assegnazione logica dei costituenti quanto una doppia struttura sintattica con una conseguente polivalenza sul piano semantico. In questo tipo di frasi avviene, pertanto, una segmentazione dell'enunciato (*phrasing*) che si adatta alle esigenze interpretative, in coerenza con il processo di disambiguazione di volta in volta effettuato. Ai soggetti è stato richiesto di produrre cinque volte una serie di quattro enunciati ottenuti a partire da due frasi ambigue. Esse sono: "la vecchia legge la regola", che può essere interpretata in due modi, a seconda che si consideri come soggetto "vecchia" oppure "legge"; e "la vecchia porta la sbarra", in cui ancora il significato cambia in base all'assegnazione del ruolo di soggetto a "vecchia" o a "porta". Ai soggetti sono state proposte una frase scritta ed un'immagine che potesse rappresentare in modo semplice e immediato la realtà descritta in una delle due possibili interpretazioni delle frasi scelte (ad

¹ Cfr. Sperti, "Il ruolo della prosodia nei casi di ambiguità sintattico-semantic della frase: la produzione", tesi di laurea triennale in Fonetica Sperimentale non pubblicata, a.a. 2005/06 e Sperti, "Il ruolo della prosodia nei casi di ambiguità sintattico-semantic della frase: la percezione", tesi di laurea specialistica in Fonologia non pubblicata, a.a. 2007/08.

esempio alla frase “La vecchia legge la regola”, per “vecchia” soggetto è stata associata l’immagine di una donna anziana nell’atto di leggere un libro di grammatica), chiedendo loro di interpretare l’enunciato in base a tale rappresentazione grafica (Fig.1).

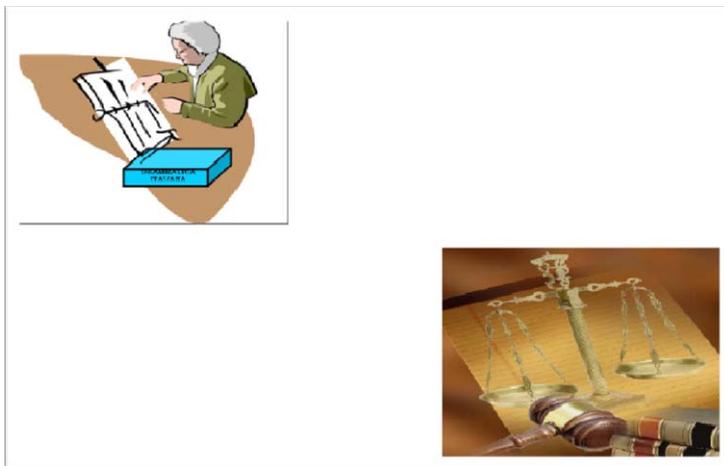


Figura 1. Immagini usate nei due esperimenti per le due interpretazioni di “La vecchia legge la regola”.

Le produzioni sono state registrate e trattate successivamente con un programma per l’analisi uditiva ed acustica (PRAAT, sviluppato da Boersma e Weenink [1996]).

La successiva analisi effettuata sulle registrazioni ha permesso di indagare da un punto di vista qualitativo e quantitativo il processo di interpretazione e di disambiguazione dei quattro possibili enunciati, con particolare attenzione agli strumenti prosodici impiegati. Molti dei parametri indagati sono stati analizzati distinguendo in base al genere dei parlanti, in modo da evidenziare eventuali differenze significative anche in rapporto a questo fattore

In particolare, i correlati presi in considerazione in quanto più significativi ai fini dell’indagine sono stati:

- i valori formantici di F1 ed F2 delle vocali medie di “vecchia” e “legge” (questa indagine, come vedremo più avanti, è stata condotta per verificare la presenza di differenze diatopiche fra i parlanti)
- le durate delle sillabe toniche, S1 ([vek]), S2 ([ledZ]) ed S3 ([re])
- le durate delle sillabe precedenti il confine sintattico, DopoS1 ([-kja]) e DopoS2 ([-dZe la])
- la *f0* finale della vocale al confine del sintagma nominale con quello verbale
- l’intensità media della vocale al confine del sintagma nominale con quello verbale
- posizione del *pitch* massimo nella sillaba tonica e in quella atona al confine di sintagma
- durata di pause al confine di sintagma e loro influenza sulle durate sillabiche.

Ciascuno di questi parametri ha dato spunto a riflessioni significative sul processo di disambiguazione di frasi polisemantiche e sull’impiego dei correlati prosodici in tale processo.

La variazione di tali parametri acustici è stata analizzata tenendo conto del cambiamento di ruolo sintattico e della variazione intersoggettiva.

È importante precisare, che le strategie di segmentazione dell'enunciato in fase di produzione sono state varie per ciascun parlante a seconda delle ripetizioni di una stessa frase. I soggetti hanno impiegato, infatti, strumenti prosodici differenti, quali durata delle sillabe toniche, contorni intonativi con variazioni di f_0 molto variegati, registri di frequenze differenti, variazioni di volume e, infine, presenza/assenza e durata di pause al confine di sintagma. Tutto ciò si evince dalle misurazioni effettuate a livello acustico sulla f_0 delle vocali toniche e di confine di sintagma, nonché sui valori di intensità delle vocali toniche. È possibile quindi affermare che l'interpretazione della frase con struttura SVO e sintagma nominale soggetto non ramificato risulti più semplice ed immediata (anche se non in modo generalizzato, come potrebbe risultare ovvio), giacché la struttura a dislocazione a sinistra o a destra con l'elemento anaforico o cataforico a seconda dei casi, presenta un grado di interpretazione e produzione più complessa e articolata, così come è possibile che le due immagini scelte per rappresentare le diverse interpretazioni non fossero ugualmente efficaci.

Le venti ripetizioni (ovvero cinque ripetizioni per quattro interpretazioni) dei dieci soggetti, come abbiamo detto, sono state analizzate tenendo conto di fattori prosodici, quali la durata delle sillabe toniche e della sillaba al confine, misurata in millesecodi (ms); la frequenza fondamentale iniziale e finale della vocale al confine sintattico tra sintagma nominale e sintagma verbale, nonché la sua intensità media misurata in decibel (dB) – indicante la variazione di volume – il valore in Hz di f_0 iniziale e finale delle vocali toniche, indicante l'andamento intonativo; la posizione del picco di f_0^2 massimo della vocale nelle sillabe toniche e nella vocale atona al confine; la durata delle pause, ove presenti.

L'analisi acustica effettuata con PRAAT sulle ripetizioni dei parlanti, quindi, è stata trattata in un secondo momento mediante elaborazione statistica. I dati raccolti a livello acustico (quali durata delle sillabe toniche e al confine, frequenza fondamentale e intensità media della vocale al confine, sia nelle ripetizioni con pausa che senza, qualità del *pitch* massimo nella sillaba al confine) sono stati successivamente rielaborati, verificando l'esistenza di differenze significative a seconda dell'interpretazione della frase, e quindi della funzione sintattica svolta da "vecchia" (e quindi di "legge", "porta", "la", "regola" e "sbarra"), e a seconda del parlante, al fine di evidenziare gli atteggiamenti del singolo parlante in rapporto al cambiamento di funzione sintattica.

In seguito ad un'analisi di tipo quantitativo si è osservato che per la durata della prima sillaba accentata (S1) nella frase La vecchia legge la regola, nei casi in cui non occorre la pausa, si registra una significatività statistica a seconda che "vecchia" sia o sostantivo o aggettivo. Emerge una situazione piuttosto omogenea anche intersoggettivamente: i dieci parlanti presentano durate sillabiche omogenee senza particolari eccezioni fra i parlanti maschi e per la maggior parte delle femmine (ad eccezione di due per le quali si può parlare di significatività, in quanto le due ragazze producono la prima sillaba con durate mediamente maggiori rispetto agli altri parlanti).

Ancora più interessante risulta il rapporto che emerge fra il parlante e la funzione sintattica di "vecchia": per 8 parlanti su 10 la prima sillaba di "vecchia" aggettivo è più breve di quella di "vecchia" sostantivo, con differenze in termini di durata anche notevoli in almeno 5 parlanti.

² Il picco di f_0 , qui indicato come *pitch* massimo, fa riferimento al picco di frequenza fondamentale raggiunto in un determinato contesto, il *pitch* minimo fornisce indicazioni sul registro di frequenza utilizzato dal parlante durante l'eloquio, la differenza tra il massimo e il minimo fornisce indicazioni, invece, sull'intervallo di frequenze impiegato (*pitch range*).

