

La raccolta e l'analisi dei dati linguistici in un'indagine su parlanti bilingui

Yasmina Moussaid

1. Introduzione

Negli studi sul bilinguismo e nella linguistica del contatto è stato fatto vasto utilizzo di tecniche statistiche e metodi quantitativi per rispondere ai vari interrogativi presi in esame. Tuttavia, poiché molte delle variabili studiate nelle ricerche condotte in contesti bilingui richiedono spesso anche un metodo qualitativo di analisi dei dati linguistici, molti degli studi recenti che si sono occupati del contatto linguistico, in circostanze di bi- e multilinguismo, hanno iniziato a combinare gli approcci quantitativi a nuovi approcci di tipo qualitativo. Alcuni tra i primi studiosi che hanno iniziato a servirsi del metodo qualitativo in questo ambito sono, tra gli altri, FISHMAN 1967 e GUMPERZ 1972, i quali intorno agli anni Settanta del secolo scorso hanno iniziato a interessarsi a nuovi aspetti legati al bilinguismo, come la diglossia, i registri stilistici, i dialetti e l'utilizzo delle lingue in ambienti multilingui dal punto di vista sociolinguistico. L'interesse verso queste nuove tematiche li ha portati a introdurre nuovi metodi qualitativi che permettessero l'osservazione dei comportamenti linguistici dei parlanti. Tra questi vi è l'osservazione diretta dei partecipanti, di cui si è servito FISHMAN 1967 in combinazione con la somministrazione di questionari per la raccolta di dati quantitativi, e la raccolta di registrazioni di interazioni verbali di parlanti bilingui di cui ha usufruito GUMPERZ 1972 per studiare come i parlanti creano, definiscono e riproducono il valore simbolico delle varietà linguistiche che padroneggiano. I loro

lavori hanno aperto la strada all'utilizzo di approcci multi-metodologici nelle ricerche sul bilinguismo, che per indagare gli usi linguistici dei parlanti bilingui hanno iniziato a servirsi sempre più di metodi come l'osservazione dei partecipanti, la registrazione di interazioni e le interviste.

Questi ultimi tre metodi appena citati, sono stati tutti utilizzati per realizzare il lavoro di ricerca che sarà illustrato nei paragrafi successivi. In una prima parte, descriverò brevemente lo studio condotto, i suoi obiettivi e le varie fasi di ricerca, mentre, nella seconda parte, presenterò alcune considerazioni sulla metodologia utilizzata, mettendo in evidenza, in particolare, i vantaggi e le difficoltà emerse durante una ricerca che si è avvalsa di una multi-metodologia per la raccolta di dati linguistici.

2. Metodologia e obiettivi della ricerca

La ricerca ha previsto due fasi di lavoro, entrambe volte a esaminare alcune caratteristiche linguistiche e sociolinguistiche di parlanti bilingui. In entrambe le fasi, gli informatori coinvolti sono stati parlanti bilingui italo-arabofoni, appartenenti a una fascia di età compresa tra i 17 e i 27 anni, e di seconda generazione, in quanto nati in Italia da genitori di madrelingua araba. Pertanto, essendo figli di parlanti madrelingua di arabo, hanno avuto modo di acquisire la lingua araba in ambiente familiare e apprendere, fin da bambini, la lingua italiana, grazie all'inserimento nei contesti scolastici e nei diversi contesti sociali. Si tratta dunque di un bilinguismo precoce che i soggetti considerati hanno mantenuto nel tempo, dal momento che, fino ad oggi, hanno tutti avuto modo di alternare costantemente le due lingue conosciute, con la possibilità di utilizzare la lingua araba sia in ambiente familiare sia mantenendo contatti con il Paese di origine e la lingua italiana nei restanti contesti sociali, tendenzialmente non familiari.

Il metodo scelto per condurre la prima indagine è stato quello delle interviste semi-strutturate, le quali, in parte, hanno previsto domande mirate per ottenere i dati linguistici ricercati e, in parte, erano costitu-

ite da domande aperte e discorsive per lasciare spazio anche alla raccolta di informazioni linguistiche non previste dallo studio (Appendice A). Le interviste hanno coinvolto, singolarmente, 11 parlanti italo-arabofoni e si sono svolte in luoghi pubblici. La lingua in cui sono state somministrate è quella italiana e la durata delle interviste, essendo costituite anche da domande aperte, è risultata piuttosto varia: la più breve ha avuto una durata di venti minuti, mentre l'intervista più lunga è durata circa un'ora. Le interviste sono state registrate e successivamente trascritte e grazie alle informazioni raccolte è stato possibile ricavare i dati linguistici di interesse per questa indagine.