In generale, sono emerse importanti differenze fra i parlanti e le loro strategie prosodiche, fornendo ancora una volta rilevanti spunti di riflessione. La discriminazione va ovviamente effettuata tenendo conto del sesso dei parlanti e delle conseguenti differenze parametriche che ciò comporta, soprattutto per la frequenza fondamentale della voce (ciò nonostante, ai fini di una corretta interpretazione dei grafici proposti, è necessario notare che la realtà indagata è presentata in essi unitariamente senza differenziazioni di genere).

Anche per la durata della seconda sillaba accentata (S2) nella frase “La vecchia legge la regola”, nei casi in cui non occorre la pausa, si registra una significatività statistica per la durata della sillaba tonica di “legge” nei casi in cui “vecchia” sia o sostantivo o aggettivo. Altrettanto dinamico risulta il rapporto fra il parlante e la funzione sintattica di “vecchia”: per 6 parlanti su 10 la sillaba tonica di “legge” sostantivo è più lunga di quella di “legge” verbo, con differenze in termini di durata anche notevoli in almeno 5 parlanti. Emerge una situazione relativamente omogenea per i rapporti intersoggettivi in quanto emergono differenze in termini di durata per quattro parlanti maschi e due femmine su dieci.

Infine anche per la durata della terza sillaba accentata (S3) nella stessa frase, nei casi in cui non occorre la pausa, si registra una significatività statistica per la durata della sillaba tonica di “regola” nei casi in cui “vecchia” sia o sostantivo o aggettivo. Il rapporto fra il parlante e la funzione sintattica di “vecchia” è significativo nel determinare la durata della sillaba per 7 parlanti su 10 con una maggiore durata della sillaba tonica di “regola” verbo rispetto a quella di “regola” sostantivo, con differenze anche notevoli in almeno 5 parlanti. I dati variano molto da soggetto a soggetto in quanto emergono differenze significative in termini di durata per quattro parlanti maschi e quattro femmine su dieci.

In generale, si può osservare che fra le durate delle tre sillabe emerge una sostanziale differenza, in particolare fra la durata della prima e della seconda e la durata della terza, tenendo tuttavia conto del fatto che si tratta di una sillaba chiusa nei primi due casi [vek] e [ledZ]³ e di una aperta nel terzo [re]. Da segnalare, ancora, la differenza piuttosto rilevante fra la durata di S1 in “vecchia” aggettivo e quella di S1 in “vecchia” sostantivo.

A questo punto è opportuno fare delle considerazioni generali sull’indagine pilota che ha preceduto l’attuale, sui dati raccolti e sul loro trattamento, nonché sui i risultati che da tale indagine sono derivati, al fine di dare una rappresentazione più chiara possibile dello scenario procedurale, teorico e sperimentale da cui si è partiti per l’attuale lavoro.

Ricapitolando, dunque, l’attenzione si è focalizzata soprattutto su alcuni dei numerosi parametri acustici segmentali e soprasedimentali che caratterizzano la catena fonica durante la produzione. Le variazioni di frequenze, di durate sillabiche, di frequenza fondamentale e del suo correlato percettivo, l’intensità, della posizione del picco di frequenza nella sillaba e di durate prepausali, ove impiegate, analizzate in considerazione al cambiamento del ruolo sintattico provocato da un possibile doppio *phrasing* della frase, hanno dimostrato su un piano pratico quanto supposto da un punto di vista teorico.

In particolare da quanto osservato sui risultati ottenuti, si può affermare che esiste nella produzione di frasi potenzialmente ambigue una preminenza della categoria sintattica del soggetto logico tesa a disambiguare la polisemia. Tale preminenza si realizza attraverso durate sillabiche maggiori, possibile presenza di pause al confine sintagmatico fra soggetto e verbo, spostamenti “strategici” del *pitch* nelle sillabe atone e toniche del

³ In questo caso si utilizza la notazione SAMPA, che supplisce quella IPA nei casi di difficoltà legate alla codifica ASCII (*American Standard Code for Information Interchange*) di alcuni font.

soggetto (soprattutto in presenza di pausa), modificazioni dell'andamento melodico realizzato mediante variazioni di frequenza fondamentale ed intensità al confine.

Inoltre, si può notare la rilevanza di un'altra categoria sintattica strettamente collegata al sintagma nominale, il predicato verbale. Tale elemento della frase presenta una relativa prominenza, tradotta in termini acustici, nell'interpretazione della frase con prolessi dell'oggetto attraverso elemento clitico. Su questo piano di difficoltà maggiore dell'interpretazione, la prosodia risulta necessaria nel processo di disambiguazione, come dimostrato dalle modulazioni effettuate dal parlante su elementi sintattici della frase considerati strategici ai fini della comprensione e dell'interpretazione: il soggetto e il verbo, ed anche l'oggetto.

L'analisi ha di fatto evidenziato non solo una certa coerenza nell'andamento dei singoli parametri acustici e prosodici, in relazione al processo teso a disambiguare le due frasi, ma anche una realtà variegata di soluzioni possibili emerse dalle considerazioni relative alle strategie dei singoli individui o dei gruppi di parlanti (si vedano le differenze osservate in base al genere).

Dall'analisi dei dati raccolti risulta:

- la rilevanza del sintagma nominale con il ruolo di soggetto della frase durante la produzione, tradotto in termini acustici mediante un allungamento della sillaba tonica del sostantivo ed una durata relativamente minore della sillaba tonica del verbo e dell'aggettivo precedente.
- la preminenza del verbo nell'interpretazione della frase con elemento clitico rappresentato dal pronome "la", in termini di durata sillabica.
- l'allungamento della sillaba al confine fra sintagma nominale e verbale e di quella prepausale nei casi in cui occorre la pausa.
- la maggiore durata della pausa, ove presente, al confine in relazione all'assegnazione dei sintagmi e al grado di difficoltà interpretativa della frase.
- l'aumento della frequenza fondamentale e/o dell'intensità, in concomitanza con la presenza di confine sintattico, in particolare di confine con pausa.
- dal punto di vista qualitativo, la prevalenza di un picco di f_0 realizzato all'interno della sillaba al confine, con l'importante eccezione di alcuni casi in cui si registra o una anticipazione nella sillaba tonica o un ritardo sul finale di sillaba.

L'insieme di questi risultati ha evidenziato la validità dei presupposti teorici di base, in quanto ha dato conferma tanto della complessità dell'uso degli strumenti prosodici nel processo interpretativo, quanto delle variegate possibilità di combinazione degli stessi, tese alla sempre maggiore efficacia dello svolgimento dell'incarico comunicativo. Risultava, pertanto, auspicabile l'approfondimento di quanto osservato sul processo di disambiguazione, con lo scopo di mettere in evidenza sia le variazioni prosodiche più rilevanti in tale processo, sia l'esistenza di correlazioni sistematiche nelle modificazioni dei parametri già presi in considerazione nel primo lavoro.

3. Il metodo: strumenti e procedura

3.1. Prima fase: la produzione

L'indagine è stata condotta su dieci parlanti, cinque maschi e cinque femmine di età compresa tra i 20 e i 31 anni, tutti studenti universitari e residenti in otto diversi comuni della provincia di Lecce del Salento occidentale. Si è cercato di prendere in considerazione una porzione piuttosto omogenea di parlanti sia a livello diatopico che diastratico, in

quanto tutti e dieci i ragazzi hanno origini salentine, con genitori salentini e un'istruzione secondaria relativamente omogenea di tipo liceale.

Nella fase preliminare dell'esperimento sono state individuate due frasi potenzialmente ambigue la cui bivalenza sintattica può dare luogo a due diverse interpretazioni semantiche; esse sono: "la vecchia legge la regola" e "la vecchia porta la sbarra". Tali frasi, denominate *garden-path sentences* (cfr. Cacciari [2001]), esprimono livelli di ambiguità frasale in cui uno o più elementi sintattici possono avere due funzioni grammaticali e logiche, rispettivamente sia di sostantivo (soggetto) che di aggettivo (attributo); sia di sostantivo (soggetto) che di verbo; sia di articolo che di pronome (complemento oggetto).