In questa prima fase della ricerca, l'obiettivo, da una parte, è stato quello di indagare il legame tra la lingua e l'identità dei parlanti coinvolti e di identificare eventuali cambiamenti a livello della personalità e nel modo di pensare in relazione alle lingue padroneggiate; dall'altra, è stato quello di valutare la consapevolezza linguistica mostrata dai bilingui coinvolti, ovvero quanto questi fossero consapevoli del proprio bilinguismo, specialmente per quanto riguarda il passaggio da una lingua all'altra e le influenze linguistiche tra le due lingue.

Nella seconda fase di questo studio, sono partita dai risultati emersi dalla prima indagine e da quanto i bilingui avevano riferito durante le interviste della prima indagine riguardo agli usi che fanno delle lingue padroneggiate. Lo scopo è stato quello di osservare alcuni fenomeni linguistici che si possono manifestare nei discorsi bilingui, come l'alternanza linguistica, le scelte linguistiche e i vari fenomeni di contatto linguistico, oltre a riprendere anche l'aspetto della consapevolezza linguistica già osservato durante la prima indagine.

I bilingui che hanno preso parte alla seconda indagine della ricerca sono stati 5 e sono stati individuati tra gli 11 italo-arabofoni che hanno partecipato alla prima indagine. In questo caso, poiché l'indagine non si è limitata all'analisi delle informazioni direttamente riferite dai parlanti bilingui – come è accaduto nella prima indagine – ma ha voluto approfondire anche alcuni aspetti linguistici della lingua parlata dei partecipanti, la metodologia scelta ha richiesto ai partecipanti di realizzare delle registrazioni audio di discorsi bilingui per poterne ricavare dei *corpora* linguisticamente analizzabili. Ai bilingui è stato chiesto

di raccontare un episodio, aneddoto o aspetto inerente a determinati contesti proposti, immaginando di riferirsi a un interlocutore in grado di comprendere sia la lingua araba che la lingua italiana, in modo da poter alternare, il più possibile naturalmente, le due lingue. In totale sono stati raccolti un'ora e dieci minuti di discorsi bilingui, che sono stati trascritti manualmente, realizzando dei *corpora*, la cui analisi ha permesso di studiare – come dallo scopo della ricerca – le scelte linguistiche dei parlanti, le alternanze di codice e fenomeni di contatto linguistico.

In sintesi, gli obiettivi generali della ricerca, quindi di entrambe le indagini, sono stati i seguenti:

1. Riflettere sulla relazione tra lingue e identità nel caso di bilingui italo-arabofoni di seconda generazione, in base alle informazioni fornite dagli intervistati.
2. Valutare la consapevolezza linguistica dei partecipanti relativamente ai loro usi linguistici.
3. Osservare, in termini linguistici e attraverso l'analisi dei *corpora* raccolti, le scelte linguistiche, le alternanze tra le due lingue ed eventuali fenomeni di contatto.

3. Riflessioni sulla metodologia

In questo paragrafo sarà commentata la metodologia utilizzata, mettendo in evidenza anche alcuni dei dati linguistici più interessanti emersi nel corso del lavoro di ricerca grazie alla scelta di un approccio multi-metodologico. In particolare, poiché ogni tecnica di ricerca tende a presentare sia dei punti di forza sia delle limitazioni, e poiché lo stesso si può dire per quanto riguarda l'approccio combinato di cui si è servito questo studio, seguiranno osservazioni sia sui limiti dovuti alla scelta di questo tipo di metodologia, sia sui punti di interesse emersi proprio grazie all'utilizzo di un approccio simile.

3.1 Difficoltà e limiti

Come sopra accennato il metodo di ricerca scelto non ha portato soltanto a ottenere dei dati utili ai fini dello studio condotto, ma ha anche mostrato dei limiti che si sono manifestati sotto forma di difficoltà riscontrate durante il lavoro di ricerca.