Bisogna osservare che, una volta selezionate, le due frasi sono state associate ad una rappresentazione grafica mediante icone generalmente riconosciute come universali per quei significanti. Questo vuol dire che i parlanti sono stati guidati graficamente nella loro scelta e nel processo di disambiguazione almeno ad un primo livello. In altre parole, una volta decifrata la giusta assegnazione dei sintagmi, il parlante ha realizzato tale codifica nella produzione, impiegando quegli elementi prosodici che costituiscono il fulcro della nostra ricerca e che quindi in tal modo potevano essere osservati in maniera più diretta, in quanto in un primo momento i dieci parlanti partivano con lo stesso livello interpretativo fornito dall'immagine somministrata ed uguale per tutti e non selezionata, invece, da un patrimonio iconico e referenziale del tutto personale tipico di ognuno di noi.

Le icone che abbiamo scelto per ciascuna delle interpretazioni possibili sono: una donna anziana che legge un libro di grammatica; la bilancia e i codici giuridici in quanto icone universali del diritto e della legge occidentali; una donna anziana che trasporta una sbarra di ferro; una vecchia porta che sbarra l'accesso ad un percorso.

Le frasi corredate dalle suddette rappresentazioni a colori sono state raccolte in numero di cinque per ogni interpretazione, più un campione da sottoporre ai soggetti per prova iniziale, mescolate casualmente ed inserite singolarmente in un unico raccoglitore.

In seguito, sono state somministrate ai dieci parlanti, ai quali ne è stata richiesta la semplice lettura ad alta voce. Ai soggetti è stata raccomandata una produzione che fosse naturale, chiara e quanto meno condizionata da fattori esterni, quali tensione emotiva e disturbi sonori.

Singolarmente ogni parlante ha letto le 24 frasi, la cui produzione è stata registrata e campionata in laboratorio con l'ausilio dello specifico programma GoldWave (freq. 22 KHz). Durante la fase di produzione ogni parlante ha effettuato una discriminazione semantica delle quattro frasi, disambiguandone la potenziale polivalenza sintattica. Tale scelta è stata realizzata mediante espedienti soprasegmentali e prosodici molto differenti sia a livello intersoggettivo (tra parlanti o gruppi di parlanti) che ad un livello interno al parlante stesso (variazioni nelle cinque produzioni o in base all'interpretazione), risultando estremamente interessante e variegata e offrendo spunti di riflessione qualitativa su numerosi fattori, quali campi di esistenza delle vocali, contorni intonativi, picchi tonali, durata sillabica, pause.

Tali inconsapevoli scelte effettuate da ogni parlante ad ogni produzione hanno fornito dati importanti sul processo cognitivo che sottostà all'interpretazione semantica di frasi ambigue e sulle maggiori o minori difficoltà incontrate dal singolo soggetto durante questo processo, spesso anche molto differenti sia qualitativamente che quantitativamente.

Prima di vedere nel dettaglio a quali risultati e considerazioni si è giunti nell'ambito della ricerca, è utile osservare come le strategie di segmentazione dell'enunciato siano state varie per ciascun parlante con delle modifiche anche per i singoli parlanti a seconda delle ripetizioni di una stessa frase.

I parlanti hanno impiegato, come vedremo, strumenti prosodici differenti, quali durata delle sillabe toniche, contorni intonativi con variazioni di f_0 molto variegato, volumi e registri di frequenze differenti e pause al confine di sintagma. Tutto ciò si evince dalle misurazioni effettuate a livello acustico, sulle f_0 delle vocali toniche e di confine di sintagma, nonché dai livelli di intensità delle stesse.

È estremamente interessante notare, pertanto, come ogni parlante impieghi strategie di *phrasing* personali e grazie all'esperienza effettuato è stato possibile notare ciò che avviene nella fase precedente la produzione, ovvero nella fase in cui i parlanti hanno effettuato un processo di riconoscimento iconico prima e una trasformazione di tale informazione in termini semantici e di senso dopo. È bene osservare, infatti, che in questa fase ogni parlante ha presentato reazioni differenti tanto al riconoscimento dell'immagine, quanto all'assegnazione di essa al significato della frase potenzialmente ambigua. Ma ciò che è più interessante è che ogni parlante ha avuto reazioni e difficoltà differenti in più in casi e ripetizioni diverse.

Quasi tutti i parlanti hanno incontrato delle titubanze nella produzione di frasi in cui il sintagma nominale era costituito da [la vecchia porta] e in generale quando [vecchia] svolgeva la funzione di aggettivo. La maggior parte di essi, pur riconoscendo l'icona assegnata alla frase, produceva nell'altro senso, dimostrando una certa propensione per una scelta sintattica più comoda.

Abbiamo osservato anche incertezze e titubanze nella fase del riconoscimento grafico di alcune frasi e pertanto le stesse sono state ripetute alla fine della sequenza. Inoltre, i parlanti alla fine delle registrazioni hanno mostrato interesse per quanto prodotto e hanno fornito spontaneamente osservazioni proprio sulla fase interpretativa e sulle difficoltà incontrate nella disambiguazione, nel riconoscimento più o meno immediato dell'immagine e sulla tendenza ad interpretare la sequenza sintagmatica in un modo piuttosto che in un altro.

Si può, pertanto, affermare che l'interpretazione della frase con struttura SVO e sintagma nominale soggetto non ramificato risulti più semplice ed immediato (anche se non in modo generalizzato, come potrebbe risultare ovvio), giacché la struttura a dislocazione a sinistra o a destra con l'elemento anaforico o cataforico a seconda dei casi, presenta un grado di interpretazione e produzione più complessa e articolata.

3.2 Seconda fase: l'analisi acustica

3.2.1 Analisi delle vocali medie

Prima di procedere all'analisi acustica delle variazioni soprasedimentali delle produzioni, è stata svolta un'indagine acustica sulle vocali medie toniche. Anche il timbro delle vocali, infatti, potrebbe rappresentare un correlato della diversa interpretazione degli enunciati ed essere quindi rilevante in fase di attribuzione del significato.

L'analisi si spiega anche nell'ambito della complessa ricerca sul variegato vocalismo salentino. È stato, infatti, produttivo ed interessante analizzare i valori formantici e le loro variazioni da parlante a parlante. La ricerca sul vocalismo, in particolare sulle vocali medie, e i fenomeni di metaforia dialettale si trova ormai ad un punto notevolmente avanzato, sebbene nuovi scenari si aprano continuamente ed aumentino gli spunti di ricerca, come è emerso anche dall'indagine. Il vocalismo tonico del Salento meridionale (di quell'area, cioè, che parte dalla linea Gallipoli-Maglie-Otranto e arriva a S.M. di Leuca) si presenta molto frammentato e disomogeneo (soprattutto quello medio). I principali risultati a cui si è giunti mostrano che /e/ ed /o/ toniche sono

effettivamente delle vocali medie [E] ed [O]; il fatto nuovo è che nel triangolo compreso fra S.M. di Leuca a sud, Ruffano ad ovest e Andrano ad est sono stati individuati dei processi metafonetici che interessano le medie [E] quando sono seguite da *-i/*. Identici processi, anche se in modo differenziato e non in tutti i punti, interessano le [O] quando sono seguite da *-i/* oppure anche da *-u/*, sia in sillaba libera che in sillaba chiusa. Tali processi producono l'innalzamento in [e] ed [o] delle rispettive medie [E] ed [O]; bisogna aggiungere che il condizionamento delle vocali medie in funzione delle vocali alte atone è stato rilevato anche nell'italiano parlato del Salento meridionale. Mancando nell'italiano termini con */u/* finale, la metaforesi è stata registrata in forme con [E] oppure [O] toniche seguite da *-i/*. Dunque nell'italiano regionale del Salento meridionale, e più precisamente nell'italiano dell'area a sensibilizzazione metafonetica, */e/* ed */o/* toniche, realizzate dai parlanti come [E] / [O], diventano [e] / [o] quando sono seguite da una *-i/* (Canepari [1983; 1992] e Grimaldi [2003]).

Anche questa ricerca sottolinea quanto sia consistente il divario fra il piano teorico, fonologico e quello pratico, fonetico fornito dalle misurazioni sul campo e dall'analisi dei dati raccolti: nel momento in cui la realtà linguistica astratta incontra i parlanti e le loro varieguate strategie, retroterra culturali e produzioni raramente rispondenti alla teoria, essa mostra quanto questo divario sia difficile da colmare, se non impossibile.

Dai dati raccolti su tre ripetizioni per ciascuna interpretazione di tre dei dieci parlanti emerge che le due vocali medie */e/* ed */o/* presentano delle variazioni acustiche relativamente significative da più punti di vista ed in base a differenti parametri. Le due vocali sono state, infatti, analizzate e successivamente trattate statisticamente in considerazione di parametri quali il tipo di frase ed il parlante. L'analisi acustica, qui come più avanti considerando parametri diversi, è stata effettuata con PRAAT.

Per quanto riguarda il tipo di frase, si intende evidenziare la differenza di produzione in base all'interpretazione (come cambiano i valori formantici di */e/* in "vecchia" soggetto e "legge" verbo rispetto a quelli di */e/* in "vecchia" aggettivo e "legge" soggetto), mentre il rapporto frase-parlante mira a sottolineare il fatto che il comportamento linguistico può variare intersoggettivamente (i tre parlanti possono effettuare scelte diverse). In questo caso, si è indagato il rapporto fra i parlanti e la parola: vale a dire, il rapporto fra il parlante e la produzione della vocale a seconda della parola, tenendo conto sia della sua natura semantica e fonetica ("vecchia" vs. "legge"), sia del ruolo sintattico svolto nei quattro casi (sogg. – verb. – agg.), andando ad osservare quali variazioni possono esserci in relazione non solo al cambiamento di parola ma anche di funzione sintattica nei casi di polivalenza di una stessa. Per ricavare indicazioni sul timbro delle vocali è stata effettuata un'indagine acustica sui valori delle prime due formanti. Le stesse misurazioni sono state effettuate sui valori della prima formante (F1) e su quelli della seconda (F2), dando spesso risultati divergenti ed interessanti.

3.2.2 *Analisi acustica delle variazioni prosodiche*

In seguito, le venti ripetizioni dei dieci soggetti sono state analizzate tenendo conto di fattori prosodici, quali la durata delle sillabe toniche, misurata in millesecodi (ms); la durata della sillaba al confine; la frequenza fondamentale iniziale e finale della vocale al confine sintattico tra sintagma nominale e sintagma verbale, nonché la sua intensità media misurata in decibel (db) – indicante la variazione di volume – il valore in Hz di *f0* iniziale

e finale delle vocali toniche, indicante l'andamento intonativo; la posizione del *pitch*⁴ massimo della vocale nelle sillabe toniche e nella vocale atona al confine; le pause misurate in millesecondi, ove presenti.

Ricapitolando, dunque, la nostra attenzione si è focalizzata soprattutto su alcuni dei numerosi parametri acustici segmentali e soprasedimentali che caratterizzano la catena fonica durante la produzione. I correlati presi in considerazione in quanto più significativi ai fini dell'indagine sono stati:

- i valori formantici di F1 ed F2 delle vocali medie di “vecchia”, “legge” e “porta”
- le durate delle sillabe toniche, S1 ([vek]), S2 ([ledZ] o [por]) ed S3 ([re] o [zbar])
- le durate delle sillabe precedenti il confine sintattico, DopoS1 ([-kja]) e DopoS2 ([-dZe la] o [-ta la])
- la *f0* finale della vocale al confine del sintagma nominale con quello verbale
- l'intensità media della vocale al confine del sintagma nominale con quello verbale
- posizione del *pitch* massimo nella sillaba tonica e in quella atona al confine di sintagma
- durata di pause al confine di sintagma e loro influenza sulle durate sillabiche.

Ciascuno di questi parametri ha dato spunto a riflessioni significative sulla realtà che ci si era proposti di indagare: il processo di disambiguazione di frasi polisemantiche e l'impiego dei correlati prosodici in tale processo.

Le variazioni di frequenze, di durate sillabiche, di frequenza fondamentale e del suo correlato percettivo, l'intensità, della posizione del picco di frequenza nella sillaba e di durate prepausali, ove impiegate, analizzate in considerazione al cambiamento del ruolo sintattico provocato da un possibile doppio *phrasing* della frase, hanno dimostrato su un piano pratico quanto supposto da un punto di vista teorico.

In particolare da quanto osservato sui risultati ottenuti, si può affermare che esiste nella produzione di frasi potenzialmente ambigue una preminenza della categoria sintattica del soggetto logico tesa a disambiguare la polisemia. Tale preminenza si realizza attraverso durate sillabiche maggiori, possibile presenza di pause al confine sintagmatico fra soggetto e verbo, spostamenti “strategici” del *pitch* nelle sillabe atone e toniche del soggetto (soprattutto in presenza di pausa), modificazioni dell'andamento melodico realizzato mediante variazioni di frequenza fondamentale ed intensità al confine.

Inoltre, si può notare la rilevanza di un'altra categoria sintattica strettamente collegata al sintagma nominale, il predicato verbale. Tale elemento della frase presenta una relativa preminenza, tradotta in termini acustici, nell'interpretazione della frase con prolessi dell'oggetto attraverso elemento clitico anaforico. Su questo piano di difficoltà maggiore dell'interpretazione, la prosodia risulta necessaria nel processo di disambiguazione, come dimostrato dalle modulazioni effettuate dal parlante su elementi sintattici della frase considerati strategici ai fini della comprensione e dell'interpretazione: il soggetto e il verbo, ed anche l'oggetto.

L'analisi ha di fatto evidenziato non solo una certa coerenza nell'andamento dei singoli parametri acustici e prosodici, in relazione al processo teso a disambiguare le due frasi, ma anche una realtà variegata di soluzioni possibili emerse dalle considerazioni relative alle strategie dei singoli individui o dei gruppi di parlanti: si vedano le differenze osservate in base al genere. Sia i parlanti maschi che le parlanti, infatti, hanno prodotto le

⁴ Il *pitch* massimo fa riferimento al picco di frequenza fondamentale raggiunto in un determinato contesto, il *pitch* minimo indica il registro di frequenza utilizzato dal parlante durante l'eloquio, la differenza tra il massimo e il minimo (il delta) fornisce indicazioni, invece, sull'intervallo di frequenze impiegato (*pitch range*).

frasi impiegando spesso i correlati prosodici a loro disposizione in modo differente, in un intreccio di variazioni di volta in volta personali ma tutte ugualmente valide.

3.3 Terza fase: la percezione

Il corpus di questa fase dell'indagine sperimentale consiste in 24 stimoli uditivi ricavati dalle manipolazioni di una frase originale, registrata nella prima fase dedicata alla produzione, realizzata da una parlante salentina di Tuglie (comune a 10 km a Nord-est di Gallipoli) partecipante all'esperimento. La scelta fra le due frasi originarie ("La vecchia legge la regola" e "La vecchia porta la sbarra") è ricaduta sulla prima perché ritenuta più lineare da un punto di vista acustico e uditivo e più interessante rispetto all'analisi dei dati raccolti nella fase di produzione per quel che riguarda strategie intonative e ritmiche di disambiguazione.

Per la definizione dei parametri di manipolazione dello stimolo originale sono stati riesaminati i valori di durata sillabica, di pausa, di f_0 e di intensità registrati sulle parlanti femmine (cinque) in fase di produzione nel primo esperimento. Grazie a questa analisi è stato possibile definire i contorni intonativi e le caratteristiche prosodiche in genere impiegati dalle parlanti per disambiguare la frase in due interpretazioni. Le conclusioni a cui si è giunti in questa fase sui dati delle parlanti nella disambiguazione della frase "La vecchia legge la regola" possono essere così sintetizzate:

- Significatività delle durate sillabiche delle toniche [vek], [ledZ], [re] anche in base alla presenza/assenza di pausa
- Significatività delle durate sillabiche sia prepausali che non delle atone [kja] e [ge]
- Significatività della frequenza fondamentale iniziale e finale delle vocali toniche di [vek] e [ledZ] anche in base alla presenza/assenza di pausa
- Significatività dell'intensità al confine di sintagma anche in base alla presenza/assenza di pausa

Sulla base di queste premesse l'attenzione si è focalizzata sulle durate sillabiche, la frequenza fondamentale e la pausa. Questi tre fattori infatti sono risultati responsabili dei paradigmi prosodici usati dalle parlanti sulle quali sono state condotte delle medie da cui ricavare i passi di manipolazione.

Uno di questi fattori, la pausa, si è rivelato di particolare rilevanza sia in questo studio che in quello dedicato alla produzione. Si tratta di un fenomeno prosodico strettamente correlato all'intonazione, in quanto può occorrere con essa per segnalare la presenza di confini sintattici, alla fine di un turno di conversazione, di una frase o di un sintagma.

Le ricerche sulla realizzazione delle pause silenti o vuote in testi letti hanno dimostrato che la loro distribuzione dipende sistematicamente dalla strutturazione sintattica dell'enunciato in clausole principali coordinate e subordinate⁵. All'interno delle frasi, abbiamo visto, la produzione delle pause può disambiguare la struttura sintagmatica, in cooccorrenza con l'andamento ascendente o discendente di f_0 e con l'allungamento della sillaba prepausale.