A questo proposito, la difficoltà principale ha riguardato la gestione della *varietà* dei dati linguistici ricavati dalle risposte ricevute durante le interviste e dall'analisi dei discorsi bilingui. In entrambe le fasi della ricerca, infatti, la varietà dei dati linguistici raccolti è stata evidente. Per quanto riguarda i dati ricavati dalle interviste, questa varietà è stata frutto delle dichiarazioni soggettive dei bilingui coinvolti, specie quando le domande riguardavano l'aspetto identitario e della consapevolezza linguistica (Appendice A). Inoltre, anche nei dati ricavati dalla domanda relativa alla lingua utilizzata nei vari contesti proposti la varietà delle informazioni ricavate è stata evidente, in quanto ogni partecipante ha fornito una propria combinazione delle lingue che utilizza nei vari contesti che si differenziava da quella degli altri partecipanti. Allo stesso modo, durante l'analisi dei *corpora*, la disomogeneità ha riguardato le scelte linguistiche compiute dai parlanti, ma anche le tipologie di alternanze linguistiche e di code-switching che si sono manifestate. Pertanto, non è stato immediato, né lineare, categorizzare e interpretare i dati ottenuti attraverso la combinazione di due metodi qualitativi in una ricerca che si è occupata anche di aspetti legati alla sfera soggettiva dei parlanti bilingui. Di conseguenza, per categorizzare i dati e ottenere dei risultati – per quanto possibile – omogenei, è stato necessario condurre un'accurata analisi, non solo dei singoli dati, ma anche di una loro comparazione e dei contesti e degli atteggiamenti dei bilingui coinvolti nello studio.

3.2. Punti di interesse e strumenti della ricerca

Probabilmente, la difficoltà dovuta alla varietà dei dati ottenuti, descritta nel sottoparagrafo precedente, potrebbe essere stata meno evidente utilizzando un approccio non combinato o facendo uso di

metodi quantitativi in ricerche di questo tipo. Tuttavia, la scelta di un approccio di ricerca qualitativo, in uno studio sul bilinguismo interessato ad aspetti come la consapevolezza e l'identità linguistica dei parlanti, sembra essere stata la scelta più adeguata, in quanto ha permesso di osservare anche aspetti comportamentali e sociolinguistici dei bilingui presi in considerazione. Osservazioni di questo genere, infatti, non sarebbero state possibili scegliendo un approccio quantitativo. Inoltre, il metodo qualitativo ha permesso, allo stesso tempo, di raccogliere anche alcuni dati quantitativi utili, che generalmente sono raccolti attraverso i metodi quantitativi.

In particolare, nella prima fase della ricerca – ovvero quella delle interviste – chiedendo ai partecipanti di raccontare la loro biografia linguistica, sono state raccolte informazioni precise sui bilingui coinvolti, come l'età dei partecipanti, le lingue conosciute, l'età di esposizione alle due lingue e il paese di origine dei genitori in quanto parlanti di madrelingua araba. È stato, così, possibile realizzare dei grafici di carattere quantitativo come il seguente, ricavato dalla domanda che chiedeva ai partecipanti quale lingua utilizzassero in questi determinati contesti.

Come è possibile notare dal grafico (Figura 1) l'approccio metodologico scelto ha permesso di ricavare ugualmente dati quantitativi e percentuali precise riguardo all'uso che i bilingui fanno delle due lingue. In altri termini, la scelta del metodo delle interviste, oltre a dare la possibilità di ottenere dati quantitativi simili a quelli sopra rappresentati, che in genere si ottengono con la somministrazione di questionari, ha permesso di integrarli con ulteriori informazioni linguistiche, ricavate sia dal contenuto delle risposte date dagli intervistati, sia dall'osservazione diretta dei comportamenti dei partecipanti. Inoltre, l'osservazione diretta dei partecipanti è stata particolarmente utile per indagare l'aspetto della consapevolezza linguistica. Quest'ultima non è stata valutata solamente attraverso quanto i bilingui hanno riferito esplicitamente a parole, ma anche attraverso altre caratteristiche comportamentali, come le esitazioni, le ripetizioni, le pause e i silenzi, che si sono manifestate mentre i partecipanti rispondevano alle domande poste. Per esempio, il fatto che le risposte date alla domanda: "Utilizzi