Infine le pause possono essere prodotte anche all'interno dei sintagmi sintattici, in dipendenza dalla lunghezza dei costituenti sintattici, calcolata come numero di parole-contenuto o di sillabe, e anche dalla velocità di eloquio.

⁵ Questa sistematica presenza della pausa viene, infatti, prescritta nel linguaggio scritto dai diversi segni di punteggiatura (punto, virgola, punto e virgola, puntini di sospensione, ecc.).

Non solo la distribuzione delle pause è importante, ma anche la loro durata, in quanto è stato dimostrato che, sia nella lettura che nel parlato spontaneo, la gerarchia della durata delle pause grammaticali corrisponde alla gerarchia dei costituenti intonativi. La durata delle pause di esitazione, invece, non presenta alcuna sistematicità, poiché i disturbi dei processi di pianificazione che le causano possono dipendere da fattori idiosincratici e momentanei quali le conoscenze linguistiche del soggetto, la difficoltà del compito, la situazione pragmatica, ragioni paralinguistiche, ecc (cfr. Bertinetto e Magno Caldognetto [1993]).

La frase scelta, come accennato, si è caratterizzata per una maggiore variabilità di scelte intonative (che vanno da ascendenti-discendenti-ascendenti, a discendenti-ascendenti, a discendenti progressive ad ascendenti-discendenti). Pertanto in base ai valori medi ricavati sulle durate sillabiche, sulla *f0* e sulla pausa è stata selezionata una ripetizione delle cinquanta totali, in cui questi parametri fossero più vicini possibile alla media. La frase è stata in seguito progressivamente manipolata tenendo conto del seguente ordine:

- Due step di manipolazione nell'interpretazione con "vecchia" soggetto per la durata della prima sillaba tonica [vek], della prima posttonica [kja], della seconda tonica [ledZ] e della seconda posttonica [ge], aumentate rispettivamente di 16,5 e 4,5 ms per le prime due sillabe e diminuite di 15 ms per le altre due sillabe.
- Tre step di manipolazione nell'interpretazione con "vecchia" soggetto per le frequenze fondamentali iniziali e finali della prima sillaba tonica [vek] e della seconda [ledZ], aumentate rispettivamente di 7,5 e 20 Hz la prima e diminuite di 8 e 11 Hz la seconda.

Le cinque manipolazioni così ottenute sono state incrociate per parametri modificati ottenendo così altre sei manipolazioni in cui agli step di durata si aggiungevano quelli di frequenza. Le undici manipolazioni finali insieme alla frase originale sono state infine trattate con l'aggiunta di pausa (190 ms) dopo "vecchia". Pertanto ai soggetti sono stati somministrati in tutto ventiquattro stimoli.

Le manipolazioni sono state elaborate nel tentativo di fornire un progressivo passaggio dall'interpretazione con "vecchia" soggetto ad una con "vecchia" aggettivo sulla base del campione raccolto nel primo studio. In questo diaframma alcuni stimoli trovandosi a metà fra un'interpretazione e l'altra hanno comportato non poche difficoltà in fase di percezione, come vedremo a breve. Di fatto le scelte effettuate nella fase di preparazione del corpus corrispondono all'esigenza di indagine alla base di questo lavoro ovvero allo studio più approfondito dei meccanismi e delle strategie che entrano in gioco nella fase di percezione dell'ambiguità, che si distingue in maniera piuttosto netta rispetto a quella di produzione, in quanto al processo di rielaborazione del messaggio linguistico concorrono fattori uditivi di adattamento a strategie ritmiche ed intonative specifiche del parlante che in casi di ambiguità possono rivelarsi più dinamiche e a volte caratterizzate da una certa marcatezza, rispetto a casi di enunciati "lineari".

In questo studio insieme ai fattori sopraelencati, anche l'organizzazione sintattica è stata determinante per l'intonazione. Non è detto infatti che una frase SVO venga realizzata con un unico andamento intonativo. Questa funzione di individuazione dei costituenti sintattici è alla base della disambiguazione di casi in cui una diversa intonazione e una diversa posizione della pausa distinguono due frasi apparentemente uguali nella lingua scritta, ma differenziate in termini di struttura sintattica, proprio come nella frase presa in considerazione in questo lavoro.

È bene ricordare inoltre che lo studio pilota aveva mostrato il ruolo che questo elemento del sintagma nominale/verbale (a seconda del *phrasing*) svolgeva nella produzione avvenuta subito dopo il breve processo di disambiguazione: allungamenti e pause precedevano in maniera significativa il "la" pronome.

In questo studio, già ad un'analisi qualitativa e descrittiva si evince che l'ascolto delle due interpretazioni ha creato non pochi problemi di discriminazione. I soggetti in molti casi hanno spontaneamente riferito al termine della sessione di aver trovato difficoltà a collocare pragmaticamente il riferimento pronominale, in coerenza con la necessità di un rapporto lineare tema-rema.

In questa fase del lavoro è stato impiegato il metodo di *eye tracking* (cfr. Cooper [1974]). Nella ricerca scientifica attuale vi sono due orientamenti principali per quanto concerne le misurazioni di movimenti oculari. Quelle che misurano la posizione dell'occhio in relazione alla testa e quelle che misurano l'orientamento dell'occhio nello spazio, o "punto di vista". Quest'ultimo metodo di *eye tracking* è usato specialmente nell'identificazione di elementi in una scena visiva, per es. in applicazioni interattive. In questo studio, infatti, ad un test di identificazione di tipo uditivo è stato associato il contemporaneo monitoraggio dei movimenti oculari attraverso l'impiego di un *eye tracker* da tavolo con videocamera ad infrarossi. Essa invia la luce al centro dell'occhio che la riflette permettendo alla camera di memorizzare le coordinate oculari e registrare i movimenti nel momento in cui il soggetto compie un determinato compito. Molti dei moderni *eye tracker* (come quello qui impiegato) usano il contrasto per localizzare il centro della pupilla e la luce infrarossa per creare il riflesso corneale (CR). Il vettore tra queste due caratteristiche può essere usato per computerizzare l'intersezione dello sguardo (*gaze*) con una superficie dopo una semplice calibrazione da effettuare per ogni soggetto.

Lo scopo dell'*eye tracking* è molto spesso di stimare la direzione dello sguardo e misurare i suoi cambiamenti. Una calibrazione accurata e affidabile dello strumento è pertanto essenziale, come vedremo, per ottenere dati sui movimenti validi e ripetibili. Le tecniche sono oggetto di continuo progresso e miglioramento, tuttavia l'interpretazione dei risultati richiede ancora uno studio approfondito e maggiore esperienza poiché errori di discriminazione o calibrazione possono inficiare l'intera indagine e fornire dati fuorvianti.

Si può trovare un ampio assortimento di studi sull'*eye tracking* nei campi interdisciplinari di neuroscienze e psicologia (cfr. Cacciari [2001]). Grazie a questo metodo, in effetti, è possibile trarre delle osservazioni sul livello di attenzione che un soggetto possiede nei confronti di quello che sta osservando, in determinati compiti cognitivi, sul suo modo di trattare le informazioni contenute nelle pagine, sulle sue strategie di esplorazione, sui possibili problemi che può incontrare. Il vantaggio dell'*eye tracking* è quello di fornire risultati oggettivi e riproducibili con una metodologia di rilevazione di dati ad invasività zero. Si tratta, infatti, di una tecnica in grado di registrare la dilatazione e la contrazione delle pupille, realizzando un effettivo tracciamento oculare che definisce l'intero percorso effettuato dall'occhio durante la visione. Essa nasce per scopi clinici, con l'obiettivo di capire come funzionano i meccanismi della visione umana, individuare che cosa si sta guardando in ogni momento o con quale livello di attenzione, attraverso la registrazione della dilatazione e contrazione delle pupille (cfr. Young e Sheena [1975], Rayner [1978; 1992; 1998], Daneman e Reingold [1993], Allopenna et al. [1998], d'Ydewalle et al. [1998], Henderson e Hollingworth [1998], De Luca et al. [1999], Anders [2001], Van Orden et al. [2001], Wooding [2002], Ashby e Clifton [2004], Hirotani et al. [2005], Rele e Duchowski [2005], Weber et al. [2006], Land [2006], Di Nocera et al. [2007]).

In questa fase della ricerca, l'analisi dei movimenti oculari, dunque, è servita a studiare alcuni aspetti della risoluzione dell'ambiguità. La disambiguazione, infatti, è un aspetto della comprensione e della produzione del linguaggio che ha interessato largamente la ricerca, con realizzazioni sperimentali diverse e su sistemi linguistici differenti. La letteratura a disposizione è molto ampia e pertanto si cercherà di delineare il

percorso storico che dagli anni '90 ha compiuto quest'ambito della ricerca evidenziando i metodi impiegati e i risultati più significativi a cui si è giunti anche ai fini di questo lavoro (cfr. Rayner et al. [1983], Frazier e Rayner [1987], Tanenhaus e Carlson [1989], Altmann et al. [1992], McDonald et al. [1994], De Vincenzi e Job [1995], Tanenhaus e Trueswell [1995], Hirscherberg e Avesani [1997], Gamsey et al. [1997], Snedeker e Trueswell [2002; 2003], Spivey et al. [2002]).

A questo primo livello di analisi è seguito il trattamento statistico dei dati grezzi che permette un'indagine più approfondita e dettagliata. I risultati sono stati trattati statisticamente mediante il test ANOVA con una soglia di significatività pari a: $p < 0.05$ e con il test post-hoc di Fisher, grazie al quale è possibile confrontare le componenti dei fattori che sono risultati significativi.

I dati sono stati considerati da più punti di vista e con criteri differenti al fine di verificare la significatività di fissazioni e saccadi (nei termini di durata e di numero), in relazione all'interpretazione della frase fornita in base alle caratteristiche dello stimolo.

Sui dati registrati quindi sono state condotte due analisi di tipo qualitativo: la prima sulle fissazioni di tutti i soggetti per alcuni stimoli apparsi particolarmente interessanti per numero e durata di fissazioni ricevute e la seconda sullo *scanpath* di alcuni soggetti per i quali si avevano più dati puliti a disposizione e dai quali si potevano ricavare maggiori informazioni sul percorso compiuto dall'occhio sullo schermo risalendo quindi anche a movimenti saccadici, fissazioni e *interfix*.

La prima analisi descrittiva ha permesso di evidenziare il "luogo" in cui è avvenuta la fissazione, facendo emergere anomalie, indecisioni e cambiamenti di interpretazione. Gli stimoli esaminati sono quelli risultati più significativi per durata di fissazione fra quelli in cui è stata manipolata la durata sillabica e la frequenza fondamentale ed è stata aggiunta la pausa.

Da questa analisi emerge che:

- Alcune fissazioni di natura interpretativa si sono verificate già sulla parola "vecchia", sulla vocale tonica ma anche sulla sillaba atona.
- Alcune fissazioni di cambio di interpretazione si sono verificate sia sulla sillaba tonica che sulla atona di "legge".
- Anche l'occorrenza della pausa fra "vecchia" e "legge" ha causato l'insorgenza di fissazioni come anche l'articolo/pronome "la" prima di "regola".
- Alcune fissazioni si trovano all'inizio dello stimolo (nella maggior parte dei casi nella zona inferiore dello schermo), ma danno indicazione solo della posizione dell'occhio al momento della visualizzazione dell'immagine poiché in quell'istante temporale non vi era stata ancora elaborazione uditiva.
- La maggioranza delle fissazioni si registra comunque a fine stimolo su "regola" o dopo lo stimolo uditivo e durante la permanenza dell'immagine sullo schermo (1,5 s); questo dato dimostra che i soggetti parlanti hanno voluto o hanno potuto dare una risposta interpretativa sullo stimolo solo dopo averlo ascoltato nella sua interezza mentre durante l'ascolto hanno fornito risposte parziali più o meno volontarie e chiare (fissazioni non "fuori area").

Procedendo in questo modo è stato possibile osservare ancor più da vicino il processo interpretativo dello stimolo uditivo e l'associazione ad un'interpretazione visiva. In conseguenza a queste considerazioni è stato ritenuto opportuno approfondire l'analisi su alcuni soggetti per i quali era più facile individuare lo *scanpath* (la sequenza di saccadi e fissazioni) durante la sessione sperimentale (Fig. 2).

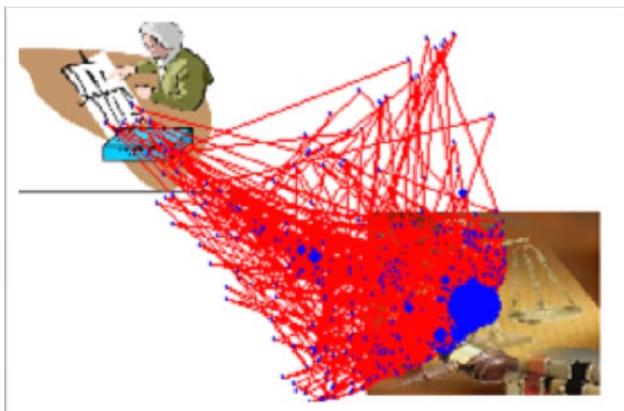


Figura 2. Immagine tratta dal programma di analisi Fixplot sulle fissazioni effettuate da uno dei soggetti partecipanti all'esperimento: sono visibili lo *scanpath* e le fissazioni sulle aree di interesse (Aoi).

Da questa seconda analisi si possono ricavare le seguenti considerazioni:

- massiccia attività oculare: per ogni stimolo i movimenti oculari sono stati numerosi e vari, i soggetti esaminati hanno mostrato di reagire agli stimoli visivo-uditivi muovendosi nello spazio visivo in modo attivo e partecipativo;
- maggiore confidenza e minore margine di errore (fissazioni fuori area) con l'incarico sperimentale direttamente proporzionale al passare del tempo;
- aumento delle saccadi precedenti fissazioni nelle ultime sessioni randomizzate in direzione delle due immagini;
- presenza nell'*interfix* sia di saccadi nelle stesse aree delle fissazioni che precedono sia di movimenti oculari in direzione di aree esterne allo schermo;
- la perdita dei dati risulta spesso seguire fissazioni lunghe e può quindi essere dovuta a naturale chiusura delle palpebre per umidificazione dell'occhio o per stanchezza del soggetto.

Un'altra indagine è stata condotta successivamente sulle durate medie di fissazione tenendo conto dell'Aoi (*Area of Interest*): questo ha permesso di calcolare delle medie e delle percentuali relative alla durata complessiva di fissazione per ogni soggetto in determinate zone dello schermo. Per quanto riguarda le zone "focali" di questo esperimento, quelle corrispondenti alle immagini delle due interpretazioni della frase, è stato possibile determinare la quantità di tempo media in cui il soggetto ha guardato l'immagine relativa a "vecchia-soggetto" oppure quella relativa a "vecchia-aggettivo". In generale bisogna sottolineare che sul totale delle fissazioni registrate il 63% è andato alle immagini target, il resto alle altre zone dello schermo. Questo dato come è stato già evidenziato non deve sorprendere e resta in linea con gli studi in cui sono state notate discrepanze ed "errori" simili, in realtà dovuti a indecisione ma anche ad una predisposizione del soggetto non corrispondente ad un segnale interpretativo (cfr. Cooper [1974]).

Di queste fissazioni il 68% è andato all'immagine di "vecchia-aggettivo" e il restante 32% all'immagine di "vecchia-soggetto". Per quanto riguarda le durate, in percentuale è risultato che in media il tempo di fissazioni è stato molto simile per entrambe le interpretazioni e si attesta intorno al 30% per ciascuna della durata totale di

ogni sessione somministrata ai soggetti. Questo significa che, come vedremo anche dall'analisi statistica, le fissazioni per "vecchia-soggetto" se minori in numero rispetto all'altra interpretazione sono state più lunghe per durata.