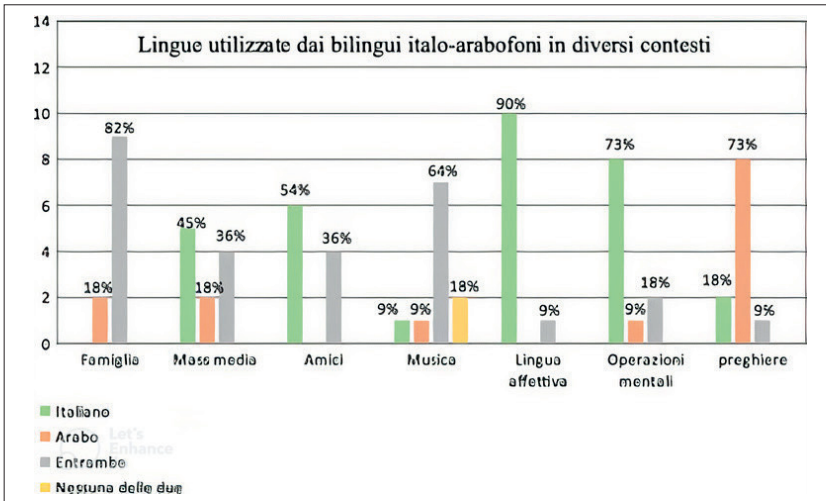


Figura 1 Grafico riportante in che misura i partecipanti italo-arabofoni coinvolti (asse delle ordinate=numero di partecipanti) utilizzano le lingue padroneggiate nei contesti proposti (asse delle ascisse)

maggiormente la lingua italiana o la lingua araba?” spesso non siano state immediate ma abbiano richiesto un tempo riflessione, o abbiano comportato delle esitazioni da parte degli informatori, è stato interpretato come segno di bassa consapevolezza linguistica.

Un ulteriore strumento interessante, legato all'utilizzo delle interviste, è quello delle domande aperte, le quali hanno permesso agli intervistati di rispondere liberamente e di raccontare talvolta anche aneddoti personali riguardo ai loro usi linguistici, fornendo quindi anche nuovi spunti utili per fare riflessioni ulteriori che inizialmente non erano previste dal lavoro di ricerca. Inoltre, è stato possibile spiegare le domande e il loro significato, chiarire eventuali fraintendimenti e richiedere esempi, se necessario. In particolare, nel caso di una delle domande riguardanti le alternanze linguistiche – ovvero quella che chiedeva ai partecipanti se alternassero le due lingue solo all'interno della stessa frase oppure anche all'interno della stessa parola – alcuni intervistati hanno chiesto chiarimenti su come interpretare la doman-

da posta. Questi ultimi, dopo aver inizialmente riferito di non alternare le due lingue all'interno di un unico termine, hanno poi risposto positivamente alla domanda e suggerito degli esempi su questa tipologia di contatto linguistico, di cui solitamente fanno uso. Infine, grazie a questo metodo, si è creata una relazione con i partecipanti che, in alcuni casi, li ha sollecitati a fornire informazioni e dati linguistici anche successivamente al momento dell'intervista. Per quanto riguarda l'aspetto della consapevolezza linguistica, per esempio, successivamente alle interviste, alcuni bilingui hanno specificato che pur riflettendo per giorni su alcune delle domande poste risultava comunque difficile stabilire in modo definitivo quando e quanto utilizzano una lingua anziché l'altra, determinare il momento del passaggio dall'una all'altra, nonché cercare a capire con quale lingua pensano.

In merito alla seconda parte della ricerca, invece, la metodologia adoperata – che, come già anticipatamente descritto nel § 2, è consistita nella realizzazione di *corpora* ricavati dalle registrazioni audio di discorsi bilingui – è stata proficua nei seguenti termini.

In primo luogo, ascoltando le registrazioni audio è stato possibile individuare le scelte linguistiche operate dai bilingui per parlare dei contesti proposti e conseguentemente realizzare, per ogni partecipante, una tabella come la seguente rappresentante le varie scelte linguistiche (tabella 1).

	Lingua di base	Presenza di elementi dell'altra lingua
Familiare/casalingo	Entrambe	
Universitario/lavorativo	Italiano	Sì
Mass media e social media	Italiano	Sì
Operazioni mentali	Italiano	Sì
Tempo libero	Italiano	Sì
Religioso/preghiere	Italiano	Sì
Amici	Entrambe	
Emozioni	Entrambe	

Tabella 1 Esempio di tabella riportante le scelte linguistiche adottate durante le registrazioni raccolte

Pertanto, anche in questo caso, il metodo scelto ha permesso di raccogliere dati di tipo quantitativo.

In secondo luogo, l'aver richiesto, in questa seconda fase, delle registrazioni di lingua parlata a ognuno dei partecipanti che avevano preso parte alla fase delle interviste – per poi ricavarne dei *corpora* linguisticamente analizzabili – ha permesso di individuare e analizzare le commutazioni di codice osservando direttamente i loro usi linguistici quotidiani e non soltanto attraverso i dati che avevano riferito durante le interviste. Quello che segue è uno degli esempi di alternanza linguistica estratti dai *corpus* realizzati.

Esempio 1

Ana ma 'raftš fin glsti nti, comunque la prima, quella dove c'è la lampada e dove ci sono *stilūāt* [...].

L'enunciato appena riportato rappresenta uno dei vari dati linguistici, ricavati dai *corpora*, che “completano” le informazioni fornite dagli intervistati. L'esempio, infatti, mette in evidenza – proprio come avevano riferito i partecipanti – come le commutazioni di cui essi fanno uso si possano manifestare sia a livello interfrasale, sia a livello intrafrasale. In altri termini, le alternanze linguistiche, nei dati raccolti, si sono presentate sia come singole parole pronunciate in una lingua all'interno di una proposizione pronunciata nell'altra lingua – come la parola *stilūāt* nell'es. 1 – sia come sequenze conversazionali più ampie come accade nella prima parte dell'enunciato dove l'intera proposizione iniziale è stata pronunciata in lingua araba. Un'ulteriore tipologia di *code-switching* che i bilingui hanno riferito di utilizzare, e che è stata successivamente riscontrata analizzando i *corpora*, è quella che consiste nella ripetizione dello stesso contenuto in entrambe le lingue utilizzate dai partecipanti, con lo scopo di risultare più chiari o di enfatizzare un determinato contenuto (cfr. MOUSSAID 2023). Infine, la scelta di questa metodologia, basata sull'analisi di *corpora* bilingui, è risultata efficace anche per proseguire con l'approfondimento della consapevolezza linguistica dei bilingui affrontato nella prima fase della ricerca attraverso le interviste e l'osservazione diretta degli atteggiamenti dei

bilingui. Più specificatamente, in questa seconda parte della ricerca, la consapevolezza linguistica è stata indagata attraverso la comparazione dei dati forniti dagli intervistati, relativamente alle lingue che hanno riferito di utilizzare nei vari contesti quotidiani (vedi grafico figura 1), con i dati ricavati dai *corpora* rispetto alle lingue che hanno effettivamente utilizzato nei loro discorsi per parlare di quegli stessi contesti.

Conclusioni

Qualsiasi indagine o ricerca quando giunge al termine lascia spazio a nuovi interrogativi, non solo in termini di nuove domande scientifiche a cui cercare di dare risposte, ma anche in termini metodologici per cercare di migliorare la qualità dei risultati delle future ricerche. Nel caso della ricerca illustrata finora, l'approccio combinato e i due metodi qualitativi scelti per condurre lo studio sembrano essere stati i più adatti per la raccolta e l'analisi dei dati in questo tipo di ricerca su parlanti bilingui. Infatti, hanno permesso di dare delle risposte alle domande su cui vertevano le due indagini, non solo analizzando singolarmente i dati ottenuti da ognuna delle metodologie utilizzate, ma anche confrontando le informazioni linguistiche ottenute dagli informatori con l'analisi linguistica dei dati ricavati dai *corpora* di lingua parlata realizzati. Tuttavia, come sottolineato nel § 3.1, l'utilizzo di un approccio combinato ha avuto anche dei limiti e presentato delle difficoltà, come quella della disomogeneità dei dati raccolti. Per quanto riguarda la metodologia utilizzata, infatti, uno dei principali interrogativi sorti durante il lavoro di ricerca riguarda proprio la gestione delle differenze che hanno caratterizzato i dati raccolti. Tali differenze corrispondono all'esigenza di trovare il metodo più adatto per analizzare e categorizzare dati linguistici come quelli qui esaminati, nonostante la loro varietà e nonostante il coinvolgimento della sfera soggettiva dei bilingui considerati, come è accaduto in questa ricerca che ha indagato aspetti come l'identità linguistica e la consapevolezza linguistica nell'ambito del bilinguismo.