4. Risultati e Conclusioni

In maniera schematica i risultati delle analisi descritte fin qui possono essere schematizzati come segue:

- Ad un livello generale emerge una sostanziale differenza nella scelta delle interpretazioni che sono state attribuite agli stimoli.
- I dati sulle fissazioni dimostrano che in numero sono maggiori quelle per l'interpretazione "vecchia-aggettivo", mentre per quanto riguarda la durata prevale l'interpretazione "vecchia-soggetto": l'immagine corrispondente a questa interpretazione ha ricevuto fissazioni più lunghe.
- Per quanto riguarda i singoli stimoli, in generale, si evidenziano interessanti differenze riguardanti la manipolazione. Vi sono, infatti, alcuni elementi che hanno condizionato la scelta del soggetto nell'attribuzione di un significato allo stimolo ascoltato. Dai dati sulle fissazioni si evidenzia in ogni caso un'incertezza decisionale anche per quanto riguarda i controlli. In alcuni casi questi risultati hanno contraddetto in parte le aspettative di partenza per cui gli stimoli di controllo sarebbero dovuti essere chiari in modo netto ed univoco (uno per ciascuna interpretazione). I restanti 22 stimoli avrebbero creato, come di fatto è stato, incertezze nell'interpretazione in maniera graduale e differente a seconda delle manipolazioni che su di essi erano state effettuate.
- Dai risultati emerge che la pausa è stata un elemento disambiguante di portata elevata in quanto più della metà degli stimoli contenenti il silenzio dopo "vecchia" ha ricevuto l'interpretazione "vecchia-soggetto"; questo è un dato che conferma le aspettative, in quanto la pausa in questa posizione dava alla frase una connotazione ritmica tipica dell'andamento intonativo della frase con struttura SVO.
- La manipolazione di durata sillabica ha influito sulla disambiguazione, anche se in maniera contenuta, in direzione dell'interpretazione con "vecchia" in funzione di soggetto. Il dato è confermato per entrambi gli step di manipolazione della durata. Anche questo aspetto è in linea con le aspettative, in quanto le manipolazioni di durata erano state effettuate tenendo conto del campione raccolto nello studio pilota sulle produzioni di "vecchia-soggetto".
- L'elemento disambiguante in contrasto era invece costituito dalla variazione di frequenza fondamentale. Gli effetti di tale manipolazione hanno confermato ma anche superato le aspettative. Era atteso un dato significativo riguardante il cambiamento di tono su "vecchia" e su "legge" nell'attribuzione del significato. Quello che emerge è che la variazione di frequenza è prevalsa sulla pausa nell'interpretazione data alla frase: in un terzo dei casi in cui alla variazione di frequenza era associata la pausa l'interpretazione prevalente è stata quella per "vecchia-aggettivo". D'altra parte però il dato contraddice le aspettative per stimoli in cui la variazione di frequenza (tre step) era presentata singolarmente come uno elemento manipolato e per gli stimoli in cui era associata alla variazione di durata sillabica: solo in tre casi su nove l'interpretazione di "vecchia" è stata quella di aggettivo, nei restanti due terzi l'interpretazione della frase è rimasta quella a struttura SVO, persino per i controlli.
- Altri dati importanti emergono dalle classi di durata scelte per suddividere le fissazioni sopra il valore minimo stabilito come indice di avvenuta attività cognitiva: a prevalere

sono fissazioni di durata compresa fra 1000 e 3000 ms, dimostrando di essere indice di una scelta interpretativa effettuata; è vero anche che molte sono le fissazioni di ordine inferiore, comprese fra i 50 e gli 80 ms. Queste brevi fissazioni insieme alle tante saccadi registrate sono indice dell'attività motoria oculare verificatasi sullo schermo durante l'ascolto: la presenza di un così variegato *scanpath* deve essere tenuta presente nella considerazione dei dati riguardanti le fissazioni e sul processo decisionale effettuato dai soggetti.

- Anche l'analisi effettuata sulla latenza di fissazioni in riferimento ad un elemento target (la seconda sillaba tonica) ha fornito indicazioni interessanti: è emersa una latenza maggiore per le interpretazioni date per "vecchia-aggettivo". In questi casi la fissazione in media è avvenuta dopo lo stimolo target e spesso dopo "la" articolo/pronome. Per "vecchia-soggetto" la latenza è minore e in media la fissazione diretta a questa interpretazione è stata effettuata in corrispondenza del target o subito dopo, quindi anche prima che l'ascolto della parola "legge" terminasse. I dati sulla latenza sono molto vari e differenziati, ma forniscono indicazione di come effettivamente sia avvenuta la comprensione ed eventualmente la disambiguazione dello stimolo. È emerso che in alcuni casi la fissazione sull'immagine corrispondente è avvenuta solo alla fine dell'ascolto, durante il periodo di permanenza dell'immagine sullo schermo, mentre prima il soggetto aveva effettuato fissazioni "fuori area". In altri casi tuttavia si può affermare che a livello di motilità oculare è avvenuto uno spostamento legato ad un'attività cognitiva di interpretazione dopo l'ascolto della pausa e delle vocali toniche.

Dall'analisi effettuata emergono punti che contraddicono le aspettative e questo risultato potrebbe essere dovuto ad aspetti procedurali che hanno influito in questo senso. Ciononostante lo studio fornisce indicazioni importanti per una prima considerazione su come avviene il processo di interpretazione dell'ambiguità. Anche in questo studio, come era emerso in precedenza osservando il trattamento dell'ambiguità in produzione, i soggetti fanno riferimento a ragioni non solo linguistiche ma anche pragmatiche nella scelta dell'interpretazione ed in generale nel delicato processo di percezione. Nel comprendere il messaggio uditivo e nell'interpretarlo i soggetti adattano le loro intuizioni alla situazione comunicativa e spesso finiscono per esserne influenzati. Nella prima indagine emerge una propensione per la struttura non marcata della frase, come confermano vari studi (cfr. fra gli altri Chomsky e Halle [1968], Nespor e Vogel [1986], Benincà, P. [1988], Avesani et al. [1995], Selkirk [1995], Avesani [1997], Avesani e Vayra [1997]) e come è in un certo senso linguisticamente naturale; in questo lavoro lo stesso aspetto sembrerebbe essersi ripresentato anche nella fase in cui si ascoltava, e non si può escludere che si sia trattato in fase di interpretazione anche di una predisposizione, di un "voler sentire". È vero però che la marcatezza è una caratteristica tipica della lingua e soprattutto di quella orale: l'uso di varianti sintattiche, di ordine tematico e di contorni intonativi, insieme a molti altri strumenti a disposizione del parlante, si rivela connaturato ed indispensabile alla comunicazione stessa.

Il metodo di *eye tracking* utilizzato ha il vantaggio di fornire un quadro sostanzialmente chiaro di ciò che avviene a livello cognitivo alla presentazione di uno stimolo visivo e di uno uditivo. Si tratta di uno strumento preciso in grado di registrare ogni variazione che avviene nel sistema attenzionale e trova un riflesso nei movimenti oculari. È auspicabile, pertanto, l'approfondimento di quanto è emerso finora, utilizzando la stessa metodologia o integrandola con altri strumenti sperimentali.

L'ambiguità è un elemento tipico della lingua e non sempre passa inosservato, soprattutto nei casi in cui si pone ai limiti dell'incomunicabilità. Alla disambiguazione prendono parte, come si è visto, elementi linguistici e soprassessuali ma soprattutto

situazionali, di contesto e di referenza. Quello che mancava ai soggetti in questo caso era proprio il corredo di altre informazioni che avrebbe permesso loro di attribuire un'interpretazione piuttosto che un'altra utilizzando un sistema tematico, come avviene nelle naturali situazioni comunicative. Ma questo probabilmente non avrebbe permesso di soddisfare lo scopo principale di questo lavoro, ovvero quello di risalire alla portata che i soli elementi prosodici sono in grado di fornire al processo di disambiguazione, che spesso richiede in base alla sua complessità un apporto ed un carico attenzionale notevoli.