Appendice A

Domande delle interviste

Parte i – Identità linguistica

1. Come ti chiami e come pronunci il tuo nome quando qualcuno ti chiede come ti chiami? In italiano o in arabo?
2. Raccontami la tua biografia linguistica.
3. Di quale cultura senti maggiormente di appartenere? (italiana o araba)
4. Ti senti una persona diversa quando utilizzi una lingua anziché l'altra? Perché?
5. Cambia il tuo modo di pensare in base alla lingua utilizzata?
6. Ci sono cambiamenti nel tuo modo di atteggiarti e comportarti quando parli in italiano rispetto all'arabo o viceversa?
7. Utilizzi atteggiamenti e comportamenti tipicamente italiani quando parli in arabo o viceversa?
8. Dimmi un vantaggio e una difficoltà nel crescere tra due mondi così differenti?
9. Pensa a un posto in cui ti senti a casa... Quale lingua si parla in questo posto?
10. "Parlare due lingue è come vivere due vite". Sei d'accordo? Per te è così?

Parte ii – Consapevolezza linguistica

11. Parli più arabo o più italiano?
12. Quando passi da una lingua all'altra, ci fai caso? Lo fai spesso?
13. Ti capita di mescolare le due lingue? Solo all'interno della stessa frase oppure anche all'interno della stessa parola? Mi sai fare qualche esempio?

Yasmina Moussaid

14. Ti faccio riferimento ad alcuni contesti e mi dici quale lingua utilizzi:
- Famiglia
 - Televisione e musica
 - Amici
 - Lingua affettiva (sentimenti): in quale lingua esprimi le tue emozioni e i tuoi sentimenti più profondi? Se sei arrabbiato quale lingua utilizzi?
 - Operazioni mentali e calcoli matematici
 - Preghiere
15. Pensi che parlare una seconda lingua sia un vantaggio oppure un inconveniente? Da quale punto di vista? Dimmi un'occasione in cui conoscere due lingue è un notevole beneficio. Perché?
16. In quale lingua pensi?
17. In quale lingua sogni?
18. Ti capita di parlare da solo/a? Parli più arabo o più italiano?

Bibliografia

- FISHMAN 1967 = JOSHUA A. FISHMAN, *Bilingualism with and without diglossia, diglossia with and without bilingualism*, In «The Journal of Social Issues», 23, 1967.
- GUMPERZ 1982 = JOHN GUMPERZ, *Discourse strategies*, Cambridge, Cambridge University Press, 1982.
- MOUSSAID 2023 = YASMINA MOUSSAID, *Il contatto linguistico tra identità e consapevolezza: uno studio su bilingui italo-arabofoni*, in «DILEF. Rivista digitale del Dipartimento di Lettere e Filosofia», 2, 2023, pp. 1-18.

Riassunto Indipendentemente dal tipo di osservazione che si vuole condurre, o dalla domanda scientifica a cui si cerca di dare risposta, le ricerche linguistiche comprendono

La raccolta e l'analisi dei dati linguistici

una fase di analisi dei dati, che ha inizio con la raccolta dei dati linguistici, prevede una loro elaborazione e termina con la presentazione dei risultati finali dell'analisi condotta. Ciò accade anche nell'ambito delle ricerche linguistiche condotte in contesti bilingui, o multilingui, come quella qui osservata, la quale è stata condotta con lo scopo di raccogliere e analizzare informazioni linguistiche legate al bilinguismo di parlanti italo-arabofoni. Le osservazioni riportate nei seguenti paragrafi vogliono riflettere sulle fasi di trattamento dei dati linguistici di tale ricerca, che ha combinato l'analisi delle informazioni ricavate da interviste semi-strutturate – rivolte a parlanti bilingui – con l'analisi di *corpora* bilingui, con l'intento di indagare alcune caratteristiche principali del bilinguismo, come le scelte linguistiche dei parlanti coinvolti, i fenomeni di contatto, e ulteriori questioni di tipo sociolinguistico e socioculturale, come l'identità e la consapevolezza linguistica.

Abstract Regardless of the type of observation a linguist wishes to conduct and regardless of which scientific question he seeks to answer, linguistic research includes different stages of data analysis, such as collecting data, processing them, and finally drawing conclusions based on the findings obtained. These phases are equally involved when the research studies are conducted in bilingual or multilingual contexts, as happened in the research study proposed in this paper, which was performed aiming to collect and analyse linguistic data related to the bilingualism of Italian-Arabic speakers. The following paragraphs discuss and reflect on the different stages needed to process the linguistic data of this study, which combined two different qualitative methods. In particular, the study integrated the analysis of linguistic information obtained from semi-structured interviews with the analysis of bilingual *corpora*. The purpose of the research was to investigate some features of bilingualism, such as the linguistic choices made by the speakers involved, language contact phenomena, and some additional sociolinguistic and sociocultural issues, as linguistic identity and language awareness.