Bibliografia

- Allopenna, P.D., Magnuson, J.S., Tanenhaus, M.K. 1998, *Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping model*, in «Journal of memory and language», 38, pp. 419-439.
- Altmann, G.T.M., Garnham, A., Dennis, Y. 1992, *Avoiding the garden path: Eye movements in context*, in «Journal of Memory and Language», 31, pp. 685-712.
- Anders, G. 2001, *Pilot's attention at location during approach and landing eye- and head-tracking research in an A330 full flight simulator*, in «International symposium on aviation psychology (ISAP)», Columbus, OH.
- Ashby, J., Clifton Jr., C. 2004, *The prosodic property of lexical stress affects eye movements during silent reading*, in «Cognition», 96, pp. B89-B100.
- Avesani, C. 1997, *Quantificatori, negazione e costituzione sintattica. Costruzioni potenzialmente ambigue e il ruolo della prosodia*, Atti del XXXI Congresso Internazionale di Studi della SLI, Padova.
- Avesani, C., Hirschberg J. e Prieto P. 1995, *The intonational disambiguation of potentially ambiguous utterances in English, Italian and Spanish*, in «Proceedings of the XIIIth International Congress of Phonetic Sciences», Stoccolma, agosto 1995.
- Avesani, C., Vayra M. 1997, *Costruzioni marcate e non marcate in Italiano. Il ruolo dell'intonazione*, Atti del XXXI Congresso Internazionale di Studi della SLI, Padova.
- Beckman, M. E., Pierrehumbert, J. B. 1986, *Intonational structure in Japanese and English*, in «Phonology Yearbook», 3, pp. 255-309.
- Benincà, P. 1988, *L'ordine degli elementi nella frase e le costruzioni marcate*, in L. Renzi (a cura di), *Grande grammatica italiana di consultazione, I. La frase. I sintagmi nominale e preposizionale*, Il Mulino, Bologna, pp. 129-194.
- Bertinetto, P.M. 1981, *Strutture prosodiche della lingua italiana. Accento, quantità, sillaba, giuntura, fondamenti metrici*, Accademia della Crusca, Firenze.
- Bertinetto, P.M., Magno Caldognetto, E. 1993, *Ritmo e intonazione*, in A.A. Sobrero (a cura di), *Introduzione all'italiano contemporaneo, II. La variazione e gli usi*, Laterza, Roma-Bari, pp. 141-185.
- Boersma, P., Weenik, D. 1996, *Praat, a system for doing phonetics by computer*, version 3.4. Institute of Phonetic Sciences of the University of Amsterdam, Report 132. (updated version on www.praat.org).
- Cacciari, C. 2001, *Psicologia del linguaggio*, Il Mulino, Bologna.
- Canepari, L. 1983, *Italiano standard e pronuncia regionale*, Padova: CLEUP.
- Canepari, L. 1992, *Manuale di pronuncia italiana*, Bologna: Zanichelli.
- Chomsky, N., Halle, M. 1968, *The Sound Pattern of English*, Harper & Row, New York.
- Cooper, R.M. 1974, *The control of eye fixation by the meaning of spoken language*, in «Cognitive Psychology», 6, pp. 84-107.
- D'Ydewalle, G., Desmet, G., Van Rensbergen, J. 1998, *Film Perception: the processing of film cuts*, in G. Underwood (ed.), *Eye guidance in reading and scene perception*, Elsevier, Oxford, pp. 357-368.
- Daneman, M., Reingold, E.M. 1993, *What eye fixations tell us about phonological recoding during reading*, in «Canadian Journal of Experimental Psychology», 47, pp. 153-178.
- De Luca, M., Di Pace, E., Judica, A., Spinelli, D., Zoccolotti, P. 1999, *Eye movement patterns in linguistic and non-linguistic surface dyslexia*, in «Neuropsychologia», 37, pp. 1407-1420.
- De Vincenzi, M., Job, R. 1995, *An investigation of Late Closure: the role of syntax, thematic structure, and pragmatics in initial and final interpretation*, in «Journal of Experimental Psychology: Learning, Memory and Cognition», 21, 5, pp. 1303-1321.
- Di Nocera, F., Camilli, M., Terenzi, M. 2007, *A random glance to the flight deck: pilot's scanning strategies and the real-time assessment of mental workload*, in «Journal of Cognitive Engineering and Decision Making», 3, pp. 271-285.
- D'Imperio, M., Gili Fivela, B. 2004, *How many levels of phrasing? Evidence from two varieties of Italian*, in J.Local, R.Ogden, R.Temple (eds.), *Phonetic Interpretation, Papers in Laboratory Phonology VI*, Cambridge University Press, Cambridge, pp. 130-144.
- Duchowski, A. 2007, *Eye tracking Methodology*, Springer-Verlag, London.
- Frazier, L., Rayner, K., (1987), *Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences*, in «Journal of Memory and Language», 26, pp. 505-526.
- Frenck-Mestre, C., Pynte, J. 1997, *Syntactic ambiguity resolution while reading in second and native languages* in «Quarterly Journal of Experimental Psychology, SOA», pp. 119-148.

- Gamsey, S. M., Pearlmutter, N. J., Myers, E., Lotocky, M. A., (1997), *The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences*, in «Journal of Memory and Language», 37, pp. 58-93.
- Gili Fivela, B. 1999, *The prosody of left-dislocated topicalized constituents in Italian read speech*, in «Proceedings of the European Conference On Speech Communication and Technology (EuroSpeech'99)», Budapest, Hungary, 5-9 September 1999, vol.1, pp. 531-534.
- Grimaldi M. 2003, *Nuove ricerche sul vocalismo tonico del Salento meridionale*, Edizioni dell'Orso, Alessandria.
- Henderson, J. M., Hollingworth, A. 1998, *Eye movements during scene viewing: an overview*, in G. Underwood (ed.), *Eye guidance in reading and scene perception*, Elsevier, Oxford, pp. 269- 294.
- Hirotnani, M., Frazier, L., Rayner, K. 2006, *Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements*, in «Journal of Memory and Language», 54, pp. 425-443.
- Hirscherberg, J., Avesani, C. 1997, *The role of prosody in disambiguating potentially ambiguous utterances in English and Italian*, in A. Botinis, G. Kouroupetroglu, G. Karayiannis (eds.), *Intonation: Theory, Models and Applications*, Proceedings of an ESCA Workshop, September 18-20, 1997, Athens, pp. 189-192.
- Land, M.F. 2006, *Eye movements and the control of actions in everyday life*, in «Progress in Retinal and Eye Research», 25, pp. 296-324.
- Mazuka, R., Itoh, K., & Kondo, T. 1997, *Processing down the garden path in Japanese: Processing of sentences with lexical homonyms*, in «Journal of Psycholinguistic Research», 26, pp. 207-228.
- McDonald, M., Pearlmutter, N.J., Seidenberg, M.S. 1994, *Lexical nature of syntactic ambiguity resolution*, in «Psychological reviews», 101, pp. 676-703.
- Nespor, M., Vogel, I. 1986, *Prosodic Phonology*, Foris Publication Dordrecht, Holland/Riverton-USA.
- Rayner, K. 1978, *Eye movements in reading and information processing*, in «Psychological Bulletin», 85, pp. 618-660.
- Rayner, K. 1992, *Eye movements and Visual Cognition: Scene Perception and Reading*, Springer-Verlag (Springer Series in Neuropsychology), New York.
- Rayner, K. 1998, *Eye Movements in Reading and Information Processing: 20 Years of Research*, in «Psychological Bulletin», 124, 3, pp. 372-422.
- Rayner, K., Carlson, A., Frazier, L. 1983, *The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences*, in «Journal of Verbal Learning and Verbal Behavior», 22, pp. 358-374.
- Rele, R., Duchowski, A. 2005, *Using eye tracking to evaluate alternative search results interfaces*, in «Proceedings of the human factors and ergonomics society», 49th annual meeting, Orlando, FL.
- Selkirk, E. 1995, *Sentence prosody: intonation, stress, and phrasing*, in John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, pp. 550-569, Cambridge, MA, and Oxford, UK: Blackwell
- Snedeker, J., Trueswell, J. 2002, *Using prosody to avoid ambiguity: Effects of speaker awareness and referential context*, in «Journal of Memory and Language», 48, pp. 103-130.
- Spivey, M. J., Tanenhaus, M.K., Eberhard, K. M., Sedivy, J. C. 2000, *Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution*, in «Cognitive Psychology», 45, pp. 447-481.
- Tanenhaus M.F., Carlson, G. 1989, *Lexical structure and language comprehension*, in W. Marslen-Wilson (a cura di), *Lexical representation and process*, The MIT Press, Cambridge, Mass., pp. 529-561.
- Tanenhaus, M.K., Trueswell, J.C. 1995, *Sentence comprehension*, in J.L. Miller e P.D. Eimas (a cura di), *Speech, language and communication*, Academic Press, New York, pp. 217-262.
- Van Orden, K.F., Limbert, W., Makeig, S., Jung, T.P. 2001, *Eye activity correlates of workload during a visuospatial memory task*, in «Human Factors», 43 (1), pp. 111-121.
- Watson, D., Gibson, E. 2004, *The relationship between intonational phrasing and syntactic structure in language production*, in «Language and cognitive processes», 19, pp. 713-755.
- Weber, A., Grice, M., Crocker, M.W. 2006, *The role of prosody in the interpretation of structural ambiguities: A study of anticipatory eye movements*, in «Cognition», 99, pp. B63-B72.
- Wooding, D. 2002, *Fixation Maps: quantifying eye movement traces*, in «Eye tracking Research & Application (ETRA) Symposium», New Orleans.
- Young, L.R., Sheena, D. 1975, *Eye-movement measurement techniques*, in «American Psychologist», 30, pp. 315-330.
- Zagar, D., Pynte, J., Rativeau, S. 1997, *Evidence for early-closure attachment on first-pass reading times in French*, in «Quarterly Journal of Experimental Psychology», 50A, pp. 421-438.